Estimating portfolio loss distribution by Monte Carlo: an Interacting Particle System approach

Stéphane Crépey*, **R.** Carmona

Preliminary Version
May 2008

Contents

*The research of S. Crépey benefited from the support of the Europlace Institute of Finance and the "Chair Risque de crédit" (Fédération Bancaire Française).

Introduction

We want to compute by simulation credit portfolio loss distributions (at fixed points in time, e.g., $T = 5y$, in view for instance of pricing CDO tranches by simulation). For the high levels of the loss distribution we are thus facing a rare events simulation problem (many not so rare accounting for contagion effects, but yet).

In factor copulae models the conditional loss distribution can be recovered analytically by various means, and the (unconditional) distribution follows by numerical quadrature. Alternatively, from the Monte Carlo point of view, it is possible to apply various variance reduction techniques to the simulation of the conditional distribution, or the unconditional distribution can be sampled directly (but slowly).
For pricing CDO tranches by Monte Carlo the most commonly used Monte Carlo variance reduction technique is control variate (using the portfolio loss at maturity as control variable).

In more general dynamic models of credit risk the loss distribution can be computed by numerical resolution of the related forward Kolmogorov equations. However practical resolution by deterministic numerical schemes is precluded by the curse of dimensionality for models of (Markovian) dimension d greater than a few units. So for high d (as in general bottom-up models) simulation approaches are the only way to go.

But, given the complexity and variety of models at hand, it is not easy to devise generic and efficient variance reduction schemes. In this work we consider application to this problem of the so-called Interacting Particle System approach (Pierre Del Moral and Josselin Garnier, *Genealogical particle analysis of rare events* [4], Annals of Applied Probability, 2005; see also the monograph [3]).

Application of IPS to the Computation of CDO Tranche Spreads was already considered in Douglas Vestal, René Carmona and Jean-Pierre Fouque, *Interacting Particle Systems for the with Rare Defaults* (paper under submission), yet in a *structural model* of credit risk. Here we consider application of IPS to computation of a credit portfolio loss distribution in *reduced-form intensity models* of credit risk.

1 Interacting Particle Systems for the Computation of Rare Events

The purpose of this review section is to give a crash course on Feynman-Kac path measures and their subsequent interacting particle system (IPS for short) interpretation. The basic material is borrowed from [3] and the actual application to the Monte Carlo computation of probabilities of rare events from [4]. Technical details are included for the sake of completeness and to ease the introduction of a specific set of notations. The IPS method of [4] was already used in the computation of CDO tranche spreads based on structural models in [10]. The point of view of the present paper is different: we use a reduced-form approach, and for this reason, the dimension and the complexity of the IPS which we bring to bear in the computation of small probabilities are significantly smaller, thus streamlining the

application to CDO valuation.

1.1 Twisted Feynman-Kac Expectations

The problem at hand is the computation of small probabilities of events relative to a (possibly time inhomogeneous) Markov chain $\{X_n\}_n$, X_n being a random element taking values in a general measurable space $(E_n, \mathcal{E}_n)$ which can change with n. Roughly speaking, the method proposed in [4] is based on the deformation of the Markov chain successive transitions by way of mutations and selections in order to force the chain into the rare events of interest. This strategy is reminiscent of classical importance sampling. However, the main difference is that while the Monte Carlo sample of an importance sampling computation are generated from the *twisted* distribution, the Monte Carlo samples used in an IPS Monte Carlo computation are generated under the original distribution of the chain. In other words, the knowledge of the distribution of the underlying Markov chain is not really necessary. As we are about to see, all we need to have in order to implement the IPS Monte Carlo computations is a *black box* capable of generating Monte Carlo samples from the distribution of the chain.

We denote by $K_n(x_{n-1}, dx_n)$ the transition kernel of the underlying Markov chain at time n, and we denote by $\{Y_n\}_n$ the historical process of $\{X_n\}_n$ defined by:

$$Y_n = (X_0, \cdots, X_n) \in F_n = E_0 \times \cdots \times E_n.$$

Now, let $M_n(y_{n-1}, dy_n)$ denote the Markov transitions associated with the inhomogeneous Markov chain $\{Y_n\}_n$. We will also use the generic notation $\mathcal{B}_b(E)$ for the space of bounded, measurable functions on the measurable space $(E, \mathcal{E})$ equipped with the uniform norm. Finally, for each integer $n \geq 0$ we consider non-negative measurable functions G_n on F_n equipped with the product σ-field, and we interpret these functions as potential functions. It is assumed in [4] that these potential functions are bounded and bounded away from zero in the sense that:

$$\sup_{(y_n, y'_n) \in F_n \times F_n} \frac{G_n(y_n)}{G_n(y'_n)} < \infty. \tag{1}$$

This restrictive boundedness assumption can be relaxed in many cases. Here, we only mention it for the sake of completeness.

Then, for any given $f_n \in \mathcal{B}_b(F_n)$, we define the Feynman-Kac expectation $\gamma_n(f_n)$ by:

$$\gamma_n(f_n) = \mathbb{E}\left\{ f(Y_n) \prod_{1 \leq i < n} G_i(Y_i) \right\}. \tag{2}$$

Note that as a non-negative linear form on a cone of non-negative functions, γ_n can be viewed as a measure. We shall denote by η_n the corresponding normalized measure which is naturally defined as

$$\eta_n(f_n) = \frac{\mathbb{E}\left\{ f(Y_n) \prod_{1 \leq i < n} G_i(Y_i) \right\}}{\mathbb{E}\left\{ \prod_{1 \leq i < n} G_i(Y_i) \right\}} = \frac{\gamma_n(f_n)}{\gamma_n(1)}. \tag{3}$$

Notice that

$$\gamma_{n+1}(1) = \gamma_n(G_n) = \eta_n(G_n)\gamma_n(1) = \prod_{i=1}^{n} \eta_i(G_i). \tag{4}$$

This seemingly innocent remark will play a crucial role in the following. Consequently, for any given bounded measurable function f_n, we have

$$\gamma_n(f_n) = \eta_n(f_n) \prod_{1 \leq i < n} \eta_i(G_i).$$

The above relationship has the merit of relating the un-normalized expectations in the left hand side to normalized *twisted* expectations in the right hand side. We will find it convenient later on to use the potential functions G_n^- defined by

$$G_n^- = \frac{1}{G_n}$$

(who incidently still satisfy the positive boundedness condition if the orginal potential functions G_n do), and the associated normalized and non-normalized measures which we denote by η_n^- and γ_n^- respectively. Given all these definitions and notation we easily get:

$$\begin{aligned} \mathbb{E}\{f_n(Y_n)\} &= \mathbb{E}\left\{ f_n(Y_n) \prod_{1 \leq i < n} G_i^-(Y_i) \prod_{1 \leq i < n} G_i(Y_i) \right\} \\ &= \gamma_n \left(f_n \prod_{1 \leq i < n} G_i^- \right) \\ &= \eta_n \left(f_n \prod_{1 \leq i < n} G_i^- \right) \prod_{1 \leq i < n} \eta_i(G_i). \end{aligned} \tag{5}$$

This shows that expectations over the original process can be computed if one can compute normalized twisted expectations. This is in fact possible, in a dynamic way because, like in classical filtering theory, it is easily checked that, the sequence of normalized twisted probability measures form a well defined *dynamical system* in the space of probability measures. Indeed we have

$$\eta_n = \Phi_n(\eta_{n-1}), \qquad \eta_1 = M_1(X_0, \cdot) \tag{6}$$

where the nonlinear operators Φ_n giving the dynamics are defined as

$$\Phi_n(\eta) = \frac{1}{\eta(G_{n-1})} \int_{F_{n-1}} \eta(dy_{n-1}) G_{n-1}(y_{n-1}) M_n(y_{n-1}, \cdot) \tag{7}$$

1.2 IPS and Empirical Estimations

1.2.1 Model Simulation

For the purpose of numerical computations and Monte Carlo estimation of expectations of the form (**??**), we introduce approximations of the above probability distributions by convex combinations of Dirac measures, and we show that the time evolution of the measures η_n given by the dynamical system (**??**) and (**??**) implies a natural time evolution for the point masses of the convex combinations of Dirac measures, hence the interpretation of these approximations as an interacting particles system.

