



Stochastic Local Intensity Loss Models with Interacting Particle Systems

Aurélien Alfonsi, Céline Labart, Jérôme Lelong

► **To cite this version:**

Aurélien Alfonsi, Céline Labart, Jérôme Lelong. Stochastic Local Intensity Loss Models with Interacting Particle Systems. *Mathematical Finance*, Wiley, 2016, 26 (2), pp.366-394. <10.1111/mafi.12059>. <hal-00786239>

HAL Id: hal-00786239

<https://hal.archives-ouvertes.fr/hal-00786239>

Submitted on 8 Feb 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Stochastic Local Intensity Loss Models with Interacting Particle Systems

Aurélien Alfonsi *

Université Paris-Est, CERMICS,
Project Team MathRisk ENPC-INRIA-UMLV, Ecole des Ponts,
6-8 avenue Blaise Pascal, 77455 Marne-la-Vallée, France

Céline Labart †

Laboratoire de Mathématiques, CNRS UMR 5127
Université de Savoie, Campus Scientifique
73376 Le Bourget du Lac, France

Jérôme Lelong ‡

Univ. Grenoble Alpes, Laboratoire Jean Kuntzmann,
BP 53, 38041 Grenoble, Cedex 09, France

February 8, 2013

Abstract

It is well-known from the work of Schönbucher (2005) that the marginal laws of a loss process can be matched by a unit increasing time inhomogeneous Markov process, whose deterministic jump intensity is called local intensity. The Stochastic Local Intensity (SLI) models such as the one proposed by Arnsdorf and Halperin (2008) allow to get a stochastic jump intensity while keeping the same marginal laws. These models involve a non-linear SDE with jumps. The first contribution of this paper is to prove the existence and uniqueness of such processes. This is made by means of an interacting particle system, whose convergence rate towards the non-linear SDE is analyzed. Second, this approach provides a powerful way to compute pathwise expectations with the SLI model: we show that the computational cost is roughly the same as a crude Monte-Carlo algorithm for standard SDEs.

Keywords : Stochastic local intensity model, Interacting particle systems, Loss modelling, Credit derivatives, Monte-Carlo Algorithm, Fokker-Planck equation, Martingale problem.

AMS classification (2010): 91G40, 91G60, 60H35

Acknowledgments: We would like to thank Benjamin Jourdain (Université Paris-Est) for fruitful discussions having stimulated this research. We would like to thank INRIA Paris Rocquencourt for allowing us to run our numerical tests on the convergence rate on their RIOG cluster. A. Alfonsi acknowledges the support of the “Chaire Risques Financiers” of Fondation du Risque. C. Labart and J. Lelong acknowledge the support of the Finance for Energy Market Research Centre, www.fime-lab.org.

*Email:alfonsi@cermics.enpc.fr

†Email:celine.labart@univ-savoie.fr

‡Email:jerome.lelong@imag.fr

1 Introduction

In equity modeling, a major concern is to get a model that fits option data. It is well-known from the work of Dupire (1994) that basically European options can be exactly calibrated by using local volatility models $\sigma(t, x)$. However, local volatility models are known to have some inadequacy to describe real markets. To get richer dynamics, Stochastic Local Volatility (SLV) models have been introduced (see Alexander and Nogueira (2004) or Piterbarg (2006)) and consider the following dynamics for the stock under a risk-neutral probability:

$$dS_t = rS_t dt + f(Y_t)\eta(t, S_t)S_t dW_t,$$

where Y_t is an adapted stochastic process. Typically, $(Y_t, t \geq 0)$ is assumed to solve an autonomous one dimensional SDE whose Brownian motion may be correlated with W . From the work of Gyöngy (1986), we know that under mild assumptions, the following choice

$$\eta(t, x) = \frac{\sigma(t, x)}{\sqrt{\mathbb{E}[f(Y_t)^2 | S_t = x]}}$$

ensures that S_t has the same marginal laws as the local volatility model with $\sigma(t, x)$, which automatically gives the calibration to European option prices. This leads to the following non-linear SDE

$$dS_t = rS_t dt + \frac{f(Y_t)}{\sqrt{\mathbb{E}[f(Y_t)^2 | S_t]}} \sigma(t, S_t) S_t dW_t.$$

Here, we stress that the law of (Y_t, S_t) steps into the diffusion term. Unless for trivial choices of Y_t and despite some attempts (Abergel and Tachet (2010)), getting the existence and uniqueness of solutions for this kind of SDE remains an open problem. Also, from a numerical perspective, the simulation of SLV models is not easy, precisely because of the computation of the conditional expectation.

In this paper, we propose to tackle a very analogous problem arising in credit risk modeling. In all the paper, we will work under a risk-neutral probabilistic filtered space $(\Omega, (\mathcal{F}_t)_{t \geq 0}, \mathcal{F}, \mathbb{P})$. As usual, \mathcal{F}_t is the σ -field describing all the events that can occur before time t and \mathcal{F} describes all the events. We consider $M \in \mathbb{N}^*$ defaultable entities (for example, $M = 125$ for the iTraxx). We assume that all the recovery rates are deterministic and equal to $1 - L_{GD}$ for all the firms within the basket. The loss process $(L_t, t \geq 0)$ is given by

$$L_t = \frac{L_{GD}}{M} X_t, \quad \forall t \geq 0$$

where X_t is the number of defaults up to time t . Clearly, X takes values in $\mathcal{L}_M := \{0, \dots, M\}$. Thanks to the assumption of deterministic recovery rates, and under the assumption of deterministic short interest rates, it is well known that CDO tranche prices only depend on the marginal laws of the loss process (see for example Remark 3.2.1 in Alfonsi (2011)). Let us assume then for a while that we have found marginal laws $(P(X_t = k), k \in \mathcal{L}_M)_{t \in [0, T]}$ which perfectly fit CDO tranche prices up to maturity $T > 0$. Then, under some mild assumptions, we know from Schönbucher (2005) that there exists a non-homogeneous Markov chain with only unit increments which exactly matches these marginal laws. Somehow, this result plays the same role for the loss as Dupire's result for the stock.

The loss model obtained with a non-homogeneous unit-increasing Markov chain is known in the literature as local intensity model. It is fully described by the local intensity λ :

$\mathbb{R}_+ \times \mathcal{L}_M \rightarrow \mathbb{R}_+$, which gives the instantaneous rate of having one more default in the basket. In the sequel, we will assume that the local intensity $\lambda : \mathbb{R}_+ \times \mathcal{L}_M \rightarrow \mathbb{R}_+$ has been calibrated to market data and perfectly matches Index and CDO tranche prices. We make the following assumptions:

- $\forall x \in \mathcal{L}_M, t \in \mathbb{R} \mapsto \lambda(t, x)$ is a càdlàg function,
- $\forall t \geq 0, \lambda(t, M) = 0$.

In particular, we have $\bar{\lambda} = \max_{x \in \mathcal{L}_M} \sup_{t \in [0, T]} \lambda(t, x) < \infty$. In this setting, the instantaneous jump rate at time t from x to $x + 1$ is given by $\lambda(t-, x)$. Thus, the local intensity model corresponds to a time-inhomogeneous Markov chain making unit jumps with this rate.

One may like however get richer dynamics than the ones given by the local intensity model. Then, we can proceed in the same way as for the stochastic volatility models. Let us consider $(Y_t)_{t \geq 0}$ a general (\mathcal{F}_t) -adapted càdlàg real process, a function $f : \mathbb{R} \rightarrow \mathbb{R}_+$ and a function $\eta : \mathbb{R}_+ \times \mathcal{L}_M \rightarrow \mathbb{R}_+$ satisfying the same assumptions as λ . We assume that the default counting process $(X_t, t \geq 0)$ has jumps of size 1 with the rate

$$\eta(t-, X_{t-})f(Y_{t-}).$$

By analogy with the equity, we name this kind of model a Stochastic Local Intensity (SLI) model. Then, it is known (see Cont and Minca (2013)) that the local intensity model with the Local Intensity (LI) $\eta(t-, x)\mathbb{E}[f(Y_{t-})|X_{t-} = x]$ has the same marginal laws as X_t . Thus, the SLI model will be automatically calibrated to CDO tranche prices if one takes:

$$\forall t > 0, x \in \mathcal{L}_M, \eta(t-, x) = \frac{\lambda(t-, x)}{\mathbb{E}[f(Y_{t-})|X_{t-} = x]}.$$

This approach has been used in the literature by Arnsdorf and Halperin (2008), and in a slightly different way by Lopatin and Misirpashaev (2008). However, up to our knowledge there is no proof in the literature of the existence nor uniqueness of such a dynamics.

The first scope of this paper is to solve this problem. At this stage, we need to make our framework precise. We assume through the paper that:

$$f : \mathbb{R} \mapsto \mathbb{R} \text{ is continuous, s.t. } \forall x \in \mathbb{R}, 0 < \underline{f} \leq f(x) \leq \bar{f} < \infty. \quad (1.1)$$

We assume that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ contains a standard Brownian motion $(W_t, t \geq 0)$, a sequence of independent uniform random variables $(U^k)_{k \in \mathbb{N}}$, and a sequence $(E^n)_{n \in \mathbb{N}}$ of independent exponential random variables with parameter $\frac{\bar{f}}{\underline{f}}$. We set $T^k = \sum_{n=1}^k E^n$ for $k \in \mathbb{N}^*$. The random variables $(T^k, U^k)_{k \in \mathbb{N}^*}$ will enable us to define a non homogeneous Poisson point process with jump intensity $\eta(t-, X_{t-})f(Y_{t-})$. We are interested in studying the following two problems in which we assume that Y_t is a process with values either in \mathbb{N} (discrete case) or in \mathbb{R} (continuous case). In the discrete case, we are interested in finding a predictable process $(X_t, Y_t)_{t \geq 0}$ such that

$$\begin{cases} X_t = x_0 + \sum_{k, T^k \leq t} \mathbf{1} \left\{ U^k \leq \frac{f(Y_{T^k-})\lambda(T^k-, X_{T^k-})}{\bar{f} \mathbb{E}[f(Y_{T^k-})|X_{T^k-}]} \right\} \\ Y_0 = y_0, \text{ and for each } k \geq 0, (Y_t, t \in [T_k, T_{k+1})) \text{ is a continuous time Markov chain} \\ \text{with transition rate } \mu_{ij}^{X_t} = \mu_{ij}^{X_{T^k}}. \end{cases} \quad (1.2)$$

For the sake of simplicity, we consider in the discrete setting that X and Y do not jump together almost surely.

In the continuous case, the corresponding problem is to solve the following stochastic differential equation:

$$\begin{cases} X_t = x_0 + \sum_{k, T^k \leq t} \mathbf{1} \left\{ U^k \leq \frac{f}{\lambda f} \frac{f(Y_{T^k-}) \lambda(T^k-, X_{T^k-})}{\mathbb{E}[f(Y_{T^k-}) | X_{T^k-}]} \right\} \\ Y_t = y_0 + \int_0^t b(s, X_s, Y_s) ds + \int_0^t \sigma(s, X_s, Y_s) dW_s + \int_0^t \gamma(s-, X_{s-}, Y_{s-}) dX_s, \end{cases} \quad (1.3)$$

for $x_0 \in \mathcal{L}_M, y_0 \in \mathbb{R}$ and given real functions b, σ and γ . This framework embeds in particular the dynamics suggested by Arnsdorf and Halperin (2008) and Lopatin and Misirpashaev (2008). Under some rather mild hypotheses on μ_{ij}, b, σ and γ , which will be specified in the corresponding sections, we will show that the above two equations admit a unique solution. In the discrete case, we are able to show that the corresponding Fokker-Planck equation has a unique solution. This can be achieved by writing the Fokker-Planck equation as an ODE which can be studied directly. In the continuous case, this approach can hardly be extended: the Fokker-Planck equation leads to a non trivial PDE. Instead, we solve this problem by introducing an interacting particle system. This technique is known to be powerful for this type of non linear problems (see Sznitman (1991) or Méléard (1996)).

The second scope of this paper is to provide a way to compute prices under SLI models. Indeed, interacting particle systems are not only theoretical tools to prove existence and uniqueness results for such equations. They give a very smart way to simulate these processes, therefore enabling us to run Monte-Carlo algorithms. This approach has been recently used by Guyon and Henry-Labordère (2011) for Stochastic Local Volatility models. For the Stochastic Local Intensity models considered in this work, the conditional expectation is much simpler to handle. This enables us to get theoretical results on the convergence and also simplifies the implementation. In fact, we show in our case under some assumptions that the rate of convergence to estimate expectations is in $O(1/N^\alpha)$ for any $\alpha < 1/2$, where N is the number of particles. On our numerical experiments, we even observe on several examples a convergence which is similar to the one of the Central Limit Theorem, which is rather usual for Interacting Particle Systems. Besides, we show that we can simulate the interacting particle system with a computational cost in $O(DN)$, where D is the number of time steps for the discretization of the SDE on Y . Thus, the computational cost is roughly the same as a crude Monte-Carlo algorithm for standard SDEs with N samples.

The paper is organized as follows. First, we study the case where Y has discrete values; this framework enables us to settle the problem and solve it by rather elementary tools. This part is independent from the rest of the paper. Second, by means of a particle system approach, we investigate the case where Y is real valued jump diffusion. Finally, we carry out numerical simulations highlighting the relevance of the particle system technique to compute pathwise expectation of the Process (1.3).

2 The SLI model when Y takes discrete values

The goal of this section is to prove the existence of a process $(X_t, Y_t)_{t \geq 0}$ satisfying (1.2). Unlike the continuous case (1.3), we can get this result by elementary means, without resorting to

an interacting particle system. To do so, we write the Fokker-Planck equation associated to the process $(X_t, Y_t)_{t \geq 0}$, which should be satisfied by $\mathbb{P}(X_t = i, Y_t = j)$ for $(i, j) \in \mathcal{L}_M \times \mathbb{N}$. We have

$$(\mathcal{E}) \begin{cases} \partial_t p(t, i, j) &= \sum_{k \neq j} \mu_{kj}^i p(t, i, k) + \mathbf{1}_{\{i \geq 1\}} \frac{\lambda(t, i-1)}{\varphi_p(t, i-1)} f(j) p(t, i-1, j) \\ &\quad - \left(\frac{\lambda(t, i)}{\varphi_p(t, i)} f(j) \mathbf{1}_{\{i \leq M-1\}} - \mu_{jj}^i \right) p(t, i, j) \\ p(0, i, j) &= 0 \quad \forall (i, j) \neq (x_0, y_0), \\ p(0, x_0, y_0) &= 1. \end{cases}$$

where p is a function from $\mathbb{R}_+ \times \mathcal{L}_M \times \mathbb{N}$ to \mathbb{R} and

$$\varphi_p(t, i) = \frac{\sum_{j=0}^{\infty} f(j) p(t, i, j)}{\sum_{j=0}^{\infty} p(t, i, j)}.$$

If we manage to prove that the Fokker-Planck equation admits a unique solution p such that

$$\forall i, j \in \mathcal{L}_M \times \mathbb{N}, \forall t \geq 0, p(t, i, j) \geq 0, \quad (2.1)$$

$$\forall t \geq 0 \sum_{i=0}^M \sum_{j=0}^{\infty} p(t, i, j) = 1, \quad (2.2)$$

then we will get that the law of a process $(X_t, Y_t)_{t \geq 0}$ satisfying (1.2) is unique. Besides, we will also get the existence of such a process. It is easy to check that a continuous Markov chain $(X_t, Y_t)_{t \geq 0}$ starting from (x_0, y_0) with transition rate matrix

$$\tilde{\mu}_{(i_1, j_1), (i_2, j_2)}(t) = \mathbf{1}_{j_1=j_2} (\mathbf{1}_{i_2=i_1+1} - \mathbf{1}_{i_2=i_1}) \frac{f(j_1) \lambda(t, i_1)}{\varphi_p(t, i_1)} + \mathbf{1}_{i_1=i_2} \mu_{j_1 j_2}^{i_1}, \quad i_1, i_2 \in \mathcal{L}_M, j_1, j_2 \in \mathbb{N},$$

where p is the solution of (\mathcal{E}) satisfies (1.2). In fact, the Fokker-Planck equation of this process

$$\begin{cases} \partial_t q(t, i, j) &= \sum_{k \neq j} \mu_{kj}^i q(t, i, k) + \mathbf{1}_{\{i \geq 1\}} \frac{\lambda(t, i-1)}{\varphi_p(t, i-1)} f(j) q(t, i-1, j) \\ &\quad - \left(\frac{\lambda(t, i)}{\varphi_p(t, i)} f(j) \mathbf{1}_{\{i \leq M-1\}} - \mu_{jj}^i \right) q(t, i, j) \\ q(0, i, j) &= 0 \quad \forall (i, j) \neq (x_0, y_0), \\ q(0, x_0, y_0) &= 1. \end{cases}$$

is linear and clearly solved by p , which gives $q \equiv p$.