We choose an integer m which we shall interpret as the number of particles. So clearly, we can think of m as being large. A particle at time n is an element

$$\xi_n^j = (\xi_{0,n}^j, \xi_{1,n}^j, \cdots, \xi_{n,n}^j) \in F_n = E_0 \times E_1 \times \cdots \times E_n.$$

where the superscript j of the particle ranges from 1 to m. We shall start with an initial configuration $\xi_1 = (\xi_1^j)_{1 \le j \le m}$ that consists of m independent and identically distributed random samples from the distribution:

$$\eta_1(d(x_0, x_1)) = M_1(X_0, d(x_0, x_1)) = \delta_{X_0}(dx_0) K_1(x_0, dx_1)$$

where we use the notation δ_x for the Dirac measure at the point x. In other words, the $\xi_1^j = (\xi_{0,1}^j, \xi_{1,1}^j) = (X_0, \xi_{1,1}^j) \in F_1 = E_0 \times E_1$ are independent and in such a way that all the $\xi_{1,1}^j$ are all independent with the same distribution $K_1(X_0, \cdot)$. Based on the transition given by the dynamic equation (**??**), we define the transition for the particles which are providing the approximation of η_2 by $\xi_{n-1} \to \xi_n$ from F_{n-1}^m into F_n^m according to the transition probability

$$\mathbb{P}\{\xi_n \in d(y_n^1, \cdots, y_n^m) \,|\, \xi_{n-1}\} = \prod_{j=1}^m \Phi_n(\mu(\xi_{n-1}))(dy_n^j), \tag{8}$$

where $\rho(\xi_{n-1})$ is the empirical measure defined by

$$\rho(\xi_{n-1}) = \frac{1}{m} \sum_{j=1}^m \delta_{\xi_{n-1}^j}$$

and $d(y_n^1, \cdots, y_n^m)$ is an infinitesimal neighborhood of the point $(y_n^1, \cdots, y_n^m) \in F_n^m$. Recalling the definition of the operators Φ_n giving the dynamics of the η_n, one can see that (**??**) is the superposition of two clearly identifiable elementary transitions, a *selection* followed by a *mutation*. In other words:

$$F_{n-1}^m \ni \xi_{n-1} \overset{\text{selection}}{\longrightarrow} \widehat{\xi}_{n-1} \in F_{n-1}^m \overset{\text{mutation}}{\longrightarrow} \xi_n \in F_n^m$$

as follows. The selection stage is performed by resampling with replacement (i.e. choosing independently) m (path) particles

$$\widehat{\xi}_{n-1}^j = (\widehat{\xi}_{0,n-1}^j, \widehat{\xi}_{1,n-1}^j, \cdots, \widehat{\xi}_{n-1,n-1}^j) \in F_{n-1},$$

with possible repetitions according to the Gibbs measure

$$\sum_{j=1}^m \frac{G_{n-1}(\xi_{0,n-1}^j, \xi_{1,n-1}^j, \cdots, \xi_{n-1,n-1}^j)}{\sum_{j=1}^m G_{n-1}(\xi_{0,n-1}^j, \xi_{1,n-1}^j, \cdots, \xi_{n-1,n-1}^j)} \delta_{(\xi_{0,n-1}^j, \xi_{1,n-1}^j, \cdots, \xi_{n-1,n-1}^j)} \tag{9}$$

Then, the mutation stage of the transition is performed by extending each selected (path) particle $\widehat{\xi}_{n-1}^j$ into a (path) particle $\xi_n^j \in F_n = F_{n-1} \times E_n$ of the form

$$\begin{aligned} \xi_n^j &= (\xi_{0,n}^j, \xi_{1,n}^j, \cdots, \xi_{n,n}^j) \\ &= (\widehat{\xi}_{0,n-1}^j, \xi_{1,n-1}^j, \cdots, \widehat{\xi}_{n-1,n-1}^j, \xi_{n,n}^j) \end{aligned}$$

where the m samples $\xi_{n,n}^j$ are independently drawn from the distributions $K_n(\widehat{\xi}_{n-1,n-1}^j, \cdot)$.

1.2.2 Convergence Results

We are now in a position to quote the theoretical result on which the Monte Carlo approximations are based. See for example [3] or [4] for details. For each fixed n we have

$$\lim_{m\to\infty} \eta_n^m = \eta_n$$

in distribution, where the empirical measures η_n^m are defined by:

$$\eta_n^m = \frac{1}{m}\sum_{j=1}^{m} \delta_{(\xi_{0,n}^j, \xi_{1,n}^j, \cdots, \xi_{n,n}^j)}.$$

This result is screaming for the introduction of the particle approximation

$$\gamma_n^m(f_n) = \eta_n^m(f_n) \prod_{1\le i<n} \eta_n^m(G_i)$$

for $\gamma_n(f_n)$. The main result of [4] which we use below states that γ_n^m is an unbiased estimator for γ_n in the sense that for any integer $p \ge 1$ and $f_n \in \mathcal{B}_b(F_n)$ with $\|f_n\| \le 1$, we have

$$\mathbb{E}\{\gamma_n^m(f_n)\} = \gamma_n(f_n),$$

and in addition

$$\sup_{m\ge 1} \sqrt{m}\mathbb{E}\{|\gamma_n^m(f_n) - \gamma_n(f_n)|^p\}^{1/p} \le c_p(n),$$

for some positive constant $c_p(n) < \infty$ whose value does not depend upon the particular choice of the function f_n.

In view of (**??**), we thus get the following unbiased (at fixed m) and asymptotically convergent (as $m \to \infty$) estimate to $\mathbb{E}\{f_n(Y_n)\}$:

$$\gamma_n^m(f_n \prod_{1\le i<n} G_i^-) = \eta_n^m(f_n \prod_{1\le i<n} G_i^-) \prod_{1\le i<n} \eta_i^m(G_i) \tag{10}$$

$$= \mathbb{E}_n^m \left\{ f_n(\xi_n) \prod_{1\le i<n} G_i^-(\xi_{0,n}, \cdots, \xi_{i,n}) \right\} \prod_{1\le i<n} \mathbb{E}_i^m G_i(\xi_i) \tag{11}$$

where for every i the notation $\mathbb{E}_i^m$ refers to expectation under the empirical distribution defined by the ξ_i^j's. More importantly, the variance of the estimator can be analyzed. In case the expectation of interest is tantamount to the probability of a "rare" event

$$\mathbb{E}\{f_n(Y_n)\} = \mathbb{E}\{\mathbb{1}_A(V(Y_n))\} = \mathbb{P}\{V(Y_n) \in A\}\ ,$$

where V is a function from E_n to $\mathbb{R}$, the conclusion is that in order to minimize this variance, one should use weight functions G favoring the occurence of the rare event without involving too large normalizing constants. Moreover, the choice of G should give rise to an algorithm that can be easily implemented.