2.1 Assumptions and notations

In this part, we assume that the transition rates satisfy the following hypothesis.

Hypothesis 1 *The intensity matrices $(\mu_{ij}^k)_{i, j \geq 0}$ for $k \in \mathcal{L}_M$ satisfy the following conditions:*

- $\forall k \in \mathcal{L}_M, \forall i, j \in \mathbb{N} \times \mathbb{N}$ such that $i \neq j$ $\mu_{ij}^k \geq 0$,
- $\forall k \in \mathcal{L}_M, \forall i \in \mathbb{N}$ $\mu_{ii}^k \leq 0$,
- $\forall k \in \mathcal{L}_M, \forall i \in \mathbb{N}$ $\sum_{j=0}^{\infty} \mu_{ij}^k = 0$ (then $\sum_{j=0, j \neq i}^{\infty} \mu_{ij}^k = -\mu_{ii}^k$).

Moreover, we assume $\forall k \in \mathcal{L}_M, \sup_{i \in \mathbb{N}} |\mu_{ii}^k| < \infty$.

We also introduce specific notations used in the discrete case.

- We define the set E as the set of real sequences indexed by $i \in \mathcal{L}_M$ and $j \in \mathbb{N}$:

$$E := \{u = (u_j^i)_{0 \leq i \leq M, 0 \leq j} : u_j^i \in \mathbb{R}\},$$

$$E_+ := \{u = (u_j^i)_{0 \leq i \leq M, 0 \leq j} : u_j^i \in \mathbb{R}_+\} \text{ and } E_+^* := \{u = (u_j^i)_{0 \leq i \leq M, 0 \leq j} : u_j^i \in \mathbb{R}_+^*\}.$$

- For $u \in E$, we set $|u| = \sum_{i=0}^M \sum_{j=0}^{\infty} |u_j^i|$.

2.2 Solving the Fokker-Planck equation

To be more concise, we rewrite the Fokker-Planck Equation (\mathcal{E}) and Conditions (2.1) – (2.2) by using a sequence of functions. Let $P := (P_j^i)_{0 \leq i \leq M, j \geq 0}$ denote a sequence such that each P_j^i is a function from \mathbb{R}_+ to \mathbb{R} . Solving (\mathcal{E}) under Constraints (2.1) – (2.2) boils down to solving

$$(\mathcal{E}') \begin{cases} P'(t) &= \Psi(t, P(t)), \\ P(0) &= P_0 \end{cases}$$

under the constraints $\forall t \geq 0, P(t) \geq 0$ and $|P(t)| = 1$. The sequence P_0 is such that $(P_0)_j^i = 0$ for $(i, j) \neq (x_0, y_0)$ and $(P_0)_{y_0}^{x_0} = 1$. Ψ is an application from $\mathbb{R}_+ \times E_+ \rightarrow E$ given by

$$(\Psi(t, x))_j^i = \sum_{k \geq 0} \mu_{kj}^i x_k^i + \mathbf{1}_{\{i \geq 1\}} \frac{\lambda(t, i-1)}{\varphi(x, i-1)} f(j) x_j^{i-1} - \mathbf{1}_{\{i \leq M-1\}} \left(\frac{\lambda(t, i)}{\varphi(x, i)} f(j) \right) x_j^i$$

where $\varphi(x, i) := \frac{\sum_{l=0}^{\infty} f(l) x_l^i}{\sum_{l=0}^{\infty} x_l^i}$.

Remark 1. Ψ is defined without ambiguity on E_+^* . When $x \in E_+$, difficulties may arise when for some fixed i , $x_j^i = 0 \forall j$. In this case, we still have $\lim_{z \in E_+^*, z \rightarrow x} \frac{z_j^i}{\varphi(z, i)} = 0$ since $\underline{f} \leq \varphi(z, i) \leq \bar{f}$ for $z \in E_+^*$. Thus, we can extend Ψ by continuity on E_+ .

We aim at solving (\mathcal{E}) in the set of summable sequences (compatible with Condition (2.2)) and get the following result.

Theorem 2. Equation (\mathcal{E}') admits a unique solution on \mathbb{R}_+ satisfying $\forall t \geq 0, P(t) \geq 0$ and $|P(t)| = 1$.

To do so, we first focus on the following differential equations:

$$(\mathcal{E}'') \begin{cases} P'(t) &= \Psi(t, (P(t))_+), \\ P(0) &= P_0. \end{cases}$$

Proposition 3. Equation (\mathcal{E}'') admits a unique solution on \mathbb{R}_+ . Moreover, the solution satisfies $\forall t \geq 0, P(t) \geq 0$ and $|P(t)| = 1$.

The proof of this proposition consists in first showing the Lipschitz property which gives the existence and uniqueness of P and then proving that $\forall t \geq 0, P(t) \geq 0$ and $|P(t)| = 1$. This proof is postponed to Appendix A.

Then, the proof of Theorem 2 becomes obvious. The unique solution P of (\mathcal{E}'') clearly solves (\mathcal{E}'), and any solution Q of (\mathcal{E}') such that $\forall t \geq 0, Q(t) \geq 0$ and $|Q(t)| = 1$ also solves (\mathcal{E}'') and thus coincides with P .

3 The SLI model when Y is real valued.

3.1 Setting and main results

We are interested in proving the existence of a process $(X_t, Y_t)_t$ solving the stochastic differential equation (1.3). More precisely, we will consider the following stochastic differential equation,

$$\begin{cases} X_t = X_0 + \sum_{k, T^k \leq t} \mathbf{1} \left\{ U^k \leq \frac{f(Y_{T^k-}) \lambda(T^k-, X_{T^k-})}{\lambda f} \right\} \\ Y_t = Y_0 + \int_0^t b(s, X_s, Y_s) ds + \int_0^t \sigma(s, X_s, Y_s) dW_s + \int_0^t \gamma(s-, X_{s-}, Y_{s-}) dX_s, \end{cases} \quad (3.1)$$

with (possibly) random initial condition (X_0, Y_0) such that $\mathbb{E}[|Y_0|^m] < \infty$ for any $m \in \mathbb{N}$. We will denote in the sequel \mathcal{L}_{init} the probability law of (X_0, Y_0) under \mathbb{P} . To get existence and uniqueness results for (3.1), we will make the following assumption on the coefficients.

Hypothesis 2 1. *The functions $b, \sigma, \gamma : \mathbb{R}_+ \times \mathcal{L}_M \times \mathbb{R} \rightarrow \mathbb{R}$ are measurable, with sub-linear growth with respect to y :*

$$\forall T > 0, \exists C_T > 0, \forall t \in [0, T], x \in \mathcal{L}_M, y \in \mathbb{R}, |b(t, x, y)| + |\sigma(t, x, y)| + |\gamma(t, x, y)| \leq C_T(1 + |y|).$$

2. *The functions $b(t, x, y)$ and $\sigma(t, x, y)$ are such that for any $x \in \mathcal{L}_M, y_0 \in \mathbb{R}$, there exists a unique strong solution for the SDE*

$$Y_t = y_0 + \int_0^t b(s, x, Y_s) ds + \int_0^t \sigma(s, x, Y_s) dW_s, \quad t \geq 0.$$

This property holds if we assume for example that:

$$\forall T > 0, \exists C_T > 0, \forall t \in [0, T], x \in \mathcal{L}_M, \begin{cases} |b(t, x, y) - b(t, x, y')| \leq C_T |y - y'| \\ |\sigma(t, x, y) - \sigma(t, x, y')| \leq C_T \sqrt{|y - y'|}. \end{cases}$$

3. *For any $x \in \mathcal{L}_M, (t, y) \mapsto \gamma(t, x, y)$ is càdlàg with respect to t and continuous with respect to y , i.e. $\gamma(t, x, y) = \lim_{s > t, s \rightarrow t, z \rightarrow y} \gamma(s, x, z)$ and $\lim_{s < t, s \rightarrow t, z \rightarrow y} \gamma(s, x, z)$ exists and is denoted by $\gamma(t-, x, y)$.*

To prove the strong existence and uniqueness of a process (X, Y) solving (3.1), we will first need to prove a weak existence and uniqueness result. To do so, we introduce the Martingale Problem associated with (3.1). We denote by $\mathcal{D}([0, T], \mathbb{R})$ the set of càdlàg real valued functions and consider:

$$E = \{(x(t), y(t))_{t \in [0, T]}, s.t. y \in \mathcal{D}([0, T], \mathbb{R}) \text{ and } x \text{ is a nondecr. càdlàg function with values in } \mathcal{L}_M\}.$$

This path space is endowed with the usual Skorokhod topology for càdlàg processes, and with the associated Borelian σ -algebra. We denote by $\mathcal{P}(E)$ the set of probability measures on E . We are looking for a probability measure $\overline{\mathbb{Q}} \in \mathcal{P}(E)$ such that \mathcal{L}_{init} is the probability law of (X_0, Y_0) under $\overline{\mathbb{Q}}$ and, for any $\phi \in C^{0,2}(\mathbb{R}^2, \mathbb{R})$,

$$\begin{aligned} \phi(X_t, Y_t) - \phi(X_0, Y_0) - \int_0^t \left\{ \frac{\lambda(u, X_u) f(Y_u)}{\mathbb{E}_{\overline{\mathbb{Q}}}[f(Y_u) | X_u]} [\phi(X_u + 1, Y_u + \gamma(u, X_u, Y_u)) - \phi(X_u, Y_u)] \right. \\ \left. + b(u, X_u, Y_u) \partial_y \phi(X_u, Y_u) + \frac{1}{2} \sigma^2(u, X_u, Y_u) \partial_y^2 \phi(X_u, Y_u) \right\} du \end{aligned} \quad (3.2)$$

is a martingale with respect to the filtration $\mathcal{F}_t = \sigma((X_u, Y_u), u \leq t)$ satisfying the usual conditions. Here, X_t and Y_t stand for the coordinate applications:

$$\begin{aligned} X_t : E &\rightarrow \mathcal{L}_M & \text{and } Y_t : E &\rightarrow \mathbb{R} \\ (x, y) &\mapsto x(t) & (x, y) &\mapsto y(t), \end{aligned}$$

and $\mathbb{E}_{\overline{\mathbb{Q}}}$ denotes the expectation under the probability measure $\overline{\mathbb{Q}} \in \mathcal{P}(E)$. Similarly, for any $\mathbb{Q} \in \mathcal{P}(E)$, we denote by $\mathbb{E}_{\mathbb{Q}}$ the expectation under the probability measure \mathbb{Q} while \mathbb{E} simply denotes the expectation under the original probability measure \mathbb{P} .

The following Theorem is the main result of the paper.

Theorem 4. *We assume that Hypothesis 2 holds and consider a \mathcal{F}_0 -measurable initial condition (X_0, Y_0) such that $\forall m \in \mathbb{N}, \mathbb{E}[|Y_0|^m] < \infty$. Then, there exists a unique probability measure $\overline{\mathbb{Q}} \in \mathcal{P}(E)$ solving the Martingale Problem (3.2). Besides, there exists a unique strong solution $(X_t, Y_t)_{t \geq 0}$ to Equation (3.1).*

To prove Theorem 4, we need the following basic result on standard SDEs with jumps. This is an easy consequence of Hypothesis 2. For the sake of completeness, we give its proof in Appendix B.

Proposition 5. *For $\mathbb{Q} \in \mathcal{P}(E)$, we set for $t \in [0, T], x \in \mathcal{L}_M$*

$$\varphi^{\mathbb{Q}}(t, x) = \frac{\mathbb{E}_{\mathbb{Q}}[f(Y_t)\mathbf{1}_{\{X_t=x\}}]}{\mathbb{Q}(X_t=x)}, \text{ when } \mathbb{Q}(X_t=x) > 0 \text{ and } \varphi^{\mathbb{Q}}(t, x) = \underline{f} \text{ otherwise.}$$

Let Hypothesis 2 hold. Then, for any $(x_0, y_0) \in \mathcal{L}_M \times \mathbb{R}$, there exists a unique strong solution $(X_t, Y_t)_{t \in [0, T]}$ to the following SDE with jumps:

$$\begin{cases} X_t = x_0 + \sum_{k, T^k \leq t} \mathbf{1} \left\{ U^k \leq \frac{\underline{f}}{\lambda \underline{f}} \frac{f(Y_{T^k-}) \lambda(T^k-, X_{T^k-})}{\varphi^{\mathbb{Q}}(T^k-, X_{T^k-})} \right\} \\ Y_t = y_0 + \int_0^t b(s, X_s, Y_s) ds + \int_0^t \sigma(s, X_s, Y_s) dW_s + \int_0^t \gamma(s-, X_{s-}, Y_{s-}) dX_s. \end{cases} \quad (3.3)$$

Proof of Theorem 4. The proof of Theorem 4 is split in three main steps.

- First, we show the existence of a probability measure $\overline{\mathbb{Q}} \in \mathcal{P}(E)$ solving the Martingale Problem (3.2). This result is obtained by considering the associated interacting particle system: we show that each particle converges in law, and that any probability measure in the support of the limiting law solves the Martingale Problem. This is done in Section 3.3.
- Second, we show the uniqueness of the probability measure $\overline{\mathbb{Q}} \in \mathcal{P}(E)$ solving (3.2). To do so, we introduce a function $\Psi : \mathcal{P}(E) \rightarrow \mathcal{P}(E)$ defined as follows. Let (X_0, Y_0) be a random variable distributed according to \mathcal{L}_{init} under \mathbb{P} . Then, we know from Proposition 5 that there exists a unique process $(X_t^{\mathbb{Q}}, Y_t^{\mathbb{Q}})_{t \in [0, T]}$ solving

$$\begin{cases} X_t^{\mathbb{Q}} = X_0 + \sum_{k, T^k \leq t} \mathbf{1} \left\{ U^k \leq \frac{\underline{f}}{\lambda \underline{f}} \frac{f(Y_{T^k-}^{\mathbb{Q}}) \lambda(T^k-, X_{T^k-}^{\mathbb{Q}})}{\varphi^{\mathbb{Q}}(T^k-, X_{T^k-}^{\mathbb{Q}})} \right\} \\ Y_t^{\mathbb{Q}} = Y_0 + \int_0^t b(s, X_s^{\mathbb{Q}}, Y_s^{\mathbb{Q}}) ds + \int_0^t \sigma(s, X_s^{\mathbb{Q}}, Y_s^{\mathbb{Q}}) dW_s + \int_0^t \gamma(s-, X_{s-}^{\mathbb{Q}}, Y_{s-}^{\mathbb{Q}}) dX_s^{\mathbb{Q}}, \end{cases} \quad (3.4)$$

and we define

$$\Psi(\mathbb{Q}) = \text{law}((X_t^{\mathbb{Q}}, Y_t^{\mathbb{Q}})_{t \in [0, T]}) \in \mathcal{P}(E).$$

As we will see in Section 3.4, Ψ (or more precisely Ψ iterated k -times) is a contraction mapping for the variation norm. Combining this result with the following Lemma gives the uniqueness of the probability measure solving (3.2).