We shall use these guidelines below for rare events of the form $V(X_n) \in [x, x+\delta x)$ with large x and small δx. In this case the asymptotic variance of the IPS estimator can in turn be estimated, at least, whenever the distribution of $V(X_n)$ given X_i for $i < n$ admits a density with respect to the Lebesgue measure. In this case we thus end-up with a confidence interval.

2 Credit Portfolio Loss Process

We consider a continuous-time d-variate Markov Chain $\widetilde{X} = (\widetilde{X}^1, \cdots, \widetilde{X}^d)$ with components in $\{0, 1, \cdots, \nu\}$, for some fixed integer ν. For simplicity we preclude simultaneous jumps of the $\widetilde{X}^l$'s. The simulation of a d-dimensional model with generator (matrix) of dimension $(\nu+1)^d$ may thus be reduced to the coupled simulation of d one-dimensional models with generator of dimension $\nu + 1$ (see Bielecki et al. [1]).

We denote by $\widetilde{t}_i$ the i^{th} jump time of $\widetilde{X}$. We also set $\widetilde{t}_0 = 0$. We model the cumulative default process L on a credit portfolio of size P as $L_t = V(\widetilde{X}_t)$, for some (integer-valued) *loss function* V. So $L_0 = 0$ and L jumps by some (integer) amount at some of the $\widetilde{t}_i$'s.

Note that many dynamic models of credit risk can be cast into this framework (after discretisation in space, if there are any space-continuous processes involved). Typical examples will be given in later sections.

Since we are interested in the time horizon $[0, T]$, we introduce (using the convention $\inf \emptyset = -\infty$)

$$n = \inf\{i \in \mathbb{N}\,;\, \widetilde{t}_i \geq T\}\ .$$

We assume that n is finite almost surely and we set further, for $i \in \mathbb{N}$,

$$t_i = \widetilde{t}_{i \wedge n} \wedge T_i^X = \widetilde{X}_{t_i}\ .$$

So $t_i = T$ iff $i \geq n$, and for $i \geq n$ we have that $X_i = \widetilde{X}_T$ and $V(X_i) = L_T$.

2.1 IPS Algorithm

Introducing the weight function (for some fixed $\alpha > 0$)

$$G_i^\alpha(y_i) = \exp(\alpha(V(x_i) - V(x_{i-1}))) = G^\alpha(x_{i-1}, x_i)\ ,$$

we propose the following IPS algorithm (in which it is enough to keep track of jump times and of the related two last components "father and son" of each path-particle, given this choice of G) for computing the loss distribution $p_l(T) = \mathbb{P}(L_T = l)$, for $l \in \mathbb{N}_P$.

Initialization: For every $j = 1, \cdots, m$, set $\xi_{0,1}^j = \widetilde{X}_0$ and simulate a pair $(t_1^j, \xi_{1,1}^j)$ starting from $\xi_{0,1}^j$ at time 0 as defined above, using the dynamics of $\widetilde{X}$ for this simulation step.

Loop: Assuming the m (time, father and son)-particles $(t_{i-1}^j, \xi_{i-2,i-1}^j, \xi_{i-1,i-1}^j)$ already simulated:

Selection: Sample independently m (father and son)-particles $(\widehat{t}_{i-1}^j, \widehat{\xi}_{i-2,i-1}^j, \widehat{\xi}_{i-1,i-1}^j)$ with possible repetitions according to the Gibbs measure defined by the $G^\alpha(\xi_{i-2,i-1}^j, \xi_{i-1,i-1}^j)\delta_{t_{i-1}^j, \xi_{i-2,i-1}^j, \xi_{i-1,}^j}$ (normalized to one);

Mutation: For every $j = 1, \cdots, m$, set $\xi_{i-1,i}^j = \widehat{\xi}_{i-1,i-1}^j$ and simulate a pair $(t_i^j, \xi_{i,i}^j)$ starting from $(\widehat{t}_{i-1}^j, \widehat{\xi}_{i-1,i-1}^j)$ using the dynamics of $\widetilde{X}$ for this simulation step.

Termination: Exit from the loop when the (random) time n is reached on every path-particle, for $i = \bar{n}$, say, and compute the following estimate of the loss probability $p_l(T) = \mathbb{P}(L_T = l)$ and of the related asymptotic variance $\sigma_l^2(T)$, for every $l \in \mathbb{N}_P$ (note that $V(\xi_{i,i}^j) = L_T^j$ on $i \geq n_i^j$) :

$$p_l^m(T) = \mathbb{E}_{\bar{n}}^m \left\{\delta_l(V(\xi_{\bar{n},\bar{n}})) \exp(-\alpha V(\xi_{\bar{n}-1,\bar{n}}))\right\} \prod_{i=1}^{\bar{n}-1} \mathbb{E}_i^m G^\alpha(\xi_{i-1,i}, \xi_{i,i}) \tag{12}$$

$$\sigma_l^{2,m}(T) = \mathbb{E}_{\bar{n}}^m \left\{\delta_l(V(\xi_{\bar{n},\bar{n}})) \exp(-2\alpha V(\xi_{\bar{n}-1,\bar{n}}))\right\} \left(\prod_{i=1}^{\bar{n}-1} \mathbb{E}_i^m G^\alpha(\xi_{i-1,i}, \xi_{i,i})\right)^2 \tag{13}$$

where for every fixed i the notation $\mathbb{E}_i^m$ refers to the empirical distribution defined by the ξ_i^j's as j ranges from 1 to m.

Remark 2.1 **(i)** In this simulation it is crucial to exploit the fact that, according to our assumptions, the components of the process $\widetilde{X}$ do not jump simultaneously. So the simulation of $\widetilde{X}$ may be done "component by component"(see Bielecki et al. [1]), in time $O(d)$. **(ii)** We are not in the standard conditions ensuring convergence of the standard deviation estimate (cf. end of section 1). However as it will be apparent in later sections it seems that this estimate works in practice for assessing significance of the related probability estimate (but it should be considered with caution as far as producing confidence intervals is concerned).

2.2 Poisson Toy Model

We first consider application of the method to the simulation of the loss distribution at $T = 5y$ of a standard Poisson process N (stopped at level $n = 125$), so (for $i < 125$) $\mathbb{P}(N_T = i) = e^{-5}\frac{5^i}{i!}$.

We ran 11 Monte Carlo loops (one standard MC loops and $\mu = 10$ IPS MC loops) of size $m = 10^5$ each, yielding eleven different estimates of the probabilities $p_i(T) = \mathbb{P}(N_T = i)$, for $i \in \mathbb{N}_n$. Then, for every i, we retained the estimator of the related probability with the highest *significance*, in the sense of the highest ratio of the estimated probability relative to the estimated standard deviation (which in our experience is practically equivalent to choosing the estimator corresponding to the value $\alpha(i)$ having given rise to the greatest number of trajectories at level i).

Table 2 displays the results obtained for the 35 first levels of the loss ($i = 0, \cdots, 34$). For higher levels of the loss, the related probabilities are too small and the generic IPS methodology is not sufficient to provide reasonable estimates, more specifically problem-dependent methodologies should be considered instead (see, e.g., Johansen, Del Moral and Doucet [8]). In the fourth column, $\alpha(i)$ (to be understood as $0.2 \times \alpha(i)$) refers to the best value of α for estimating $p_i(T)$ (best in the sense of significance as explained above, among the eleven values of α that were used), the one corresponding to the probability estimate displayed in the third column in case $\alpha(i) = 0$ and in the fifth one in case $\alpha(i) > 0$.

The exact (black curve) and estimated (MC standard estimates in blue and MC IPS estimates in red) probabilities are plotted in Figure 1.

3 Local Intensity Model

We now consider a *local intensity model* corresponding to a pure top approach of credit risk, in which the cumulative default process N of a credit portfolio with n names is modeled as a Markov point process stopped at level n (see, for instance, Laurent, Cousin and Fermanian [9] or Cont and Minca [2]). The cumulative default process N is thus a pure birth process with (risk-neutral) *local intensity* $\lambda(t, N_t)$, for a local intensity function λ such that $\lambda(t, i) = 0$ for $i \geq n$ (in order to ensure that N is stopped at level n, since there are n names in the pool). This corresponds to the case where $\widetilde{X} = N = L$ in the set-up of the previous section.

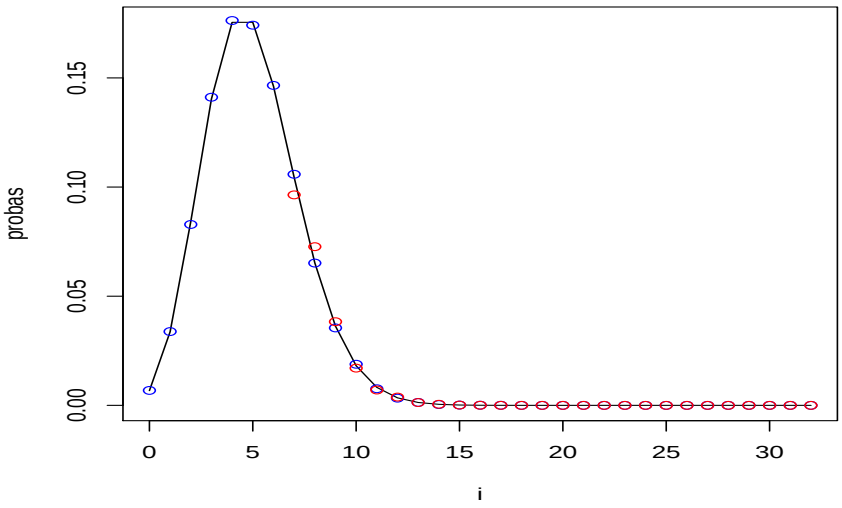

Figure 1: *Standard Monte Carlo versus IPS approach.*

i	$p(i)$	$MC(i)$	$\alpha(i)$	$IPS(i)$	$errIPS(i)$
0	6.737947e-03	6.850000e-03	0	–	1.6630140
1	3.368973e-02	3.381000e-02	0	–	0.3569782
2	8.422434e-02	8.287000e-02	0	–	1.6080120
3	1.403739e-01	1.411600e-01	0	–	0.5600074
4	1.754674e-01	1.762900e-01	0	–	0.4688223
5	1.754674e-01	1.741600e-01	0	–	0.7450786
6	1.462228e-01	1.465800e-01	0	–	0.2442792
7	1.044449e-01	0.10586	1	9.642260e-02	7.6808635
8	6.527804e-02	0.06519	1	7.269477e-02	11.3617485
9	3.626558e-02	0.03548	1	3.832182e-02	5.6699651
10	1.813279e-02	0.01884	2	1.706726e-02	5.8762642
11	8.242177e-03	0.00761	2	7.025229e-03	14.7648865
12	3.434240e-03	0.00331	2	3.798202e-03	10.5980232
13	1.320862e-03	0.00135	2	1.301592e-03	1.4588951
14	4.717363e-04	0.00039	2	4.948383e-04	4.8972237
15	1.572454e-04	0.00014	3	1.315650e-04	16.3314126

Table 1: *Standard MC versus MC IPS.*

16	4.913920e-05	0.00007	3	5.158852e-05	4.9844475
17	1.445271e-05	0.00004	3	1.453860e-05	0.5943091
18	4.014640e-06	0	3	3.570278e-06	11.0685371
19	1.056484e-06	0	3	9.457888e-07	10.4777238
20	2.641211e-07	0	3	2.389309e-07	9.5373474
21	6.288597e-08	0	4	6.775490e-08	7.7424802
22	1.429227e-08	0	4	1.578648e-08	10.4547036
23	3.107014e-09	0	4	3.165438e-09	1.8803674
24	6.472947e-10	0	4	6.513996e-10	0.6341628
25	1.294589e-10	0	4	1.149943e-10	11.1731226
26	2.489595e-11	0	4	2.563031e-11	2.9497091
27	4.610361e-12	0	4	2.805776e-12	39.1419521
28	8.232787e-13	0	4	6.662209e-13	19.0771085
29	1.419446e-13	0	5	1.453048e-13	2.3672480
30	2.365743e-14	0	5	2.589671e-14	9.4654196
31	3.815715e-15	0	5	4.039209e-15	5.8571942
32	5.962055e-16	0	5	4.274102e-16	28.3115984
33	9.033417e-17	0	5	5.303823e-17	41.2866355
34	1.328444e-17	0	6	1.845935e-18	86.1045282

The related generator (spatial generator at time t) writes

$$\mathcal{A}_t = \begin{pmatrix} -\lambda(t,0) & \lambda(t,0) & 0 & 0 & 0 \\ 0 & -\lambda(t,1) & \lambda(t,1) & 0 & 0 \\ & & \cdots & & \\ 0 & 0 & 0 & -\lambda(t,n-1) & \lambda(t,n-1) \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

So $N_0 = 0$ and N jumps by one at some (increasing) $(0,+\infty)$-valued random times $\widetilde{t}_i$. Conditionally on the information $\mathcal{F}_t = \mathcal{F}_t^N$ available at time t, the probability of a jump in the next time interval $(t, t+dt)$ is thus $\lambda(t, N_t)dt$. The related loss distribution $p_i(t) = \mathbb{P}(L_t = i)$, for $i = 0, \cdots, n$, satisfies the following *forward Kolmogorov equation* (system of ODEs), in which $\mathcal{A}^*$ represents the adjoint (transpose) of $\mathcal{A}$:

$$(\partial_t - \mathcal{A}_t^*)p = 0 \text{ on } (0,T] \,, \tag{14}$$

with initial condition $p(0) = \delta_0$ (Dirac mass at 0). Or equivalently, for $t \in [0,T]$:

$$p(t) = \exp(\int_0^t \mathcal{A}_s^* ds)\delta_0 \tag{15}$$

Computing the loss distribution p in this model can thus be achieved by various means, like Runge-Kutta resolution of (6) or numerical matrix exponentiation based on (7).

3.1 Numerical Results

We applied the IPS method in the case $\lambda(t,i) = 1 - \frac{i}{n}$ (cf. Table 3), using numerical matrix exponentiation as a benchmark. Note that this specification $\lambda(t,i) = \frac{n-i}{n}$ corresponds to a model with homogenous and independent obligors, with individual (pre-default) intensity equal to $\frac{1}{n}$.