Lemma 6. *Let $\overline{\mathbb{Q}} \in \mathcal{P}(E)$. We have $\Psi(\overline{\mathbb{Q}}) = \overline{\mathbb{Q}}$ if and only if $\overline{\mathbb{Q}}$ solves the Martingale Problem (3.2). In this case, $(X_t^{\overline{\mathbb{Q}}}, Y_t^{\overline{\mathbb{Q}}})_{t \in [0, T]}$ solves Equation (3.1).*

- Then, the existence and uniqueness of $\overline{\mathbb{Q}}$ satisfying $\Psi(\overline{\mathbb{Q}}) = \overline{\mathbb{Q}}$ and Lemma 6 automatically give the strong existence and uniqueness of (X, Y) satisfying (3.1) since we necessarily have $\mathbb{E}[f(Y_{T^k-})|X_{T^k-}] = \mathbb{E}_{\overline{\mathbb{Q}}}[f(Y_{T^k-})|X_{T^k-}]$ and thus $(X_t, Y_t)_{t \in [0, T]} = (X_t^{\overline{\mathbb{Q}}}, Y_t^{\overline{\mathbb{Q}}})_{t \in [0, T]}$. ■

Proof of Lemma 6. The direct implication is clear since $\Psi(\overline{\mathbb{Q}}) = \overline{\mathbb{Q}}$ gives that $\mathbb{E}[f(Y_t^{\overline{\mathbb{Q}}})|X_t^{\overline{\mathbb{Q}}}] = \varphi^{\overline{\mathbb{Q}}}(t, X_t^{\overline{\mathbb{Q}}})$. Thus, $(X_t^{\overline{\mathbb{Q}}}, Y_t^{\overline{\mathbb{Q}}})_{t \in [0, T]}$ solves (3.1) and in particular $\overline{\mathbb{Q}}$ solves the Martingale Problem (3.2).

Conversely, let us assume that $\overline{\mathbb{Q}}$ solves the Martingale Problem (3.2). We know from Proposition 5 that strong uniqueness holds for the SDE $(X_t^{\overline{\mathbb{Q}}}, Y_t^{\overline{\mathbb{Q}}})_{t \in [0, T]}$. From Lepeltier and Marchal (1976, Theorem II₁₃), we know that strong uniqueness implies weak uniqueness, which precisely gives $\Psi(\overline{\mathbb{Q}}) = \overline{\mathbb{Q}}$ since $\overline{\mathbb{Q}}$ solves the Martingale Problem (3.2) and therefore the Martingale problem associated with (3.4). In particular, we have $\mathbb{E}[f(Y_t^{\overline{\mathbb{Q}}})|X_t^{\overline{\mathbb{Q}}}] = \varphi^{\overline{\mathbb{Q}}}(t, X_t^{\overline{\mathbb{Q}}})$, and $(X_t^{\overline{\mathbb{Q}}}, Y_t^{\overline{\mathbb{Q}}})_{t \in [0, T]}$ solves (3.1). ■

3.2 The interacting particle system

We assume that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ carries all the random variables used below. Now, we set up the particle system related to the Martingale Problem (3.2). In the following, N will denote the number of particles, $(X_0^i, Y_0^i), i \in \mathbb{N}$ are independent random variables following the law \mathcal{L}_{init} under \mathbb{P} , and $(W_t^i, t \geq 0), i \in \mathbb{N}$ are independent standard Brownian motions. We build an interacting particle system $((X_t^{i, N}, Y_t^{i, N}), t \geq 0)_{1 \leq i \leq N}$ with the following features. For $i = 1, \dots, N$, $(X_t^{i, N}, t \geq 0)$ is a Poisson process with intensity:

$$\frac{\lambda(t-, X_{t-}^{i, N}) f(Y_{t-}^{i, N}) \sum_{j=1}^N \mathbf{1}_{\{X_{t-}^{j, N} = X_{t-}^{i, N}\}}}{\sum_{j=1}^N f(Y_{t-}^{j, N}) \mathbf{1}_{\{X_{t-}^{j, N} = X_{t-}^{i, N}\}}}, \quad (3.5)$$

and $Y_t^{i, N}$ solves the following equation:

$$Y_t^{i, N} = Y_0^i + \int_0^t b(s, X_s^{i, N}, Y_s^{i, N}) ds + \int_0^t \sigma(s, X_s^{i, N}, Y_s^{i, N}) dW_s^i + \int_0^t \gamma(s-, X_{s-}^{i, N}, Y_{s-}^{i, N}) dX_s^{i, N}. \quad (3.6)$$

In fact, we can give an explicit construction of this particle system, which will be useful later and we explain now. Let us consider $(U^{i, k})_{i, k \in \mathbb{N}^*}$ a sequence of independent uniform

variables on $[0, 1]$ and $(E^{i,k})_{i,k \in \mathbb{N}^*}$ a sequence of independent exponential random variables with parameter $\frac{\lambda f}{f}$. These variables are independent, and independent of the previously defined Brownian motions. We define the times

$$T^{i,k} = \sum_{l=1}^k E^{i,l},$$

and we can order $(T^{i,k}, i = 1, \dots, N, k \geq 1)$, such that $0 < T^{i_1, k_1} < \dots < T^{i_l, k_l} < \dots$ almost surely.

Up to the first jump of $X^{i,N}$, $Y_t^{i,N}$ is defined as the unique strong solution of

$$Y_t^{i,N} = Y_0^i + \int_0^t b(s, X_0^i, Y_s^{i,N}) ds + \int_0^t \sigma(s, X_0^i, Y_s^{i,N}) dW_s^i.$$

At time $\tau = T^{i_l, k_l}$, the process $X^{i_l, N}$ makes a jump of size 1 if

$$U^{i_l, k_l} \leq \frac{\frac{f}{\lambda f} \lambda(\tau-, X_{\tau-}^{i_l, N}) f(Y_{\tau-}^{i_l, N}) \sum_{j=1}^N \mathbf{1}_{\{X_{\tau-}^{j, N} = X_{\tau-}^{i_l, N}\}}}{\sum_{j=1}^N f(Y_{\tau-}^{j, N}) \mathbf{1}_{\{X_{\tau-}^{j, N} = X_{\tau-}^{i_l, N}\}}},$$

and does not jump otherwise. If a jump occurs, we set $Y_{\tau}^{i_l, N} = Y_{\tau-}^{i_l, N} + \gamma(\tau-, X_{\tau-}^{i_l, N}, Y_{\tau-}^{i_l, N})$ and, up to the next jump of $X^{i_l, N}$, we define $Y_t^{i_l, N}$ as the unique strong solution of

$$Y_t^{i_l, N} = Y_{\tau}^{i_l, N} + \int_{\tau}^t b(s, X_{\tau}^{i_l, N}, Y_s^{i_l, N}) ds + \int_{\tau}^t \sigma(s, X_{\tau}^{i_l, N}, Y_s^{i_l, N}) dW_s^i, \quad t \geq \tau.$$

3.3 Existence of a solution to (3.2)

We follow the analysis carried out by Méléard (1996), pages 69 and 70. We denote by $\mu^N = \frac{1}{N} \sum_{i=1}^N \delta_{(X_t^{i, N}, Y_t^{i, N})_{t \in [0, T]}}$ the empirical measure given by the particle system. It is a random variable taking values in $\mathcal{P}(E)$. We denote by $\pi^N \in \mathcal{P}(\mathcal{P}(E))$ the probability law of μ^N . For $\pi \in \mathcal{P}(\mathcal{P}(E))$, we denote by

$$I(\pi) = \int_{\mathcal{P}(E)} \mu \pi(d\mu) \in \mathcal{P}(E),$$

the mean of π . Let $F : E \rightarrow \mathbb{R}$ be a bounded function which is continuous with respect to the Skorokhod topology. It induces an application $\mathcal{P}(E) \rightarrow \mathbb{R}$ — still denoted by F with an abuse of notation — such that

$$F(\mu) = \int_E F(z) \mu(dz).$$

Since π^N is by definition the probability law of μ^N , we have by using Fubini's Theorem $\mathbb{E}(F(\mu^N)) = \int_{\mathcal{P}(E)} F(\mu) \pi^N(d\mu) = \int_E F(z) I(\pi^N)(dz)$. On the other hand, we have $F(\mu^N) = \frac{1}{N} \sum_{i=1}^N F((X_t^{i, N}, Y_t^{i, N})_{t \in [0, T]})$. By symmetry, $(X_t^{i, N}, Y_t^{i, N})_{t \in [0, T]}$ has the same law as $(X_t^{1, N}, Y_t^{1, N})_{t \in [0, T]}$ and we get that:

$$\mathbb{E}[F((X_t^{1, N}, Y_t^{1, N})_{t \in [0, T]})] = \int_E F(z) I(\pi^N)(dz). \quad (3.7)$$

Lemma 7. *The sequence $(\pi^N)_N$ is tight.*

The proof of Lemma 7 is postponed to Appendix D. Now, we can consider a subsequence π^{N_k} which converges weakly to $\pi^\infty \in \mathcal{P}(\mathcal{P}(E))$. Let $q \in \mathbb{N}^*$, $0 \leq s_1 \leq \dots \leq s_q \leq s \leq t$, $\phi \in \mathcal{C}_b^{0,2}(\mathbb{R}^2, \mathbb{R})$ and $g_1, \dots, g_q \in \mathcal{C}_b(\mathbb{R}^2, \mathbb{R})$ be bounded functions with bounded derivatives for ϕ . We set for $\mathbb{Q} \in \mathcal{P}(E)$,

$$F(\mathbb{Q}) = \mathbb{E}_{\mathbb{Q}} \left[\left(\phi(X_t, Y_t) - \phi(X_s, Y_s) - \int_s^t \left\{ \frac{\lambda(u, X_u) f(Y_u)}{\mathbb{E}_{\mathbb{Q}}[f(Y_u) | X_u]} (\phi(X_u + 1, Y_u + \gamma(u, X_u, Y_u)) - \phi(X_u, Y_u)) \right. \right. \right. \\ \left. \left. \left. + b(u, X_u, Y_u) \partial_y \phi(X_u, Y_u) + \frac{1}{2} \sigma^2(u, X_u, Y_u) \partial_y^2 \phi(X_u, Y_u) \right\} du \right) \prod_{l=1}^q g_l(X_{s_q}, Y_{s_q}) \right] \quad (3.8)$$

We have to check that $\mathbb{Q} \mapsto F(\mathbb{Q})$ is continuous with respect to \mathbb{Q} for the weak convergence. Since f is continuous by Assumption (1.1), we first notice that when \mathbb{Q}_n converges weakly to \mathbb{Q} , $\mathbb{E}_{\mathbb{Q}_n}[f(Y_t) | X_t = x]$ converges to $\mathbb{E}_{\mathbb{Q}}[f(Y_t) | X_t = x]$ when $\mathbb{Q}(X_t = x) > 0$, unless for an at most countable set of times t depending on \mathbb{Q} . Then, following Méléard (1996), it comes out that if s, t, s_1, \dots, s_q are taken outside a countable set depending on π^∞ , F is π^∞ -a.s. continuous. In this case, we have:

$$\mathbb{E}[F(\mu^{N_k})^2] \xrightarrow[k \rightarrow +\infty]{} \int_{\mathcal{P}(E)} F(\mathbb{Q})^2 \pi^\infty(d\mathbb{Q}).$$

By definition of μ^N , we have:

$$F(\mu^N) = \frac{1}{N} \sum_{i=1}^N \left[(M_t^{i,N} - M_s^{i,N}) \prod_{l=1}^q g_l(X_{s_q}^{i,N}, Y_{s_q}^{i,N}) \right], \text{ where} \\ M_t^{i,N} = \phi(X_t^{i,N}, Y_t^{i,N}) - \int_0^t \left\{ b(u, X_u^{i,N}, Y_u^{i,N}) \partial_y \phi(X_u^{i,N}, Y_u^{i,N}) + \frac{1}{2} \sigma^2(u, X_u^{i,N}, Y_u^{i,N}) \partial_y^2 \phi(X_u^{i,N}, Y_u^{i,N}) \right. \\ \left. + \frac{\lambda(u, X_u^{i,N}) f(Y_u^{i,N}) \sum_{j=1}^N \mathbf{1}_{\{X_u^{j,N} = X_u^{i,N}\}}}{\sum_{j=1}^N f(Y_u^{j,N}) \mathbf{1}_{\{X_u^{j,N} = X_u^{i,N}\}}} [\phi(X_u^{i,N} + 1, Y_u^{i,N} + \gamma(u, X_u^{i,N}, Y_u^{i,N})) - \phi(X_u^{i,N}, Y_u^{i,N})] \right\} du.$$

We observe now that $[M^{i,N}, M^{j,N}]_t = 0$ for $i \neq j$ since, by construction these martingales do not jump together and $\langle W^i, W^j \rangle_t = 0$. Therefore, we get that:

$$\mathbb{E}[F(\mu^N)^2] = \frac{1}{N} \mathbb{E} \left[(M_t^{1,N} - M_s^{1,N})^2 \prod_{l=1}^q g_l(X_{s_q}^{1,N}, Y_{s_q}^{1,N})^2 \right] \leq C/N,$$

thanks to the boundedness assumption made on functions g_l and ϕ . It comes out that $F(\mathbb{Q}) = 0$, $\pi^\infty(d\mathbb{Q})$ almost surely. This holds in fact for any function F given by (3.8), provided that s, t, s_1, \dots, s_q are taken outside a countable set depending on π^∞ . Since the process (X_t, Y_t) is càdlàg, this is sufficient to show that any measure in the support of π^∞ solves the Martingale Problem (3.2). In particular, we get the following result.

Proposition 8. *There exists a measure $\bar{\mathbb{Q}} \in \mathcal{P}(E)$ solving the Martingale Problem (3.2).*

3.4 Uniqueness of a solution to (3.2)

Let $\bar{\mathbb{Q}} \in \mathcal{P}(E)$ denote a probability measure solving the Martingale Problem (3.2). We know that such a probability exists thanks to Proposition 8. We want to show that it is indeed unique. To do so, we consider another probability $\mathbb{Q} \in \mathcal{P}(E)$ and study the total variation distance $V_T(\Psi(\bar{\mathbb{Q}}) - \Psi(\mathbb{Q}))$ between $\bar{\mathbb{Q}}$ and \mathbb{Q} over E . For $t \in [0, T]$, $\bar{\mathbb{Q}}, \mathbb{Q} \in \mathcal{P}(E)$, we denote by $\bar{\mathbb{Q}}|_{[0,t]}$ and $\mathbb{Q}|_{[0,t]}$ their restriction to the paths on the time interval $[0, t]$. We also set $V_t(\bar{\mathbb{Q}} - \mathbb{Q})$ the total variation distance between $\bar{\mathbb{Q}}|_{[0,t]}$ and $\mathbb{Q}|_{[0,t]}$.

Lemma 9. *Let $\tau := \inf\{t \geq 0, X_t^{\bar{\mathbb{Q}}} \neq X_t^{\mathbb{Q}}\}$ denote the first time when $X^{\bar{\mathbb{Q}}}$ and $X^{\mathbb{Q}}$ do not jump together. We have*

$$V_T(\Psi(\bar{\mathbb{Q}}) - \Psi(\mathbb{Q})) \leq 2\mathbb{P}(\tau \leq T).$$

Proof. Let us recall that for any signed measure η on E , the total variation of η is given by

$$V(\eta) = \eta^+(E) + \eta^-(E),$$

where $\eta = \eta^+ - \eta^-$ is the Hahn-Jordan decomposition of η . Besides, we clearly have

$$\frac{1}{2}V(\eta) \leq \tilde{V}(\eta) \leq V(\eta),$$

where $\tilde{V}(\eta) = \sup\{|\eta(A)|, A \subset E \text{ measurable}\}$.