The left part of Figure 2 displays the concentration of loss levels hit depending on the value of α used (number of hits increasing as the color ranges from blue to red). The right part of

T	n	$\lambda(t,i)$	m	μ	α **(Step)**
5y	125	$1-\frac{i}{n}$	100000	11	0.4

Table 2: *Parameter Values.*

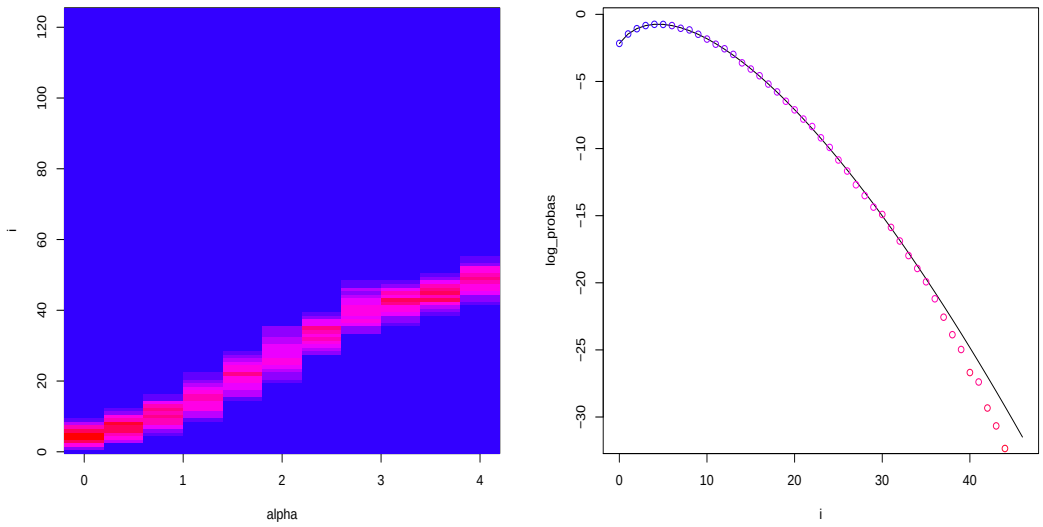

Figure 2: *Losses and Values.*

the figure displays the log-probabilities, exact (black curve) and simulated (points in color). The color of a point refers to the value of α retained for estimating the related probability, from blue for the lowest α to red for the highest.

The left and right part of Figure 3 exhibit the related errors ($\log_{10}$-relative errors, so the level 0 corresponds to a 1% relative error) and significances (estimated probabilities divided by estimated standard deviations). For comparison we plotted in Figure 4 the errors obtained in computing the loss distribution by discretization of the forward Kolmogorov equation (6), using binomial trees with 50, 250 and 1500 time steps, respectively.

Remark 3.1 Even if loops with too small number of runs (like $m = 5000$) do not ensure accuracy over the desired range of loss levels, however they can be fruitfully used for scaling the range of values of α to be used in loops with more runs.

3.1.1 Explicit IS

Figures 5 and 6 display the results obtained by an explicit importance sampling method, consisting on this problem in multiplying the involved intensities by a factor α ranging from 1 to 11, simulating 5000 trajectories in the models with scaled intensities and applying the related payoffs corrections. Color code are as before (color ranging from blue to red as the

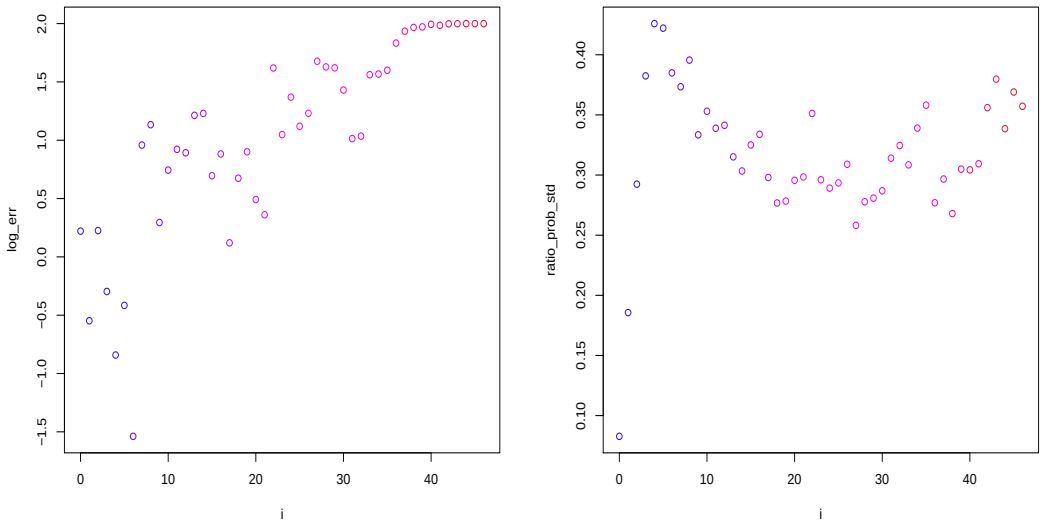

Figure 3: *Errors and Significance.*

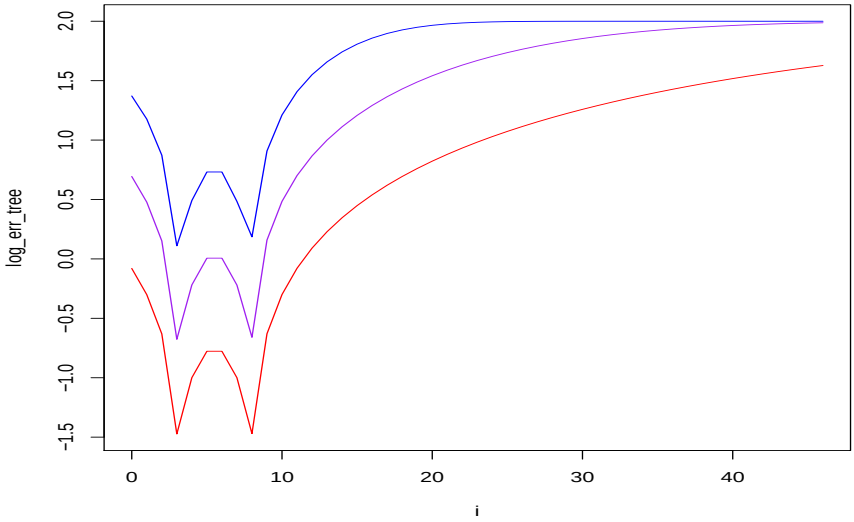

Figure 4: *Errors using binomial trees with 50, 250 and 1500 time steps (monthly, weekly and daily time steps).*

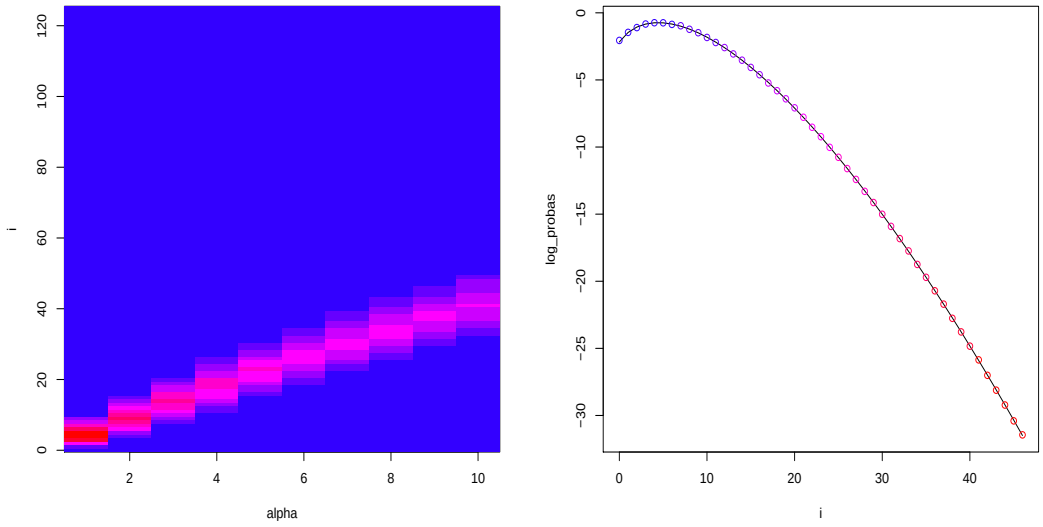

Figure 5: *Explicit IS Losses and Values* ($m = 5000$).