We have for any measurable set A of E ,

$$\begin{aligned} & \mathbb{P}((X_t^{\bar{\mathbb{Q}}}, Y_t^{\bar{\mathbb{Q}}})_{t \in [0, T]} \neq (X_t^{\mathbb{Q}}, Y_t^{\mathbb{Q}})_{t \in [0, T]}) \\ & \geq \mathbb{P}((X_t^{\bar{\mathbb{Q}}}, Y_t^{\bar{\mathbb{Q}}})_{t \in [0, T]} \in A, (X_t^{\mathbb{Q}}, Y_t^{\mathbb{Q}})_{t \in [0, T]} \notin A) \\ & = \mathbb{P}((X_t^{\bar{\mathbb{Q}}}, Y_t^{\bar{\mathbb{Q}}})_{t \in [0, T]} \in A) - \mathbb{P}((X_t^{\bar{\mathbb{Q}}}, Y_t^{\bar{\mathbb{Q}}})_{t \in [0, T]} \in A, (X_t^{\mathbb{Q}}, Y_t^{\mathbb{Q}})_{t \in [0, T]} \in A) \\ & \geq \mathbb{P}((X_t^{\bar{\mathbb{Q}}}, Y_t^{\bar{\mathbb{Q}}})_{t \in [0, T]} \in A) - \mathbb{P}((X_t^{\mathbb{Q}}, Y_t^{\mathbb{Q}})_{t \in [0, T]} \in A). \end{aligned}$$

By taking the supremum over A , we get

$$\mathbb{P}((X_t^{\bar{\mathbb{Q}}}, Y_t^{\bar{\mathbb{Q}}})_{t \in [0, T]} \neq (X_t^{\mathbb{Q}}, Y_t^{\mathbb{Q}})_{t \in [0, T]}) \geq \tilde{V}_T(\Psi(\bar{\mathbb{Q}}) - \Psi(\mathbb{Q})),$$

which gives the claim. ■

Up to time τ , we have $(X_t^{\bar{\mathbb{Q}}}, Y_t^{\bar{\mathbb{Q}}}) = (X_t^{\mathbb{Q}}, Y_t^{\mathbb{Q}})$, and we get that

$$\mathbf{1}_{\{\tau \leq t\}} - \int_0^t \mathbf{1}_{\{\tau > s\}} \lambda(s, X_s^{\bar{\mathbb{Q}}}) f(Y_s^{\bar{\mathbb{Q}}}) \left| \frac{1}{\varphi^{\bar{\mathbb{Q}}}(s, X_s^{\bar{\mathbb{Q}}})} - \frac{1}{\varphi^{\mathbb{Q}}(s, X_s^{\mathbb{Q}})} \right| ds$$

is a martingale. In particular, we have

$$\begin{aligned} \frac{d\mathbb{P}(\tau \leq t)}{dt} &= \mathbb{E} \left[\mathbf{1}_{\{\tau > t\}} \lambda(t, X_t^{\bar{\mathbb{Q}}}) f(Y_t^{\bar{\mathbb{Q}}}) \left| \frac{1}{\varphi^{\bar{\mathbb{Q}}}(t, X_t^{\bar{\mathbb{Q}}})} - \frac{1}{\varphi^{\mathbb{Q}}(t, X_t^{\mathbb{Q}})} \right| \right] \\ &\leq \frac{\bar{\lambda}f}{f^2} \mathbb{E} \left[|\varphi^{\mathbb{Q}}(t, X_t^{\bar{\mathbb{Q}}}) - \varphi^{\bar{\mathbb{Q}}}(t, X_t^{\bar{\mathbb{Q}}})| \right]. \end{aligned}$$

Now, let us observe that $\varphi^{\overline{\mathbb{Q}}}(t, x) - \varphi^{\mathbb{Q}}(t, x) = \frac{\mathbb{E}_{\overline{\mathbb{Q}}}[f(Y_t)\mathbf{1}_{\{X_t=x\}}] - \mathbb{E}_{\mathbb{Q}}[f(Y_t)\mathbf{1}_{\{X_t=x\}}]}{\overline{\mathbb{Q}}(X_t=x)} + \frac{\mathbb{E}_{\mathbb{Q}}[f(Y_t)|X_t=x]}{\overline{\mathbb{Q}}(X_t=x)}[\mathbb{Q}(X_t = x) - \overline{\mathbb{Q}}(X_t = x)]$. Thus, we get that $|\varphi^{\overline{\mathbb{Q}}}(t, x) - \varphi^{\mathbb{Q}}(t, x)| \leq \frac{1}{\overline{\mathbb{Q}}(X_t=x)} \left(|\mathbb{E}_{\overline{\mathbb{Q}}}[f(Y_t)\mathbf{1}_{\{X_t=x\}}] - \mathbb{E}_{\mathbb{Q}}[f(Y_t)\mathbf{1}_{\{X_t=x\}}]| + \bar{f}|\mathbb{Q}(X_t = x) - \overline{\mathbb{Q}}(X_t = x)| \right)$, and therefore (we use here that $\overline{\mathbb{Q}}$ is invariant, and is the law of $(X_t^{\overline{\mathbb{Q}}}, Y_t^{\overline{\mathbb{Q}}})_{t \geq 0}$)

$$\begin{aligned} \mathbb{E}[|\varphi^{\mathbb{Q}}(t, X_t^{\overline{\mathbb{Q}}}) - \varphi^{\overline{\mathbb{Q}}}(t, X_t^{\overline{\mathbb{Q}}})|] &\leq \sum_{x \in \mathcal{L}_M} |\mathbb{E}_{\overline{\mathbb{Q}}}[f(Y_t)\mathbf{1}_{\{X_t=x\}}] - \mathbb{E}_{\mathbb{Q}}[f(Y_t)\mathbf{1}_{\{X_t=x\}}]| + \bar{f}|\mathbb{Q}(X_t = x) - \overline{\mathbb{Q}}(X_t = x)| \\ &\leq 2\bar{f}V_t(\Psi(\overline{\mathbb{Q}}) - \Psi(\mathbb{Q})). \end{aligned}$$

To check the last inequality, one has to observe that for a simple nonnegative function $f(y) = \sum_{i=1}^n f_i \mathbf{1}_{\{A_i\}}$, where $A_i \subset \mathbb{R}$ are Borel sets, we have $|\mathbb{E}_{\overline{\mathbb{Q}}}[f(Y_t)\mathbf{1}_{\{X_t=x\}}] - \mathbb{E}_{\mathbb{Q}}[f(Y_t)\mathbf{1}_{\{X_t=x\}}]| \leq \sum_{i=1}^n f_i |\overline{\mathbb{Q}}(X_t = x, Y_t \in A_i) - \mathbb{Q}(X_t = x, Y_t \in A_i)| \leq \bar{f}V_t(\overline{\mathbb{Q}} - \mathbb{Q})$. By passing to the limit, this property holds for any bounded measurable (hence continuous) function f .

To sum up, we have for $t \in [0, T]$,

$$V_T(\Psi(\overline{\mathbb{Q}}) - \Psi(\mathbb{Q})) \leq 2\mathbb{P}(\tau \leq T) \leq 4\frac{\bar{\lambda}\bar{f}^2}{\underline{f}^2} \int_0^T V_s(\overline{\mathbb{Q}} - \mathbb{Q}) ds \leq 4\frac{\bar{\lambda}\bar{f}^2}{\underline{f}^2} TV_T(\overline{\mathbb{Q}} - \mathbb{Q}),$$

since $V_0(\Psi(\overline{\mathbb{Q}}) - \Psi(\mathbb{Q})) = 0$ and $s \mapsto V_s(\overline{\mathbb{Q}} - \mathbb{Q})$ is nondecreasing. Let us recall that the set of bounded countably additive measures on E endowed with the total variation norm is a Banach space. If $4\frac{\bar{\lambda}\bar{f}^2}{\underline{f}^2}T < 1$, we get by the Banach fixed point theorem that Ψ has a unique fixed point that is necessarily $\overline{\mathbb{Q}}$. Otherwise, we get by iterating that:

$$V_T(\Psi^{(k)}(\overline{\mathbb{Q}}) - \Psi^{(k)}(\mathbb{Q})) \leq 4\frac{\bar{\lambda}\bar{f}^2}{\underline{f}^2} \int_0^T V_s(\Psi^{(k-1)}(\overline{\mathbb{Q}}) - \Psi^{(k-1)}(\mathbb{Q})) ds \leq \frac{\left[4\frac{\bar{\lambda}\bar{f}^2}{\underline{f}^2}T\right]^k}{k!} V_T(\overline{\mathbb{Q}} - \mathbb{Q}).$$

When k is large enough, $\frac{\left[4\frac{\bar{\lambda}\bar{f}^2}{\underline{f}^2}T\right]^k}{k!} < 1$ and $\Psi^{(k)}$ is a contraction mapping. Thus, $\Psi^{(k)}$ has a unique fixed point that is necessarily $\overline{\mathbb{Q}}$.

This concludes the proof of Theorem 4. Besides, using the notations of Section 3.3, we get that any convergent subsequence of π^N should converge to $\delta_{\overline{\mathbb{Q}}}(d\mathbb{Q})$. This gives the weak convergence of π^N towards $\delta_{\overline{\mathbb{Q}}}(d\mathbb{Q})$. By (3.7), we get that any particle converges in law towards $\overline{\mathbb{Q}}$.

3.5 Convergence speed towards $\overline{\mathbb{Q}}$

Now that we have proved that each particle converges to the invariant probability measure, we are interested in characterizing the speed of convergence of the interacting particle system towards this measure. This question is of practical importance, since one would like to use the following approximation

$$\mathbb{E}_{\overline{\mathbb{Q}}}[F((X_t, Y_t)_{t \in [0, T]})] \approx \frac{1}{N} \sum_{i=1}^N F((X_t^{i, N}, Y_t^{i, N})_{t \in [0, T]}), \quad (3.9)$$

and have an estimate of the error involved.

First, we need to introduce some additional notations. We consider the same particle system $(X_t^{i,N}, Y_t^{i,N})_{t \geq 0}$ as in Section 3.2, constructed with the random variables $T^{i,k}$, $U^{i,k}$ and $(W_t^i, t \geq 0)$. With these variables, we construct now the processes $(\bar{X}_t^i, \bar{Y}_t^i)_{t \geq 0}$ as the unique solution of

$$\begin{cases} \bar{X}_t^i = \bar{X}_0^i + \sum_{k, T^{i,k} \leq t} \mathbf{1} \left\{ U^{i,k} \leq \frac{f}{\lambda f} \frac{f(\bar{Y}_{T^{i,k}-}^i) \lambda(T^{i,k-}, \bar{X}_{T^{i,k}-}^i)}{\varphi^{\mathbb{Q}}(T^{i,k-}, \bar{X}_{T^{i,k}-}^i)} \right\} \\ \bar{Y}_t^i = \bar{Y}_0^i + \int_0^t b(s, \bar{X}_s^i, \bar{Y}_s^i) ds + \int_0^t \sigma(s, \bar{X}_s^i, \bar{Y}_s^i) dW_s^i + \int_0^t \gamma(s-, \bar{X}_{s-}^i, \bar{Y}_{s-}^i) d\bar{X}_s^i. \end{cases} \quad (3.10)$$

By construction, the law of $(\bar{X}_t^i, \bar{Y}_t^i)_{t \in [0, T]}$ is the invariant probability law $\bar{\mathbb{Q}}$, since $\Psi(\bar{\mathbb{Q}}) = \bar{\mathbb{Q}}$. By using the same argument as in Lemma 9, we have:

$$V_T(\mathcal{L}((X_t^{1,N}, Y_t^{1,N})_{t \in [0, T]}) - \bar{\mathbb{Q}}) \leq 2\mathbb{P}((X_t^{1,N}, Y_t^{1,N})_{t \in [0, T]} \neq (\bar{X}_t^1, \bar{Y}_t^1)_{t \in [0, T]}) = 2\mathbb{P}(\tau^1 \leq T),$$

where $\tau^1 = \inf\{t \geq 0, \bar{X}_t^1 \neq X_t^{1,N}\}$. We also set for $i = 2, \dots, N$, $\tau^i = \inf\{t \geq 0, \bar{X}_t^i \neq X_t^{i,N}\}$.

Proposition 10. *Let us assume that \mathcal{L}_{init} is such that $\mathbb{P}(X_0 = x) > 0$ for any $x \in \mathcal{L}_M$. Then, there is a constant $K > 0$ such that*

$$\mathbb{P}(\tau^1 \leq T) \leq \frac{K}{\sqrt{N}}.$$

Proof. By construction, the processes \bar{X}^1 and $X^{1,N}$ may become different at the times $T^{1,k}$ if $U^{1,k} \leq \frac{f}{\lambda f} \frac{\lambda(T^{1,k-}, \bar{X}_{T^{1,k}-}^1) f(\bar{Y}_{T^{1,k}-}^1)}{\varphi^{\mathbb{Q}}(T^{1,k-}, \bar{X}_{T^{1,k}-}^1)}$ and $U^{1,k} > \frac{f}{\lambda f} \frac{\lambda(T^{1,k-}, X_{T^{1,k}-}^{1,N}) f(Y_{T^{1,k}-}^{1,N})}{\sum_{j=1}^N f(Y_{T^{1,k}-}^{j,N}) \mathbf{1}_{\{X_{T^{1,k}-}^{j,N} = X_{T^{1,k}-}^{1,N}\}}}$, or conversely.

Thus, $\mathbf{1}_{\{\tau^1 \leq t\}} - \int_0^t \mathbf{1}_{\{\tau^1 > s\}} \frac{f}{\lambda f} \left| \frac{\lambda(s, \bar{X}_s^1) f(\bar{Y}_s^1)}{\varphi^{\mathbb{Q}}(s, \bar{X}_s^1)} - \frac{\lambda(s, X_s^{1,N}) f(Y_s^{1,N})}{\sum_{j=1}^N f(Y_s^{j,N}) \mathbf{1}_{\{X_s^{j,N} = X_s^{1,N}\}}} \right| ds$ is a martingale and we

have

$$\mathbb{P}(\tau^1 \leq T) = \frac{f}{\lambda f} \mathbb{E} \left[\int_0^T \mathbf{1}_{\{\tau^1 > t\}} \left| \frac{\lambda(t, \bar{X}_t^1) f(\bar{Y}_t^1)}{\varphi^{\mathbb{Q}}(t, \bar{X}_t^1)} - \frac{\lambda(t, X_t^{1,N}) f(Y_t^{1,N})}{\sum_{j=1}^N f(Y_t^{j,N}) \mathbf{1}_{\{X_t^{j,N} = X_t^{1,N}\}}} \right| dt \right].$$

Let us observe that on $\{\tau^1 > t\}$, we have $(\bar{X}_t^1, \bar{Y}_t^1) = (X_t^{1,N}, Y_t^{1,N})$. Therefore,

$$\mathbb{P}(\tau^1 \leq T) \leq \frac{f}{\lambda f} \mathbb{E} \left[\int_0^T \mathbf{1}_{\{\tau^1 > t\}} \left| \frac{1}{\varphi^{\mathbb{Q}}(t, \bar{X}_t^1)} - \frac{1 + \sum_{j=2}^N \mathbf{1}_{\{X_t^{j,N} = \bar{X}_t^1\}}}{f(Y_t^{1,N}) + \sum_{j=2}^N f(Y_t^{j,N}) \mathbf{1}_{\{X_t^{j,N} = \bar{X}_t^1\}}} \right| dt \right]. \quad (3.11)$$

Now, we study $\left| \frac{1}{\varphi^{\mathbb{Q}}(t, x)} - \frac{1 + \sum_{j=2}^N \mathbf{1}_{\{X_t^{j,N} = x\}}}{f(Y_t^{1,N}) + \sum_{j=2}^N f(Y_t^{j,N}) \mathbf{1}_{\{X_t^{j,N} = x\}}} \right|$ for $x \in \mathcal{L}_M$, and set

$$\bar{q}_t(x) = \bar{\mathbb{Q}}(X_t = x).$$

When $\bar{q}_t(x) > 0$, we have

$$\begin{aligned}
& \left| \frac{1}{\varphi^{\mathbb{Q}}(t, x)} - \frac{1 + \sum_{j=2}^N \mathbf{1}_{\{X_t^{j,N}=x\}}}{f(Y_t^{1,N}) + \sum_{j=2}^N f(Y_t^{j,N}) \mathbf{1}_{\{X_t^{j,N}=x\}}} \right| \leq \left| \frac{\bar{q}_t(x) - \frac{1}{N} \left(1 + \sum_{j=2}^N \mathbf{1}_{\{X_t^{j,N}=x\}} \right)}{\mathbb{E}_{\mathbb{Q}}[f(Y_t) \mathbf{1}_{\{X_t=x\}}]} \right| \\
& + \frac{1}{\underline{f} \mathbb{E}_{\mathbb{Q}}[f(Y_t) \mathbf{1}_{\{X_t=x\}}]} \left| \mathbb{E}_{\mathbb{Q}}[f(Y_t) \mathbf{1}_{\{X_t=x\}}] - \frac{1}{N} \left(f(Y_t^{1,N}) + \sum_{j=2}^N f(Y_t^{j,N}) \mathbf{1}_{\{X_t^{j,N}=x\}} \right) \right| \\
& \leq \frac{1}{\underline{f} \bar{q}_t(x)} \left| \bar{q}_t(x) - \frac{1}{N} \left(1 + \sum_{j=2}^N \mathbf{1}_{\{X_t^{j,N}=x\}} \right) \right| \\
& + \frac{1}{\underline{f} \bar{q}_t(x)} \left| \mathbb{E}_{\mathbb{Q}}[f(Y_t) \mathbf{1}_{\{X_t=x\}}] - \frac{1}{N} \left(f(Y_t^{1,N}) + \sum_{j=2}^N f(Y_t^{j,N}) \mathbf{1}_{\{X_t^{j,N}=x\}} \right) \right|.
\end{aligned}$$

We analyze these two terms in a similar manner. We introduce $(\bar{X}_t^{N+1}, \bar{Y}_t^{N+1})_{t \in [0, T]}$ another copy of $(\bar{X}_t^1, \bar{Y}_t^1)_{t \in [0, T]}$, which is independent from all other existing processes. We have:

$$\begin{aligned}
\left| \bar{q}_t(x) - \frac{1}{N} \left(1 + \sum_{j=2}^N \mathbf{1}_{\{X_t^{j,N}=x\}} \right) \right| & \leq \left| \bar{q}_t(x) - \frac{1}{N} \sum_{j=2}^{N+1} \mathbf{1}_{\{\bar{X}_t^j=x\}} \right| + \frac{1}{N} \left| \sum_{j=2}^N \mathbf{1}_{\{\bar{X}_t^j=x\}} - \mathbf{1}_{\{X_t^{j,N}=x\}} \right| + \frac{1}{N} \\
& \leq \left| \bar{q}_t(x) - \frac{1}{N} \sum_{j=2}^{N+1} \mathbf{1}_{\{\bar{X}_t^j=x\}} \right| + \frac{1}{N} \sum_{j=2}^N \mathbf{1}_{\{\tau^j \leq t\}} + \frac{1}{N}, \tag{3.12}
\end{aligned}$$

since \bar{X}_t^j and $X_t^{j,N}$ may be different only on $\{\tau^j \leq t\}$. Similarly, we have

$$\begin{aligned}
& \left| \mathbb{E}_{\mathbb{Q}}[f(Y_t) \mathbf{1}_{\{X_t=x\}}] - \frac{1}{N} \left(f(Y_t^{1,N}) + \sum_{j=2}^N f(Y_t^{j,N}) \mathbf{1}_{\{X_t^{j,N}=x\}} \right) \right| \\
& \leq \left| \mathbb{E}_{\mathbb{Q}}[f(Y_t) \mathbf{1}_{\{X_t=x\}}] - \frac{1}{N} \sum_{j=2}^{N+1} f(\bar{Y}_t^j) \mathbf{1}_{\{\bar{X}_t^j=x\}} \right| + \frac{\bar{f}}{N} \sum_{j=2}^N \mathbf{1}_{\{\tau^j \leq t\}} + \frac{\bar{f}}{N}. \tag{3.13}
\end{aligned}$$

We introduce

$$A(x) = \frac{1}{\bar{q}_t(x)} \left(\left| \bar{q}_t(x) - \frac{1}{N} \sum_{j=2}^{N+1} \mathbf{1}_{\{\bar{X}_t^j=x\}} \right| + \frac{1}{\underline{f}} \left| \mathbb{E}_{\mathbb{Q}}[f(Y_t) \mathbf{1}_{\{X_t=x\}}] - \frac{1}{N} \sum_{j=2}^{N+1} f(\bar{Y}_t^j) \mathbf{1}_{\{\bar{X}_t^j=x\}} \right| \right).$$

This is a random function which is independent from $(\bar{X}_t^1, \bar{Y}_t^1)_{t \geq 0}$. From (3.11), (3.12) and (3.13), we obtain by observing that τ^j and τ^1 have the same law under \mathbb{P} :

$$\mathbb{P}(\tau^1 \leq T) \leq \int_0^T \mathbb{E}[A(\bar{X}_t^1)] + \left(1 + \frac{\bar{f}}{\underline{f}} \right) \left(\frac{1}{N} \mathbb{E} \left[\frac{1}{\bar{q}_t(\bar{X}_t^1)} \right] + \mathbb{E} \left[\frac{1}{\bar{q}_t(\bar{X}_t^1)} \mathbf{1}_{\tau^2 \leq t} \right] \right) dt. \tag{3.14}$$

First, we observe that $\mathbb{E} \left[\frac{1}{\bar{q}_t(\bar{X}_t^1)} \right] = \sum_x, \text{ s.t. } \bar{q}_t(x) > 0 \frac{\bar{q}_t(x)}{\bar{q}_t(x)} \leq M + 1$. Thanks to the independence of $A(x)$ and \bar{X}_t^1 , $\mathbb{E}[A(\bar{X}_t^1) | \bar{X}_t^1 = x] = \mathbb{E}[A(x)]$. On the one hand, we have:

$$\mathbb{E} \left[\left| \bar{q}_t(x) - \frac{1}{N} \left(\sum_{j=2}^{N+1} \mathbf{1}_{\{\bar{X}_t^j=x\}} \right) \right| \right] \leq \sqrt{\mathbb{E} \left[\left(\bar{q}_t(x) - \frac{1}{N} \left(\sum_{j=2}^{N+1} \mathbf{1}_{\{\bar{X}_t^j=x\}} \right) \right)^2 \right]} = \frac{\sqrt{\bar{q}_t(x)(1-\bar{q}_t(x))}}{\sqrt{N}}.$$

On the other hand, we have similarly that:

$$\begin{aligned} & \mathbb{E} \left[\left| \mathbb{E}_{\bar{\mathbb{Q}}} [f(Y_t) \mathbf{1}_{\{X_t=x\}}] - \frac{1}{N} \left(\sum_{j=2}^{N+1} f(\bar{Y}_t^j) \mathbf{1}_{\{\bar{X}_t^j=x\}} \right) \right| \right] \\ & \leq \sqrt{\mathbb{E} \left[\left(\mathbb{E}_{\bar{\mathbb{Q}}} [f(Y_t) \mathbf{1}_{\{X_t=x\}}] - \frac{1}{N} \sum_{j=2}^{N+1} f(\bar{Y}_t^j) \mathbf{1}_{\{\bar{X}_t^j=x\}} \right)^2 \right]} \leq \sqrt{\frac{1}{N} \mathbb{E}_{\bar{\mathbb{Q}}} (\mathbf{1}_{\{X_t=x\}} f^2(Y_t))} \leq \sqrt{\frac{\bar{f}^2}{N} \bar{q}_t(x)}. \end{aligned}$$

Finally, we obtain that:

$$\mathbb{E}[A(x)] \leq \left(1 + \frac{\bar{f}}{\underline{f}} \right) \frac{1}{\sqrt{\bar{q}_t(x)} \sqrt{N}}.$$

By using the tower property of the conditional expectation, we get:

$$\mathbb{E}[A(\bar{X}_t^1)] \leq \left(1 + \frac{\bar{f}}{\underline{f}} \right) \frac{\sqrt{M+1}}{\sqrt{N}},$$

since we have $\sum_{x \in \mathcal{L}_M} \sqrt{\bar{q}_t(x)} \leq \sqrt{M+1}$ by the Cauchy-Schwarz inequality. From (3.14)

$$\mathbb{P}(\tau^1 \leq T) \leq \left(1 + \frac{\bar{f}}{\underline{f}} \right) \left(\frac{\sqrt{M+1}}{\sqrt{N}} + \frac{M+1}{N} + \int_0^T \mathbb{E} \left[\frac{1}{\bar{q}_t(\bar{X}_t^1)} \mathbf{1}_{\tau^2 \leq t} \right] dt \right). \quad (3.15)$$

So far, we have not used the assumption $\bar{q}_0(x) = \mathbb{P}(X_0 = x) > 0$ for any $x \in \mathcal{L}_M$. Since the jump intensity is bounded by $\frac{\bar{M}}{\underline{f}}$, we necessarily have $\bar{q}_t(x) \geq e^{-\frac{\bar{M}}{\underline{f}} T} \bar{q}_0(x)$, for $t \in [0, T]$. Thus, there exists a constant $C > 0$ (depending on T and \mathcal{L}_{init}) such that $\frac{1}{\bar{q}_t(x)} \leq C$ for $t \in [0, T]$, $x \in \mathcal{L}_M$. Since τ^2 and τ^1 have the same law under \mathbb{P} , this gives

$$\mathbb{P}(\tau^1 \leq T) \leq \left(1 + \frac{\bar{f}}{\underline{f}} \right) \left(\frac{\sqrt{M+1}}{\sqrt{N}} + \frac{M+1}{N} + C \int_0^T \mathbb{P}(\tau^1 \leq t) dt \right), \quad (3.16)$$

and we easily conclude by Gronwall's lemma. \blacksquare

Now, we can have an estimate of the accuracy given by the approximation (3.9). We assume that $F : E \rightarrow \mathbb{R}$ is a bounded measurable function. Then, we know by the central limit theorem that

$$\sqrt{N} \left(\frac{1}{N} \sum_{i=1}^N F((\bar{X}_t^i, \bar{Y}_t^i)_{t \in [0, T]}) - \mathbb{E}_{\bar{\mathbb{Q}}} [F((X_t, Y_t)_{t \in [0, T]})] \right) \xrightarrow{\text{law}} \mathcal{N}(0, \sigma^2),$$

where σ^2 is the variance of $F((X_t, Y_t)_{t \in [0, T]})$ under $\bar{\mathbb{Q}}$.

Since F is bounded by a constant $K > 0$, we have by Proposition 10

$$N^\alpha \left| \frac{1}{N} \sum_{i=1}^N F((\bar{X}_t^i, \bar{Y}_t^i)_{t \in [0, T]}) - \frac{1}{N} \sum_{i=1}^N F((X_t^{i, N}, Y_t^{i, N})_{t \in [0, T]}) \right| \leq \frac{K}{N^{1-\alpha}} \sum_{i=1}^N \mathbf{1}_{\{\tau^i \leq T\}},$$

which converges for the L^1 -norm to 0 when $N \rightarrow +\infty$ for $\alpha < 1/2$. Combining both results, we finally get a lower estimate of the convergence rate.

Corollary 11. *Under the assumptions of Proposition 10, we have for any bounded measurable function $F : E \rightarrow \mathbb{R}$ and any $0 < \alpha < 1/2$,*

$$N^\alpha \left| \frac{1}{N} \sum_{i=1}^N F((X_t^{i,N}, Y_t^{i,N})_{t \in [0, T]}) - \mathbb{E}_{\mathbb{Q}}[F((X_t, Y_t)_{t \in [0, T]})] \right| \rightarrow 0 \text{ in probability.}$$

Remark 12. *To prove Proposition 10, we have assumed that $\mathbb{P}(X_0 = x) > 0$ for any $x \in \mathcal{L}_M$. In fact, the same proof would work if we assumed that there existed $x_0 \in \mathcal{L}_M$ such that $\mathbb{P}(X_0 \geq x_0) = 1$ and $\mathbb{P}(X_0 = x) > 0$ for any $x \geq x_0$.*

However, in practice, it would have been nice to treat the case $\mathbb{P}(X_0 = x_0) = 1$ for some $x_0 \in \mathcal{L}_M$, since we know at the beginning how many firms have already defaulted. Heuristically from (3.15), we may hope to have for large N that $\mathbb{E} \left[\frac{1}{q_t(X_t^1)} \mathbf{1}_{\tau^2 \leq t} \right] \leq C \mathbb{P}(\tau^1 \leq t)$ since \bar{X}^1 and τ^2 are asymptotically independent, $\mathbb{E} \left[\frac{1}{q_t(X_t^1)} \right] \leq M + 1$, and τ^2 and τ^1 have the same law. This would be enough to conclude. Unfortunately, despite our investigations, we have not been able to prove this formally. However, we still observe a convergence speed of $1/\sqrt{N}$ when $X_0 = 0$ on our numerical experiments (Section 4).

4 Numerical results

In this section, we illustrate the theoretical results obtained in the previous sections. Let us recall that the local intensity (LI) model is a Markov chain with unit jumps occurring with the rate $\lambda(t-, x)$ and that the SLI model is given by equation (1.3). First, we highlight that the LI and SLI models have the same marginal distributions but different laws as processes. Second, we study the convergence of the interacting particle system and obtain numerical simulations showing a central limit theorem.

For our numerical experiments we consider two different models for the process Y described below and we assume that the local intensity λ and the function f are given by

$$\begin{aligned} \lambda(t, x) &= \bar{\lambda} \left(1 - \frac{x}{M} \right) \\ f(x) &= (x \vee \underline{f}) \wedge \bar{f} \end{aligned}$$

where the number of defaultable entities M will be taken equal to $M = 125$ from now on. This choice of $\lambda(t, x)$ corresponds to M independent default times with intensity $\frac{\bar{\lambda}}{M}$. We also assume that there is no default at the beginning, i.e. $X_0 = 0$.

1. In the framework proposed by Lopatin and Misirpashaev (2008), the process Y is a continuous process satisfying a CIR type SDE

$$dY_t = \kappa(\lambda(t, X_{t-}) - Y_t) + \sigma \sqrt{Y_t} dW_t \quad (4.1)$$

where $\kappa, \sigma > 0$. To sample such a process, we use the second-order discretization scheme for the CIR diffusion given in Alfonsi (2010).

2. In the framework of Arnsdorf and Halperin (2008), the process Y is no more continuous and may jump when a new default occurs. The process Y solves the following SDE

$$dY_t = -aY_t \log(Y_t) dt + \sigma Y_t dW_t + \gamma Y_{t-} dX_t \quad (4.2)$$

where $a, \sigma > 0$ and $\gamma \geq 0$. Remember that X only has positive jumps. For discretization purposes, note that between two jump times of X , $\log(Y)$ solves the following Ornstein Uhlenbeck SDE

$$dZ_t = (-aZ_t + \frac{1}{2}\sigma^2)dt + \sigma dW_t.$$

Even though we could have sampled exactly in this particular case the Gaussian increments of Z , we discretize Y using the Euler scheme on Z in our simulation since we would make this choice for more general SDEs on Y .

The LI model can be simulated very easily using a standard Monte-Carlo approach as it is sufficient to know how to sample a Poisson process with intensity λ ; we do not need any interacting particle system.

4.1 Practical implementation

In this part, we describe our implementation of the particle system to sample from the distribution of $(X_t, Y_t)_{0 \leq t \leq T}$. We recall that the process X has no continuous part and makes jumps of size 1. The process Y has a continuous part and may jump at the same times as X . We consider a regular time grid of $[0, T]$ with step size $h = \frac{T}{D}$: $s_k = kh$, $0 \leq k \leq D$. Assume we have already discretized Y up to time s_k , the discretization of Y at time s_{k+1} is built in the following way:

- If the process X does not jump between time s_k and time s_{k+1} we use the increment of a standard discretization scheme.
- If the process X jumps at time s with $s_k < s < s_{k+1}$, we proceed in three steps: apply the previous case between times s_k and s , integrate the jump at time s and finally apply the previous case again between time s and s_{k+1} .

This scheme ensures all the Y^i are at least discretized on the regular grid $\{s_0, s_1, \dots, s_D\}$.

Computational complexity. Studying the complexity is of prime importance when proposing a numerical algorithm.

On the one hand, there are D discretization times at which we recalculate the values of each Y^i , which requires $O(DN)$ operations. On the other hand, the average total number of proposed jump dates (ie. the number of steps in loop 5) is given by the expectation of the underlying Poisson process at time T : $NT \frac{\lambda \bar{f}}{\underline{f}}$. The computation cost of the ratio (4.3) is $O(N)$ and the complexity is $O(ND + N^2 T \frac{\lambda \bar{f}}{\underline{f}})$. Hence, for fixed model parameters, the overall complexity of our approach is bounded by $O(DN + N^2)$. In practice, N is much larger than D , and the most computationally demanding part of the algorithm is the numerical approximation of the condition expectation involved in the jump intensity.