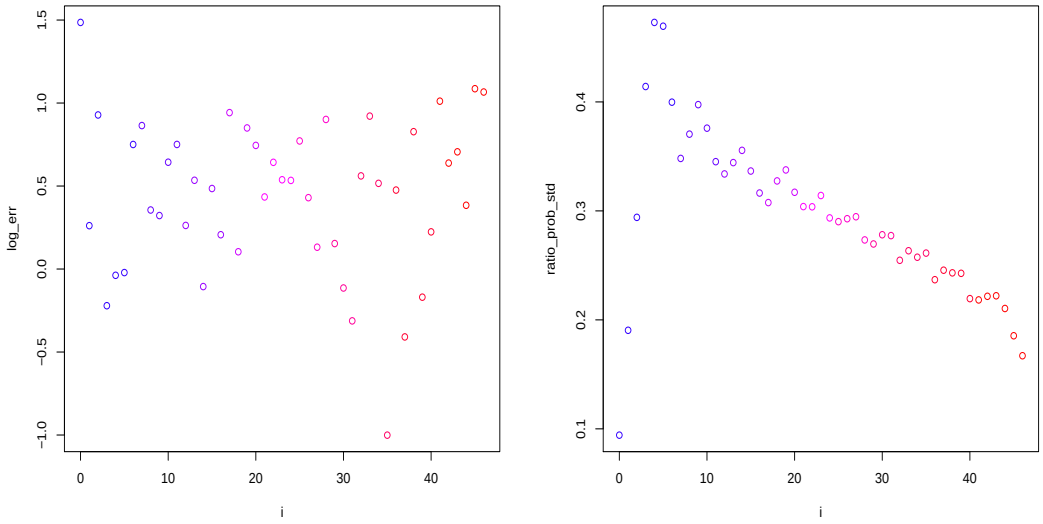

Figure 6: *Explicit IS Errors and Significance* ($m = 5000$).

factor α increases).

4 Homogenous Classes Model

An important issue to consider is whether the IPS approach allows one to cope with the so-called curse of dimensionality. Namely, we need to assess the robustness of the method as the model dimension increases – recall that in low-dimensional models the loss distribution can be recovered exactly by numerical matrix exponentiation, as we did in the previous section.

In this view we now consider a Markov chain model of credit risk as of Frey and Backhaus [5], Bielecki et al. [1], or Herbertsson [7]. Namely, the n names of a pool are grouped in d classes of $\nu = \frac{n}{d}$ homogenous obligors (assuming ν integer), and the cumulative default processes N^l, $l = 1, \cdots, d$, of the different groups are jointly modeled as a d-variate Markov point process N (so simultaneous jumps are excluded), with intensity of N^l given as $\lambda^l(t, N_t)$, for some intensity function $\lambda^l = \lambda^l(t, \iota)$, where $\iota = (i_1, \cdots, i_d) \in \mathbb{N}_\nu^d$. The related generator may thus be written in the form of a $(\nu+1)^d$-dimensional (sparse) matrix Γ. We are thus in the general set-up of section 2 with $\widetilde{X} = N$ and $L = V(N) = \sum N^l$.

For $d = 1$, we recover the local intensity model of the previous section. At the other extreme, for $d = n$, we are in effect modeling the vector of the default indicator processes of the pool names. As d varies between 1 and n, we thus get a variety of models of credit risk, ranging from pure top models for $d = 1$ to pure bottom-up models for $d = n$.

The related classes loss distribution $q_\iota(t) = \mathbb{P}(N_t = \iota)$, for $\iota \in \mathbb{N}_\nu^d$, satisfies the following forward Kolmogorov equation (system of ODEs), in which Γ^* represents the adjoint of Γ:

$$(\partial_t - \Gamma_t^*)q = 0 \text{ on } (0, T] \,, \tag{16}$$

with initial condition $q(0) = \delta_0$ (Dirac mass at $0 \in \mathbb{N}_\nu^d$). Or equivalently, for $t \in [0, T]$:

$$q(t) = \exp(\int_0^t \Gamma_s^* ds)\delta_0 \tag{17}$$

Knowing $q(t)$, the related portfolio loss distribution $p = p_i(t) = \mathbb{P}(L_t = i)$, for $i = 0, \cdots, n$, follows in a straightforward way. However practical resolution of (8) or (9) by determnistic numerical schemes is precluded by the curse of dimensionality for d greater than a few units (depending on ν). So for high d simulation approaches are the only way to go.

4.1 Numerical Results

However, observe that for λ^l given as $\lambda^l(t, \iota) = \frac{\nu - i_l}{n}$, the general Markov chain model (for arbitrary d) reduces to the local intensity model of the previous section (case of independent homogenous obligors $\lambda(t, i) = 1 - \frac{i}{n}$). Applying the IPS methodology and using the results of the previous section as a benchmark, we got the following results. The (rather expected yet) good news is that the accuracy of the method is not altered in higher dimension ($d = 5$, here).

Figures 7 and 8 are the counterparts to Figures 2 and 3, respectively.

Figures 9, 10, 11 and 12 show the results obtained using $m = 20000$ or 5000 runs by Monte Carlo loop (instead of 10^5 above).

4.1.1 Explicit IS

T	$n = d \times \nu$	$\lambda^l(t, \iota)$	μ	α **(Step)**
5y	$125 = 5 \times 25$	$\frac{\nu - i_l}{n}$	11	0.4

Table 3: *Parameter Values.*

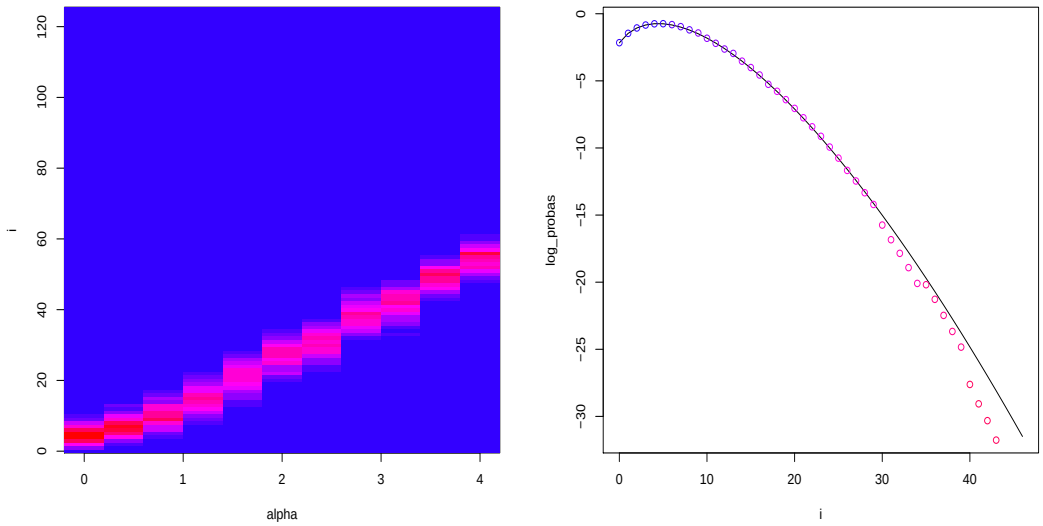

Figure 7: *Losses and Values* ($m = 10^5$).