The complexity of the interacting particle approach can be well improved if during the algorithm we keep track of the following two quantities involved in Equation (4.3)

$$D_t^j = \sum_{l=1}^N f(Y_{t-}^l) \mathbf{1}_{\{X_{t-}^l = j\}} \quad \text{and} \quad N_t^j = \sum_{l=1}^N \mathbf{1}_{\{X_{t-}^l = j\}} \quad j = 0, \dots, M. \quad (4.4)$$

Algorithm 1 (The naive particle system algorithm with complexity $O(DN + N^2)$.)

```

1:  $t = 0$  //Current date.
2:  $t_i = 0$  for all  $i = 1 \dots N$  //Last discretization date for the particle  $i$ .
3: Sample  $(X^i, Y^i)$  independently according to the initial law.
4:  $s_k = 0$  // Last date on the regular grid:  $s_k \leq t < s_{k+1}$ .
5: while  $t \leq T$  do
6:    $t' = t + \mathcal{E} \left( \frac{\lambda f}{f} N \right)$  // $t'$  is the potential next jump in the whole particle system.
7:   while  $t' > s_k + h$  do
8:      $s_k = s_k + h$ 
9:     for  $i = 1$  to  $N$  do
10:      Update the discretization of  $Y^i$  from time  $t_i$  to time  $s_k$ .
11:       $t_i = s_k$ 
12:    end for
13:     $t = s_k$ 
14:  end while
15:   $I =$  uniform r.v. with values in  $\{1, \dots, N\}$ . //Index of the particle which may jump.
16:  Compute the conditional expectation

```

$$E = \frac{\sum_{l=1}^N \mathbf{1}_{\{X^l = X^I\}}}{\sum_{l=1}^N f(Y^l) \mathbf{1}_{\{X^l = X^I\}}} \quad (4.3)$$

```

17:   $R = \frac{f}{\lambda f} \lambda(t', X^I) f(Y^I) E$  //Compute the acceptance ratio.
18:   $U =$  uniform r.v. in  $[0, 1]$ .
19:  if  $U < R$  then //We accept the jump.
20:    Discretize  $Y^I$  up to time  $t'$ .
21:     $Y^I = Y^I + \gamma(t', X^I, Y^I)$ 
22:     $X^I = X^I + 1$ 
23:     $t_I = t'$ 
24:  end if
25:   $t = t'$ 
26: end while

```

Since the processes X^l are unit-jump increasing these quantities clearly vanish for $j > \max_l X_{t-}^l$.

Let t be the last proposed jump time of the particle system. These two vectors can be easily updated at time t' which denotes the next possible jump time. If some ticks of the regular grid lie in $[t, t')$, we set t as the last discretization date in this interval and recompute vectors $(D_t^j)_j$ and $(N_t^j)_j$ using Equation (4.4). This happens D times in the algorithm and can be done with $O(M + N)$ operations as explained in Algorithm 2.

Case 1: If the proposed jump at time t' is not accepted, there is nothing to compute:

$$D_{t'}^j = D_t^j \quad \text{and} \quad N_{t'}^j = N_t^j.$$

Case 2: Otherwise, let I denote the index of the particle jumping at time t' , we use

$$\begin{cases} D_{t'}^j = D_t^j - f(Y_t^I); & D_{t'}^{j+1} = D_t^{j+1} + f(Y_{t'}^I) & \text{if } j = X_t^I, \\ N_{t'}^j = N_t^j - 1; & N_{t'}^{j+1} = N_t^{j+1} + 1 & \text{if } j = X_t^I, \\ N_{t'}^j = N_t^j; & D_{t'}^j = D_t^j & \text{otherwise.} \end{cases} \quad (4.5)$$

Algorithm 2 (The improved particle system algorithm with complexity $O(DN)$.)

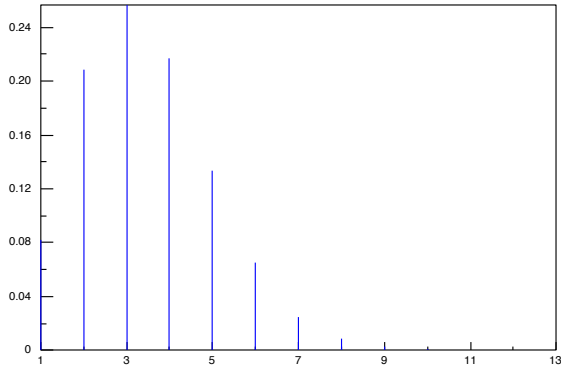
```

1:  $t = 0$ 
2:  $t_i = 0$  for all  $i = 1 \dots N$  //Last discretization date for the particle  $i$ .
3: Sample  $(X^i, Y^i)$  independently according to the initial law.
4: Set  $D^l = N^l = 0$  for  $0 \leq l \leq M$ .
5: for  $i = 1$  to  $N$  do //Calculate  $D$  and  $N$ . We can directly set  $N^0 = N$ ,  $D^0 = Nf(Y_0)$ ,
    $D^l = N^l = 0$  for  $1 \leq l \leq M$  when the initial law is a Dirac mass at  $(0, Y_0)$ .
6:    $N^{X^i} = N^{X^i} + 1$ ,  $D^{X^i} = D^{X^i} + f(Y^i)$ 
7: end for
8:  $s_k = 0$  //Last date on the regular grid.
9: while  $t \leq T$  do
10:    $t' = t + \mathcal{E}\left(\frac{\lambda f}{f} N\right)$  // $t'$  is the potential next jump in the whole particle system.
11:   while  $t' > s_k + h$  do
12:      $s_k = s_k + h$ 
13:     for  $i = 1$  to  $N$  do
14:       Update the discretization of  $Y^i$  from time  $t_i$  to time  $s_k$ .
15:        $t_i = s_k$ 
16:     end for
17:      $t = s_k$ 
18:     Reinitialize  $D^l = N^l = 0$  for  $0 \leq l \leq M$ .
19:     for  $i = 1$  to  $N$  do //Recalculate  $D$  and  $N$ .
20:        $N^{X^i} = N^{X^i} + 1$ ,  $D^{X^i} = D^{X^i} + f(Y^i)$ 
21:     end for
22:   end while
23:    $I =$  uniform r.v. with values in  $\{1, \dots, N\}$  //Index of the particle which may jump.
24:    $R = \frac{f}{\lambda f} \lambda(t', X^I) f(Y^I) \frac{N^{X^I}}{D^{X^I}}$  //Compute the acceptance ratio.
25:    $U =$  uniform r.v. in  $[0, 1]$ 
26:   if  $U < R$  then //We accept the jump.
27:      $N^{X^I} = N^{X^I} - 1$ ,  $D^{X^I} = D^{X^I} - f(Y^I)$ 
28:     Discretize  $Y^I$  up to time  $t'$ .
29:      $Y^I = Y^I + \gamma(t', X^I, Y^I)$ 
30:      $X^I = X^I + 1$ 
31:      $N^{X^I} = N^{X^I} + 1$ ,  $D^{X^I} = D^{X^I} + f(Y^I)$ 
32:      $t_I = t'$ 
33:   end if
34:    $t = t'$ 
35: end while

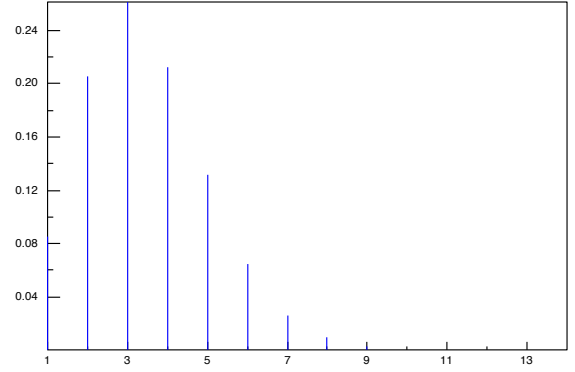
```

One should notice that in cases 2 and 3, the updating cost does not depend on the size of the vectors $(D_t^j)_j$ and $(N_t^j)_j$. Using these updating formulas, we can improve Algorithm 1 to obtain Algorithm 2. With this new algorithm we only have to compute D full approximations of the conditional expectation which has a unit a cost of $O(M + N)$ and the rest of the time we use the updating formulas (4.5), which happens on average less than $NT\frac{\bar{\lambda}\bar{f}}{\underline{f}}$ times. Then, the overall cost of this new algorithm is $O(D(M + N) + NT\frac{\bar{\lambda}\bar{f}}{\underline{f}})$. For fixed model parameters, this complexity reduces to $O(DN)$. This new algorithm has a linear cost with respect to the number of particles. Thus, we managed to propose an interacting particle algorithm with the same cost as a crude Monte–Carlo method for SDEs since M is in practice fixed and much smaller than N . The CPU times of the two algorithms are compared in the following examples.

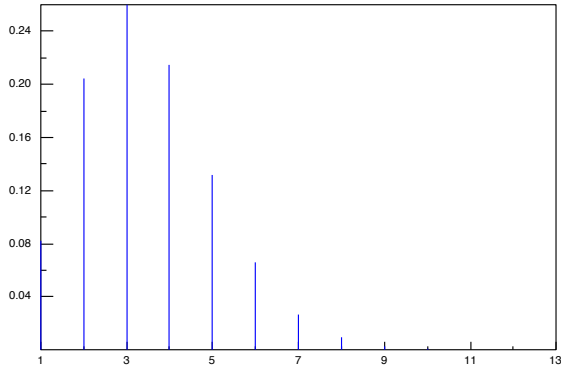
4.2 Marginal distributions



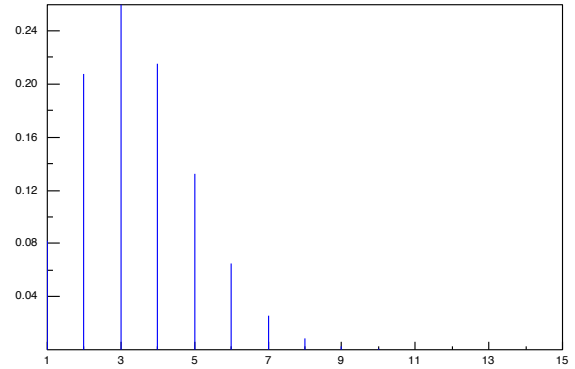
(a) Default distribution of the SLI model for Y given by Equation (4.2) with $T = 1, Y_0 = 1, a = 1, \sigma = 0.3, \gamma = 1, \bar{\lambda} = 2.5$.



(b) Default distribution of the SLI model for Y given by Equation (4.1) with $T = 1, Y_0 = 1, \kappa = 1, \sigma = 0.3, \bar{\lambda} = 2.5$.



(c) Default distribution of the LI model for $T = 1, \bar{\lambda} = 2.5$.



(d) Binomial distribution function with parameters $M, p = 1 - e^{-\bar{\lambda}T/M}$.

Figure 1: Comparison of the marginal distributions of the LI and SLI models. The simulations use $N = 50,000$ samples, $D = 100$ and $\underline{f} = 1/3, \bar{f} = 3$.

In Figures 1, we draw the probability distribution of X for the LI and SLI models for both processes Y considered in this part. The comparison of these graphs highlights how the LI and SLI models effectively mimic their marginal distributions; their probability distributions look almost the same. Since the LI model corresponds to independent default times with intensity $\frac{\bar{\lambda}}{M}$, the default distribution at time T is actually the Binomial law with parameters M and $p = 1 - \exp\left(-\frac{\bar{\lambda}}{M}T\right)$ as we can see on Figure 1(d).

	Algorithm 1	Algorithm 2
Model of Fig. 1(a)	239	1.51
Model of Fig. 1(b)	229	1.49

Table 1: Comparison of the CPU times (in seconds) of the two algorithms with the same parameters as in Figure 1.

We compare in Table 1 the computational times of the two algorithms and the gain obtained by the second approach is definitely outstanding. Algorithm 2 massively outperforms Algorithm 1 by a factor of 150. Of course, this gain will be all the more important as the number of particles increases.

Given the impressive match of the marginal distributions, we would like to numerically investigate the difference between their distributions as processes. To do so, we have computed in each model the length of the longest interval during which X does not jump defined by

$$\tau = \sup\{t \in [0, T] : \exists u \in [0, T - t], X_{u+t-} = X_u\} \quad (4.6)$$

Note that with this definition, $\tau = T$ when X does not jump on the interval $[0, T]$. The histograms of τ in the LI and SLI models are shown in Figure 2; we can see that, in the SLI, the length of the longest interval without jumps for X can be very small with a probability much higher than in the LI model (the l.h.s. of the histogram in the SLI model is fatter than in the LI model). This impression is reinforced by more quantitative observations. From the data used to plot these histograms, we have computed in Table 2 several values of the cumulative distribution function of the length of the longest interval without defaults both in the LI and SLI models. These quantities differ sufficiently to be numerically convinced that these two distributions do not match.

	$\mathbb{P}(\tau \leq T/4)$	$\mathbb{P}(\tau \leq T/8)$
SLI model	0.1911 (± 0.0033)	0.0200 (± 0.0012)
LI model	0.1645 (± 0.0033)	0.0113 (± 0.0009)

Table 2: Some values of the distribution function of the length of the longest interval without jumps for $T = 2$ and $\bar{\lambda} = 2.5$. The SLI model is defined as in Figure 2(a). The simulations use 50,000 samples. The values between braces correspond to twice the standard deviation of the estimator.

4.3 Convergence of the interacting particle system

When introducing the interacting particle system, we emphasized that it was not only a theoretical tool but that it was also of practical interest as it satisfies a strong law of large numbers. From a numerical point of view, the efficiency of the particle system depends on the

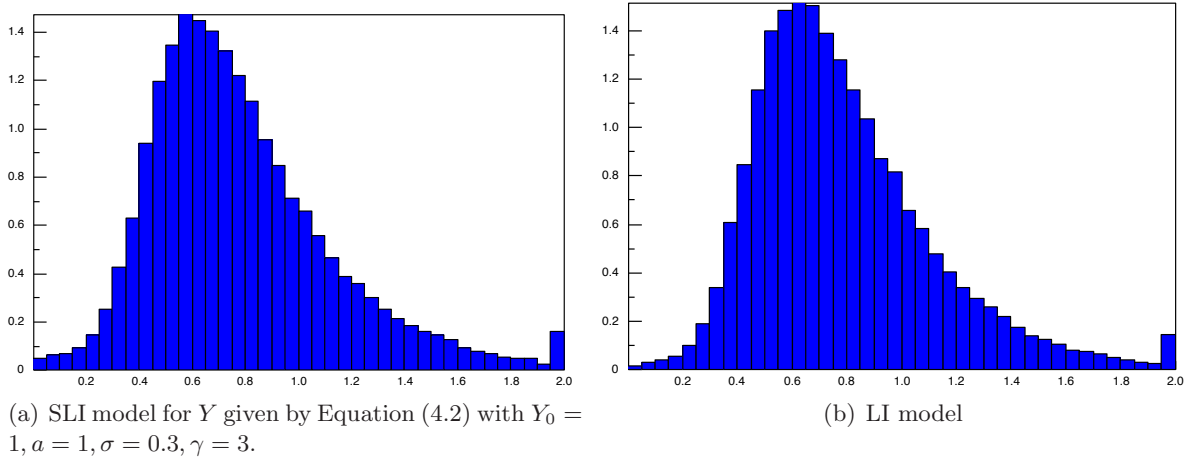


Figure 2: Histogram of the length of the longest interval without X jumping for $T = 2$ and $\bar{\lambda} = 2.5$. These histograms uses 50,000 samples.

rate at which every particle converges to the invariant probability (see Section 3.5). In that section, we proved that this convergence rate was faster than N^α for $0 < \alpha < 1/2$ where N is the number of particles. Now, we want to study this convergence rate in several examples: the first example only involves the marginal distribution of the particle system at maturity time, whereas the other two examples require the knowledge of the whole distribution and not only the marginal ones.

Number of defaults distribution. First, we start with a simple example. We want to study the convergence of the estimator of $\mathbb{P}(X_T = 3)$ computed on the particle system. We ran 5000 independent copies of the interacting particle systems and we computed the value of the estimator for each system. In Figure 3, we can see the centered and renormalised histogram of the values obtained for the empirical estimators. The histogram can be compared to the density of the standard normal distribution plotted as a solid line on Figure 3 and they match pretty well. This result suggests that a kind of central limit theorem should hold in practice even though we did not manage to prove such a result.

Asian option on the number of defaults. For our second example, we consider an Asian option on the default counting process whose price is given by

$$P = \mathbb{E} \left(\frac{1}{T} \int_0^T X_u du - K \right)_+$$

This price P will be approximated using the corresponding particle system estimate P_N , where N is the number of particles. We are interested in the limiting distribution of P_N . Because the process X has no continuous part and only makes jumps of size 1 and $X_0 = 0$, the pathwise integral can be rewritten

$$\frac{1}{T} \int_0^T X_u du = X_T - \frac{1}{T} \sum_{t \leq T \text{ s.t. } X_{t-} \neq X_t} t$$

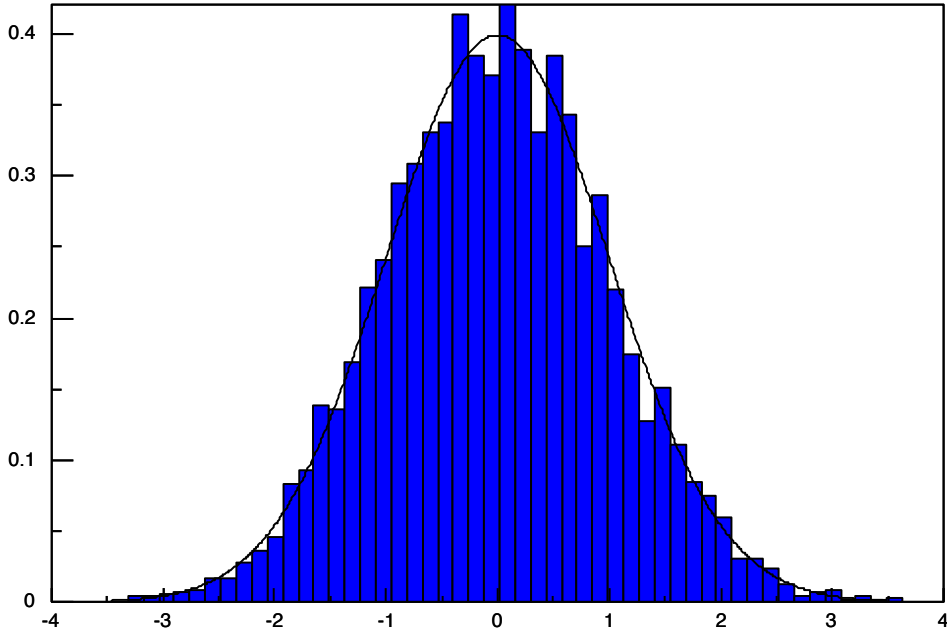


Figure 3: Centered and renormalised distribution of the estimation of $\mathbb{P}(X_T = 3)$ using the particle system when Y is given by Equation (4.2) with $T = 1, Y_0 = 1, a = 1, \sigma = 0.3, \gamma = 1, \bar{\lambda} = 2.5, \underline{f} = 1/3, \bar{f} = 3$ and $N = 10000$. The histogram was obtained using 5000 independent particle systems.

Hence, there is no need to approximate the integral, it can be computed exactly (up to the simulation of X). The example requires to sample the joint distribution of X_T and the sum of the default times.

On Figure 4, we have plotted the distribution of P_N after renormalizing and centering. As before, the solid line is the standard normal density. We can see that the limiting distribution looks very much like the Gaussian distribution, but such an histogram does not enable to determine the rate of convergence to the limiting distribution. Actually, the rate of convergence is given by the decrease rate of $\sqrt{\text{Var}(P_N)} = (\mathbb{E}(|P_N - P|^2))^{1/2}$. From a practical point of view we do not have access to P , so we have approximated it by the empirical mean \hat{P} of the data set used to build the histogram of Figure 4.

We can see on Figure 5 that the rate of decrease of $\sqrt{\mathbb{E}(|P_N - \hat{P}|^2)}$ recalls the shape of a negative power function. Then, we have decided to compute the linear regression of $-\frac{1}{2} \log \mathbb{E}(|P_N - \hat{P}|^2)$ with respect to $\log(N)$ on our simulations of P_N for N varying from 100 to 10,000 with a step size of 100, which gives a set of 100 data. The idea of the regression is to write

$$-\frac{1}{2} \log \mathbb{E}(|P_N - \hat{P}|^2) = \alpha \log(N) + \beta + \varepsilon_N$$

and to minimize the series $\sum_N \varepsilon_N^2$. The minimum is achieved for $\alpha = 0.5014, \beta = -0.1361$ and the empirical variance of the sequence (ε_n) is equal to 10^{-4} . This computation yields that the rate of convergence to the limiting distribution is \sqrt{N} . It ensues from this result combined with the analysis of the histogram 4 that a Central Limit Theorem with rate \sqrt{N} should hold.

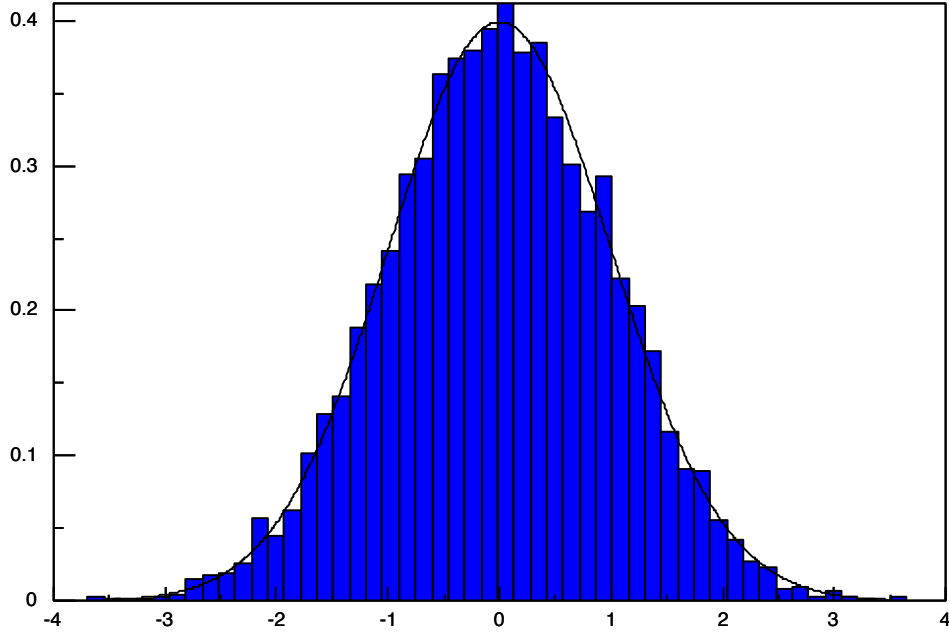


Figure 4: Centered and renormalised distribution of the estimator P_N of the Asian option price using the particle system when Y is given by Equation (4.1) with $T = 2, Y_0 = 1, a = 1, \sigma = 0.3, \kappa = 1, \underline{f} = 1/3, \bar{f} = 3$ and $N = 10000$. The histogram was obtained using 10,000 independent particle systems.

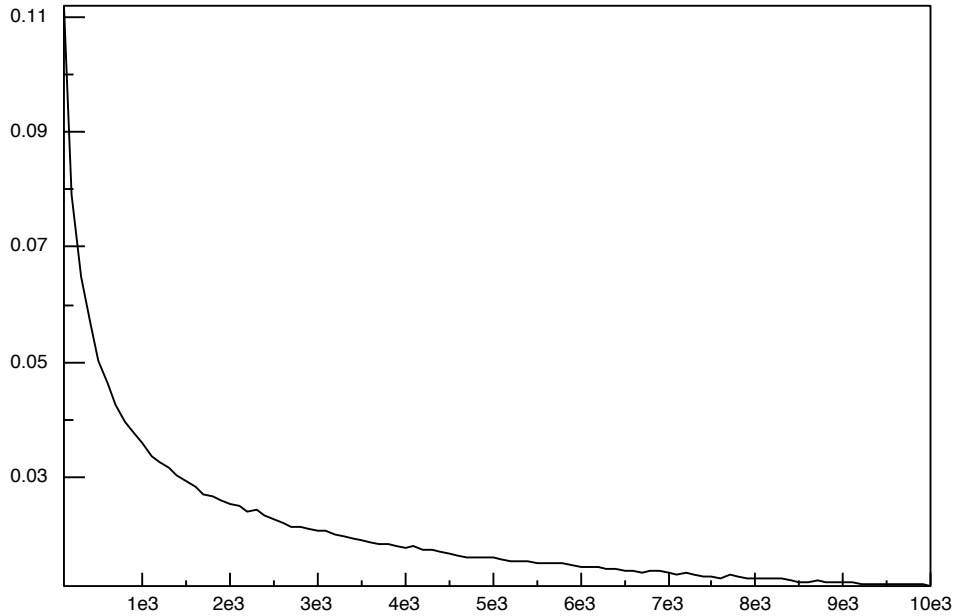


Figure 5: Convergence rate of $\sqrt{\mathbb{E}|P_N - \hat{P}|^2}$ w.r.t the number of particles N .

Longest interval without jump. In this paragraph, we are interested in the convergence rate of the estimator of the length of the longest time interval with no jump. We recall the quantity of interest already defined in Equation (4.6)

$$\tau = \sup\{t \in [0, T] : \exists u \in [0, T - t], X_{u+t^-} = X_u\}$$

and we consider its particle system estimator τ_N . To numerically sample from the distribution of τ , we need to know the joint distribution of the jumping times of X , which is a prime case of a pathwise estimator.

On Figure 6, we can see the centered and renormalized distribution of τ_N together with the standard normal density function plotted as a solid line. Again, the limiting distribution looks very much like a Gaussian distribution. Using these 5,000 independent particle systems, each with 10,000 particles, we can compute an approximation of τ , denote $\hat{\tau}$ in the following. Now, we can run independent simulations of particle systems with a number of particles N varying from 100 to 10,000. We study the rate of decrease of $\sqrt{\mathbb{E}|\tau_N - \hat{\tau}|^2}$, which according to Figure 7 shows a negative power function shape. If we linearly regress $\log \sqrt{\mathbb{E}|\tau_N - \hat{\tau}|^2}$ against $\log(N)$ we find that we can write

$$-\frac{1}{2} \log \mathbb{E}(|\tau_N - \hat{\tau}|^2) = \alpha \log(N) + \beta + \varepsilon_N$$

The linear regression yields $\alpha = 0.4865$, $\beta = 1.016$ and the empirical variance of the sequence (ε_n) is equal to 0.0004. This regression yields that the rate of convergence to the limiting distribution is \sqrt{N} , which focuses the existence of a Central Limit Theorem with rate \sqrt{N} .

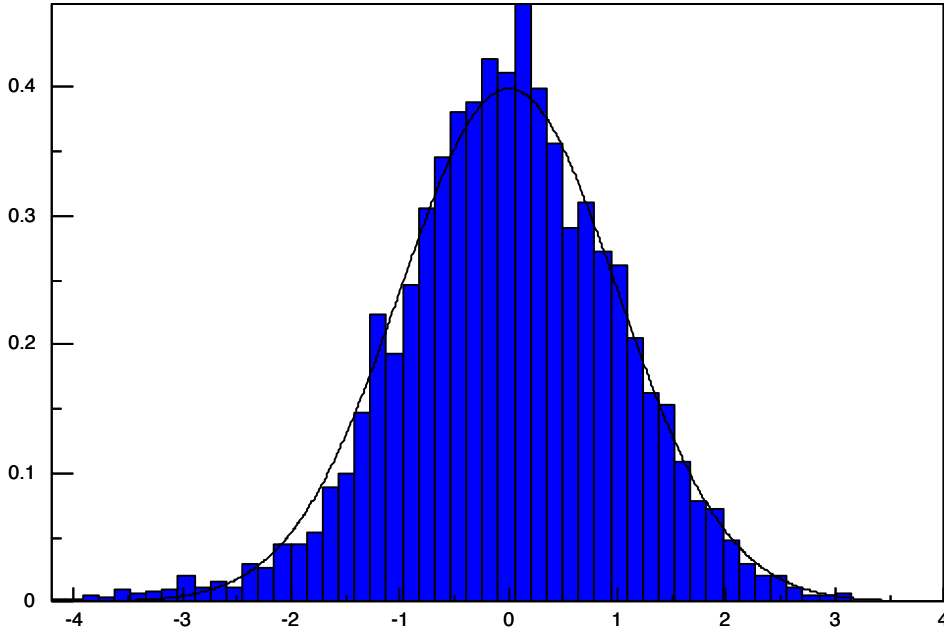


Figure 6: Centered and renormalised distribution of the estimator $\hat{\tau}$ of τ using the particle system when Y is given by Equation (4.1) with $T = 2, Y_0 = 1, a = 1, \sigma = 0.3, \kappa = 1, \underline{f} = 1/3, \bar{f} = 3$ and $N = 10,000$. The histogram was obtained using 10,000 independent particle systems.

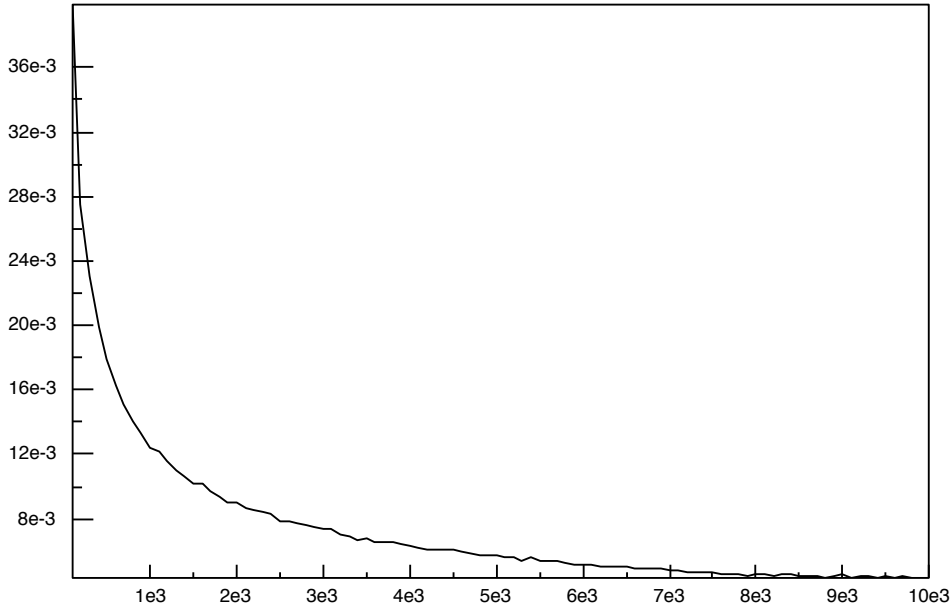


Figure 7: Convergence rate of $\sqrt{\mathbb{E}|\tau_N - \hat{\tau}|^2}$ w.r.t the number of particles N .

5 Conclusion

Local intensity models are wide spread for modelling the default counting process in credit risk. Recently, more sophisticated models with a stochastic factor involved in the intensity have been introduced in the literature. These stochastic local intensity models can be automatically calibrated to CDO tranche prices by properly choosing the local part of the intensity. This particular choice of the local intensity gives rise to a very specific family of SLI dynamics for which we have investigated the existence and uniqueness of solutions. This theoretical study has been carried out using particle systems, which turned out to be a clever tool for the numerical simulation of such dynamics. We have proved that particle Monte-Carlo algorithms based on this particle system approach almost surely converge. The theoretical study of the convergence rate enabled us to prove that the almost sure convergence took place at a rate faster than N^α for any $0 < \alpha < 1/2$. Obtaining a Central Limit Theorem type result for such particle systems remains an open question, even though we could highlight such a behaviour in all our simulations. Last, we have shown that the interacting particle system can be sampled with a computational cost in $O(ND)$, which is the same asymptotic cost as a Monte-Carlo algorithm for standard SDEs.

A Proof of Proposition 3

The scheme of the proof is the following:

- First, we prove that $\Psi_+ : t, x \mapsto \Psi(t, x_+)$ is globally Lipschitz in x . Then, (\mathcal{E}'') admits a unique solution on \mathbb{R}_+ .
- Second, we prove that the solution satisfies $\forall t \geq 0, P(t) \geq 0$ and $|P(t)| = 1$.