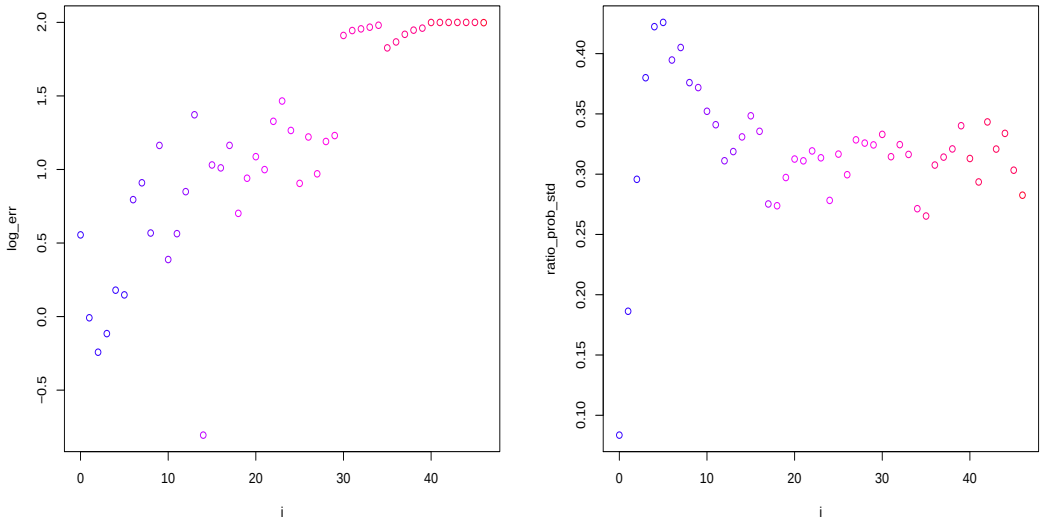

Figure 8: *Errors and Significance* ($m = 10^5$).

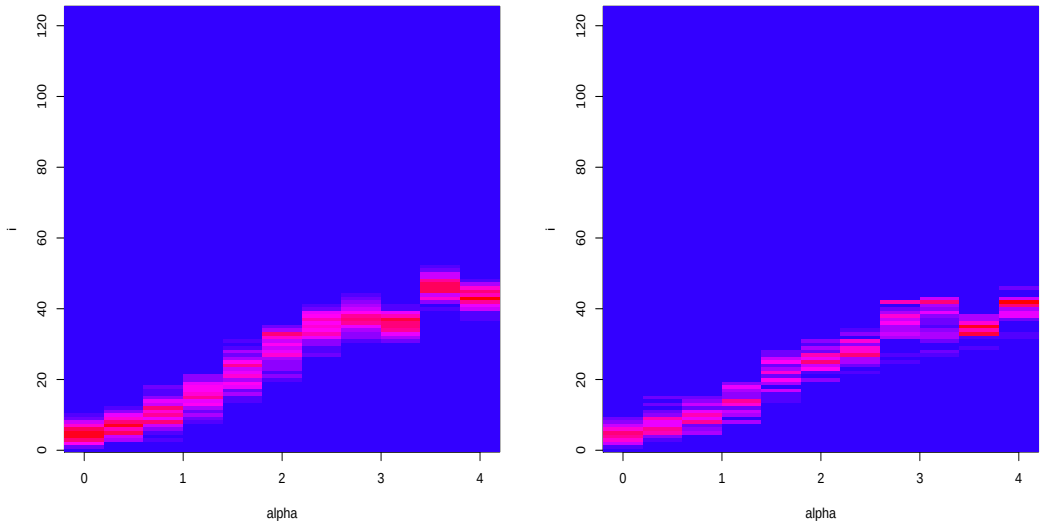

Figure 9: *Losses* $(m = 20000, 5000)$.

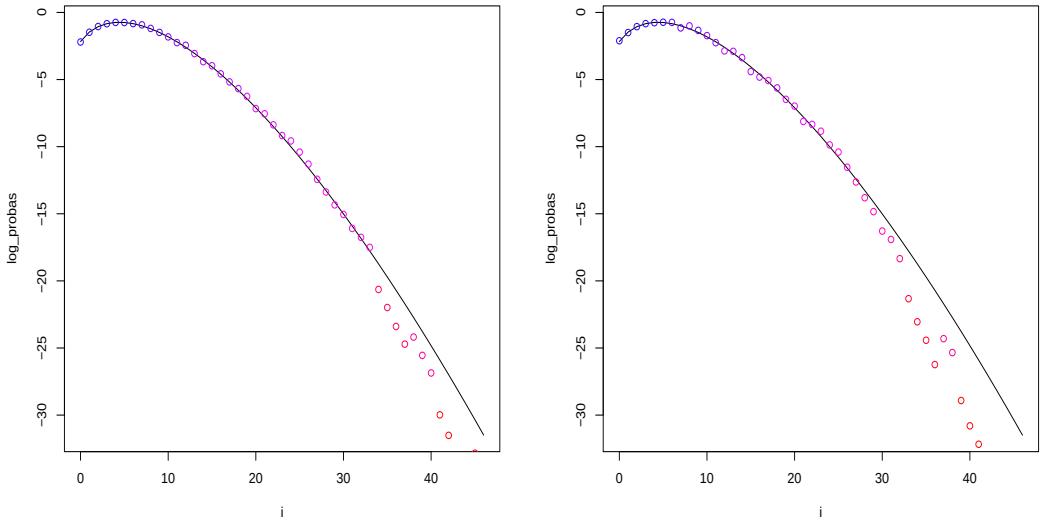

Figure 10: *Values* $(m = 20000, 5000)$.

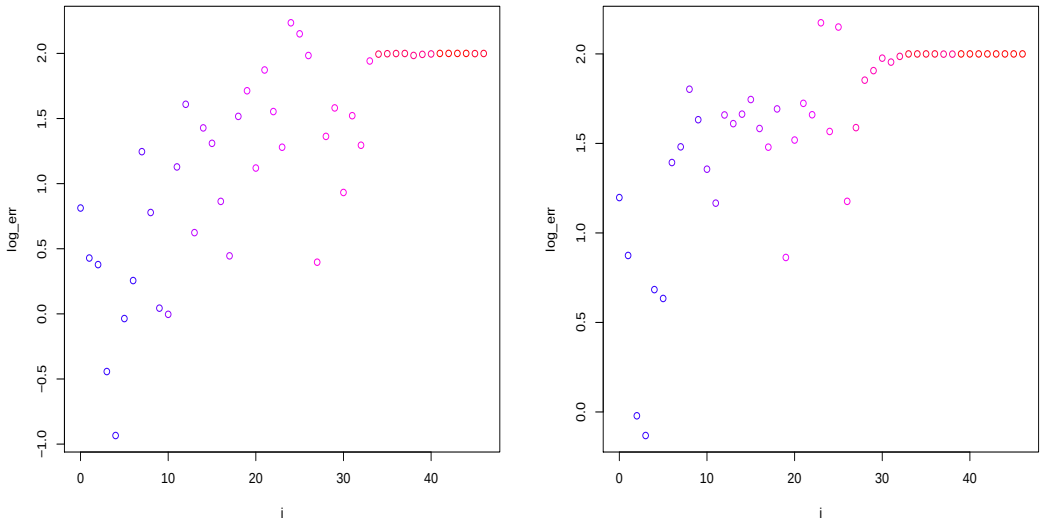

Figure 11: *Errors* $(m = 20000, 5000)$.

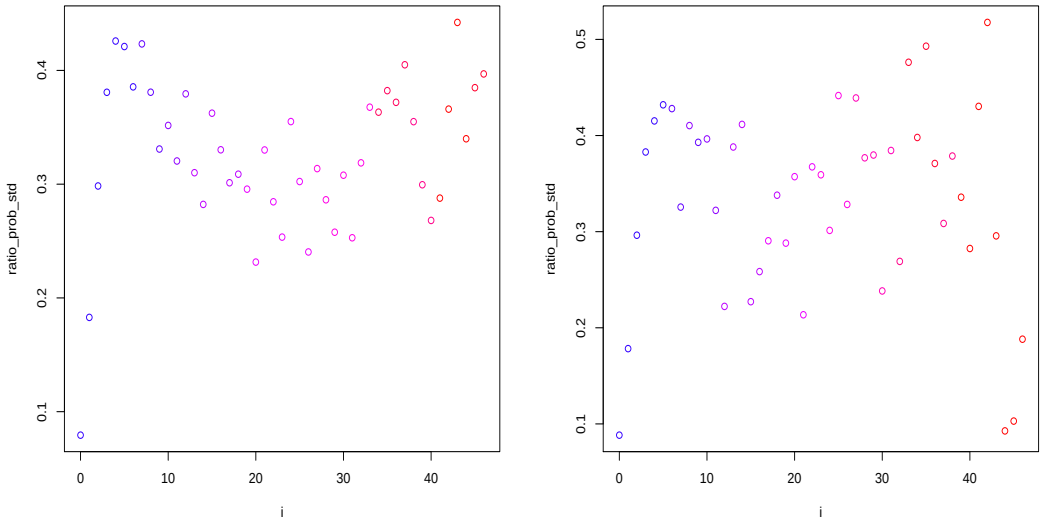

Figure 12: *Significance* $(m = 20000, 5000)$.

Figures 13 and 14 display the results obtained by an explicit importance sampling method consisting in multiplying the involved intensities by a factor α ranging from 1 (standard MC) to 11, simulating 5000 trajectories in the models with scaled intensities and applying the related payoffs corrections (cf. Figures 5 and 6).

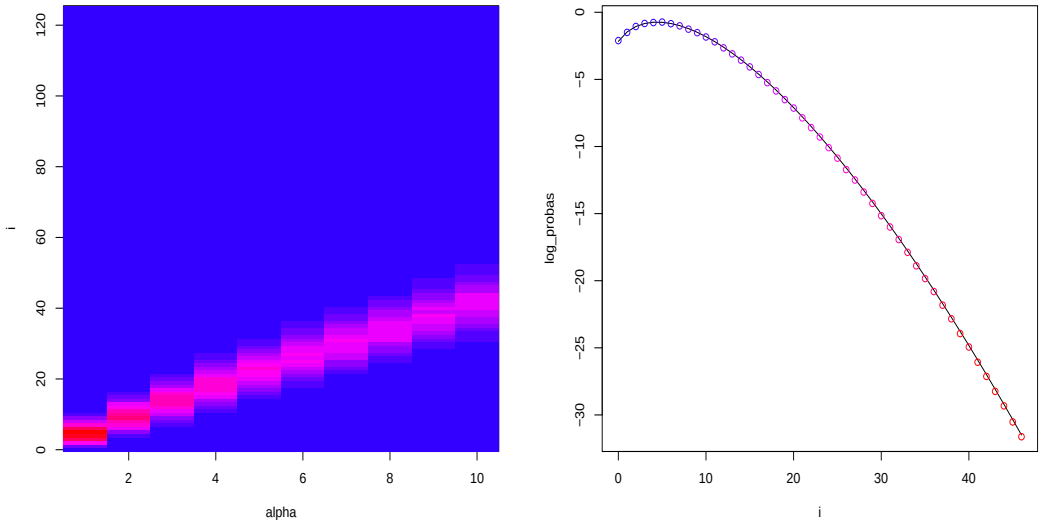

Figure 13: *Explicit IS Losses and Values ($m = 5000$).*

Conclusions and Perspectives

This work is in a preliminary stage. We need to assess further the impact of contagion and the actual application of the method to CDO (senior) tranches.

According to Del Moral and J. Garnier [4], the IPS approach is also powerful to compute conditional probabilities or expectations given the occurrence of some rare event. This could be exploited to deal with (senior) tranche options by simulation.

A pitfall of the approach is that it only allows one to estimate with some accuracy (with a few percents to fifty percents of relative error, say) loss levels with "not so small" probability, not less than $1e^{-15}$ to $1e^{-25}$, say. To get good results for "rarer events" one needs to switch to related yet different methodologies, more demanding for the user (see, e.g., Johansen, Del Moral and Doucet [8]).

References

[1] Bielecki, T.R., Crépey, S., Jeanblanc, M. and Rutkowski, M.: Valuation of basket credit derivatives in the credit migrations environment. *Handbook of Financial Engineering*, J. Birge and V. Linetsky eds., Elsevier, 2007.

[2] Cont, R., Minca, A.: Recovering Portfolio Default Intensities Implied by CDO Quotes, *Working Paper*, 2008.

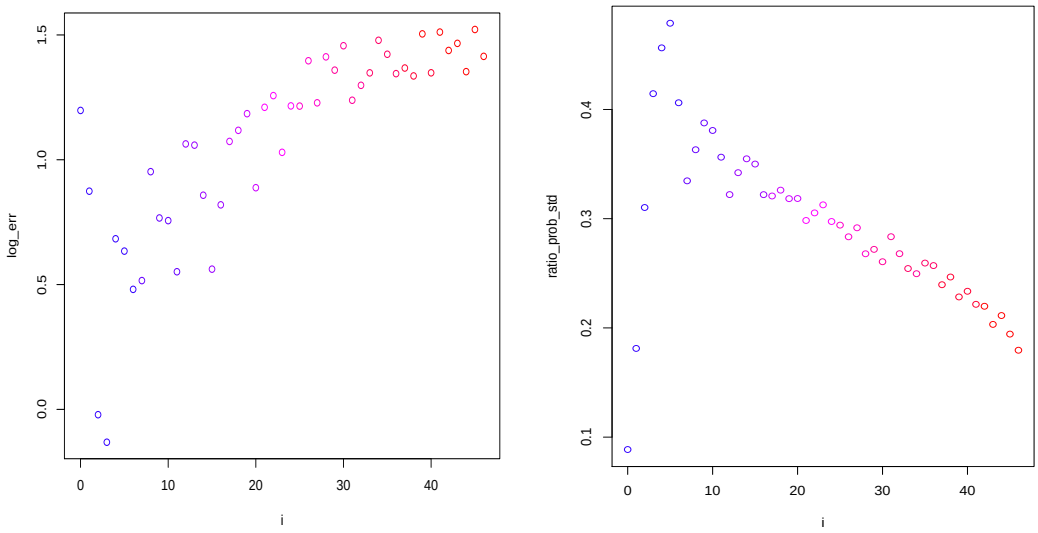

Figure 14: *Explicit IS Errors and Significance ($m = 5000$).*

[3] P. Del Moral. *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications*, Springer-Verlag, 2004.

[4] P. Del Moral and J. Garnier. Genealogical particle analysis of rare events. *Annals of Applied Probability,* 15(4):2496–2534, November 2005.

[5] Frey, R., Backhaus, J.: Pricing and Hedging of Portfolio Credit Derivatives with Interacting Default Intensities. *Working Paper*, 2007.

[6] P. Glasserman. *Monte Carlo Methods in Financial Engineering*, Springer, 2004.

[7] Herbertsson, A.: Pricing portfolio credit derivatives. *PhD Thesis*, G'oteborg University, 2007.

[8] A. M. Johansen, P. Del Moral, and A. Doucet. Sequential Monte Carlo Samplers for Rare Events. *Working Paper*.

[9] Laurent, J.-P., Cousin, A. , Fermanian, J-D.: Hedging default risks of CDOs in Markovian contagion models. *Working Paper*, 2007.

[10] D. Vestal, R. Carmona and J.-P.Fouque. Interacting Particle Systems for the Computation of CDO Tranche Spreads with Rare Defaults. *Submitted*.