Step 1: Ψ_+ is globally Lipschitz. Let us prove that for all $(x, y) \in E \times E$ there exists a constant K such that $|\Psi_+(t, x) - \Psi_+(t, y)| \leq K|x - y|$. We have

$$|\Psi_+(t, x) - \Psi_+(t, y)| = \sum_{i \in \mathcal{L}_M} \sum_{j \geq 0} |(\Psi(t, x_+))_j^i - (\Psi(t, y_+))_j^i|.$$

Bounding this quantity boils down to bound $\sum_{i \in \mathcal{L}_M} \sum_{j \geq 0} ((A_1)_j^i + 2f(j)\lambda(t, i)(A_2)_j^i)$ where

$$(A_1)_j^i := \left| \sum_{k \geq 0} \mu_{kj}^i ((x_k^i)_+ - (y_k^i)_+) \right|,$$

$$(A_2)_j^i := \left| \frac{\sum_{l=0}^{\infty} (x_l^i)_+}{\sum_{l=0}^{\infty} f(l)(x_l^i)_+} (x_j^i)_+ - \frac{\sum_{l=0}^{\infty} (y_l^i)_+}{\sum_{l=0}^{\infty} f(l)(y_l^i)_+} (y_j^i)_+ \right|.$$

Bound for A_1 . We easily get from Hypothesis 1

$$\begin{aligned} \sum_{i \in \mathcal{L}_M} \sum_{j \geq 0} (A_1)_j^i &\leq \sum_{i \in \mathcal{L}_M} \sum_{k \geq 0} \sum_{j \geq 0} |\mu_{kj}^i| |(x_k^i)_+ - (y_k^i)_+| = \sum_{i \in \mathcal{L}_M} \sum_{k \geq 0} 2|\mu_{kk}^i| |(x_k^i)_+ - (y_k^i)_+| \\ &\leq 2 \sup_{i \in \mathcal{L}_M, k \in \mathbb{N}} |\mu_{kk}^i| |x - y|. \end{aligned}$$

Bound for A_2 . To bound it, we introduce $\pm \frac{\sum_{l=0}^{\infty} (y_l^i)_+}{\sum_{l=0}^{\infty} f(l)(x_l^i)_+} (x_j^i)_+$ and $\pm \frac{\sum_{l=0}^{\infty} (y_l^i)_+}{\sum_{l=0}^{\infty} f(l)(y_l^i)_+} (x_j^i)_+$ in order to split $(A_2)_j^i$ in three terms. Each of them is bounded in the following way

$$\begin{aligned} \left| \frac{\sum_{l=0}^{\infty} (x_l^i)_+}{\sum_{l=0}^{\infty} f(l)(x_l^i)_+} (x_j^i)_+ - \frac{\sum_{l=0}^{\infty} (y_l^i)_+}{\sum_{l=0}^{\infty} f(l)(x_l^i)_+} (x_j^i)_+ \right| &\leq \frac{(x_j^i)_+}{\sum_{l=0}^{\infty} f(l)(x_l^i)_+} \sum_{l=0}^{\infty} |(x_l^i)_+ - (y_l^i)_+|, \\ \left| \frac{\sum_{l=0}^{\infty} (y_l^i)_+}{\sum_{l=0}^{\infty} f(l)(x_l^i)_+} (x_j^i)_+ - \frac{\sum_{l=0}^{\infty} (y_l^i)_+}{\sum_{l=0}^{\infty} f(l)(y_l^i)_+} (x_j^i)_+ \right| &\leq \frac{\bar{f}}{\underline{f}} \frac{(x_j^i)_+}{\sum_{l=0}^{\infty} f(l)(x_l^i)_+} \sum_{l=0}^{\infty} |(x_l^i)_+ - (y_l^i)_+|, \\ \left| \frac{\sum_{l=0}^{\infty} (y_l^i)_+}{\sum_{l=0}^{\infty} f(l)(y_l^i)_+} (x_j^i)_+ - \frac{\sum_{l=0}^{\infty} (y_l^i)_+}{\sum_{l=0}^{\infty} f(l)(y_l^i)_+} (y_j^i)_+ \right| &\leq \frac{1}{\underline{f}} |(x_j^i)_+ - (y_j^i)_+|. \end{aligned}$$

Then, $\sum_{i \in \mathcal{L}_M} \sum_{j \geq 0} f(j)\lambda(t, i)(A_2)_j^i \leq \bar{\lambda}(1 + 2\frac{\bar{f}}{\underline{f}})|x - y|$. Combining bounds on A_1 and A_2 , we get $|\Psi_+(t, x) - \Psi_+(t, y)| \leq K|x - y|$, where $K = 2 \sup_{i \in \mathcal{L}_M, k \in \mathbb{N}} |\mu_{kk}^i| + 2\bar{\lambda} \left(1 + 2\frac{\bar{f}}{\underline{f}}\right)$.

Step 2: the solution is positive with norm 1. Let $P(t)$ denote the unique solution of (\mathcal{E}'') .

First, we prove that $\forall t \in \mathbb{R}_+, \forall (i, j) \in \mathcal{L}_M \times \mathbb{N}, P_j^i(t) \geq 0$.

$P_j^i(t)$ satisfies

$$\begin{cases} (P_j^i)'(t) = \sum_{k \neq j} \mu_{kj}^i (P_k^i(t))_+ + \mathbf{1}_{\{i \geq 1\}} \frac{\lambda(t, i-1) f(j)}{\varphi((P(t))_+, i-1)} (P_j^{i-1}(t))_+ - \left(\frac{\lambda(t, i) f(j) \mathbf{1}_{\{i \leq M-1\}}}{\varphi((P(t))_+, i)} - \mu_{jj}^i \right) (P_j^i(t))_+, \\ P_j^i(0) = \delta_{ix_0} \delta_{jy_0}. \end{cases}$$

We assume that there exists $\bar{t} \geq 0$ such that $P_j^i(\bar{t}) < 0$. We also introduce $u := \sup\{s \leq \bar{t} : P_j^i(s) = 0\}$. Then, we integrate the above equation between u and \bar{t} . We get

$$(P_j^i)(\bar{t}) = \int_u^{\bar{t}} \sum_{k \neq j} \mu_{kj}^i (P_k^i(s))_+ + \mathbf{1}_{\{i \geq 1\}} \frac{\lambda(s, i-1) f(j)}{\varphi((P(s))_+, i-1)} (P_j^{i-1}(s))_+ ds.$$

The l.h.s. is strictly negative whereas the r.h.s. is non negative. Then, $P_j^i(t) \geq 0$ for all $t \in \mathbb{R}_+$.

Second, we prove $\forall t \in \mathbb{R}_+, |P(t)| = 1$. Since P is non negative, $(|P(t)|)' = \sum_{i \in \mathcal{L}_M} \sum_{j \geq 0} (P_j^i)'(t)$. Moreover, $\sum_{i \in \mathcal{L}_M} \sum_{j \geq 0} (P_j^i)'(t) = 0$. Then, $|P(t)| = |P(0)| = 1$.

B Proof of Proposition 5

In fact, we can explicitly describe the solution of (3.3) as follows. We have $(X_0, Y_0) = (x_0, y_0)$. Then, up to the first jump of X , Y_t is necessarily defined as the unique strong solution of the following SDE

$$Y_t = y_0 + \int_0^t b(s, x_0, Y_s) ds + \int_0^t \sigma(s, x_0, Y_s) dW_s, \quad t \leq T.$$

The process $(X_t, t \geq 0)$ only increases by jumps of size 1, and has at most M jumps since $\lambda(t, M) = 0$. These jumps may only occur at the times T^n and a jump do occur at time T^n if the following condition is satisfied

$$U^n \leq \frac{f}{\lambda f} \frac{\lambda(T^n-, X_{T^n-}) f(Y_{T^n-})}{\varphi^{\mathbb{Q}}(T^n-, X_{T^n-})}.$$

Observe here that for any $x \in \mathcal{L}_M$, $t \mapsto \varphi^{\mathbb{Q}}(t, x)$ is càdlàg since the process (X, Y) is càdlàg under \mathbb{Q} , and $\varphi^{\mathbb{Q}}(T^n-, X_{T^n-})$ is well defined. If the jump occurs, we have

$$Y_{T^n} = Y_{T^n-} + \gamma(T^n-, X_{T^n-}, Y_{T^n-}),$$

and Y_t is defined up to the next jump of X as the unique strong solution of the SDE

$$Y_t = Y_{T^n} + \int_{T^n}^t b(s, X_{T^n}, Y_s) ds + \int_{T^n}^t \sigma(s, X_{T^n}, Y_s) dW_s, \quad T_n \leq t \leq T.$$

C Proof of $\mathbb{E}[\sup_{t \in [0, T]} |Y_t^{1, N}|^p] < \infty$

We assume that $Y^{1, N}$ satisfies (3.6) and $X^{1, N}$ is a Poisson process with intensity (3.5). Then

$$\begin{aligned} |Y_t^{1, N}|^p &\leq 4^{p-1} \left(y_0^p + \left| \int_0^t b(s, X_s^{1, N}, Y_s^{1, N}) ds \right|^p + \left| \int_0^t \sigma(s, X_s^{1, N}, Y_s^{1, N}) dW_s \right|^p \right. \\ &\quad \left. + \left| \int_0^t \gamma(s^-, X_{s^-}^{1, N}, Y_{s^-}^{1, N}) dX_s^{1, N} \right|^p \right) \\ &\leq 4^{p-1} \left(y_0^p + T^{p-1} \int_0^t |b(s, X_s^{1, N}, Y_s^{1, N})|^p ds + \sup_{u \leq t} \left| \int_0^u \sigma(s, X_s^{1, N}, Y_s^{1, N}) dW_s \right|^p \right. \\ &\quad \left. + M^{p-1} \int_0^t |\gamma(s^-, X_{s^-}^{1, N}, Y_{s^-}^{1, N})|^p dX_s^{1, N} \right), \end{aligned}$$

since $X^{1, N}$ jumps at most M times. The r.h.s being increasing w.r.t t , we can replace in the l.h.s. $|Y_t^{1, N}|^p$ by $\sup_{u \leq t} |Y_u^{1, N}|^p$. Burkholder-Davis-Gundy inequality yields to

$$\mathbb{E} \left[\left| \sup_{u \leq t} \int_0^u \sigma(s, X_s^{1, N}, Y_s^{1, N}) dW_s \right|^p \right] \leq C_p T^{p/2-1} \mathbb{E} \left[\int_0^t |\sigma(s, X_s^{1, N}, Y_s^{1, N})|^p ds \right].$$

On the other hand, $\mathbb{E} \left[\int_0^t |\gamma(s^-, X_{s^-}^{1, N}, Y_{s^-}^{1, N})|^p dX_s^{1, N} \right] \leq \frac{\bar{\lambda} T}{T} \int_0^t \mathbb{E} \left[|\gamma(s^-, X_{s^-}^{1, N}, Y_{s^-}^{1, N})|^p \right] ds$ since the jump intensity of $X^{1, N}$ is upper bounded by $\frac{\bar{\lambda} T}{T}$. Now since b , σ , and γ have a sub linear growth with respect to y (see Hypothesis (2)), we get

$$\mathbb{E}[\sup_{t \leq T} |Y_t^{1, N}|^p] \leq C \left(\int_0^T 1 + \mathbb{E}[\sup_{u \leq s} |Y_u^{1, N}|^p] ds \right),$$

which gives the result by Gronwall's lemma.

D Proof of Lemma 7

The proof of this lemma is done in two steps. First, we claim that $(\pi^N)_N$ is tight if and only if the sequence $((X_t^{1, N}, Y_t^{1, N})_{t \in [0, T]})_N$ is tight. To check this, we first notice that $\mathcal{P}(E)$ is a Polish space. By Proposition 4.6 in Méléard (1996), $(\pi^N)_N$ is tight if and only if $(I(\pi^N))_N$ is tight. Then Prohorov's Theorem (tightness is equivalent to sequential compactness) gives with equation (3.7) the claim.

Now, we must show that $((X_t^{1, N}, Y_t^{1, N})_{t \in [0, T]})_N$ is tight. We use Aldous' criterion.

First, we have to check that, for any $\varepsilon > 0$, there exists a constant K such that $\mathbb{P}(\sup_{t \in [0, T]} |X_t^{1, N}| + |Y_t^{1, N}| > K) \leq \varepsilon$. This is trivial since $X_t^{1, N}$ is bounded and $\mathbb{E}[\sup_{t \in [0, T]} |Y_t^{1, N}|] < \infty$ (see Appendix C). Second, we have to check that for any $\varepsilon > 0$, $\eta > 0$, there exist $\delta > 0$ and n_0 such that

$$\sup_{n \geq n_0} \sup_{\tau_1, \tau_2 \in \mathcal{T}_{[0, T]}, \tau_1 \leq \tau_2 \leq \tau_1 + \delta} \mathbb{P}(|(X_{\tau_2}^{1, N}, Y_{\tau_2}^{1, N}) - (X_{\tau_1}^{1, N}, Y_{\tau_1}^{1, N})| > \eta) < \varepsilon,$$

where $\mathcal{T}_{[0, T]}$ denotes the set of stopping times taking values in $[0, T]$. We take $|(x, y)| = \max(|x|, |y|)$ and we assume without loss of generality that $0 < \eta < 1$. For convenience, we

introduce $\nu_t^{i,N} = \sum_{k=1}^{\infty} \mathbf{1}_{\{T^{i,k} \leq t\}}$: this is a Poisson process with intensity $\frac{\lambda f}{f}$. We distinguish the two cases: $\nu_{\tau_2}^{1,N} \geq \nu_{\tau_1}^{1,N} + 1$ ($X^{1,N}$ jumps in $(\tau_1, \tau_2]$) and $\nu_{\tau_2}^{1,N} = \nu_{\tau_1}^{1,N}$ ($X^{1,N}$ does not jump in $(\tau_1, \tau_2]$). We have:

$$\begin{aligned} \mathbb{P}(|(X_{\tau_2}^{1,N}, Y_{\tau_2}^{1,N}) - (X_{\tau_1}^{1,N}, Y_{\tau_1}^{1,N})| > \eta) &\leq \mathbb{P}(\nu_{\tau_2}^{1,N} \geq \nu_{\tau_1}^{1,N} + 1) \\ &\quad + \mathbb{P}(|Y_{\tau_2}^{1,N} - Y_{\tau_1}^{1,N}| > \eta, \nu_{\tau_2}^{1,N} = \nu_{\tau_1}^{1,N}) := P_1 + P_2. \end{aligned}$$

Since $\mathbb{P}(\nu_{\tau_2}^{1,N} > \nu_{\tau_1}^{1,N}) \leq \mathbb{E}[\nu_{\tau_2}^{1,N} - \nu_{\tau_1}^{1,N}] = \frac{\lambda f}{f} \mathbb{E}[\tau_2 - \tau_1]$, we get $P_1 \leq \delta \frac{\lambda f}{f}$.

P_2 is bounded by $\mathbb{P}(\int_{\tau_1}^{\tau_2} b(s, X_s, Y_s) ds + \int_{\tau_1}^{\tau_2} \sigma(s, X_s, Y_s) dW_s > \eta)$. Using Markov's inequality, we get

$$P_2 \leq \frac{1}{\eta^2} \mathbb{E} \left[\left| \int_{\tau_1}^{\tau_2} b(s, X_s^{1,N}, Y_s^{1,N}) ds + \int_{\tau_1}^{\tau_2} \sigma(s, X_s^{1,N}, Y_s^{1,N}) dW_s \right|^2 \right].$$

Moreover

$$\begin{aligned} &\mathbb{E} \left[\left| \int_{\tau_1}^{\tau_2} b(s, X_s^{1,N}, Y_s^{1,N}) ds + \int_{\tau_1}^{\tau_2} \sigma(s, X_s^{1,N}, Y_s^{1,N}) dW_s \right|^2 \right] \\ &\leq 2 \left(\mathbb{E} \left[\left| \int_{\tau_1}^{\tau_2} b(s, X_s^{1,N}, Y_s^{1,N}) ds \right|^2 \right] + \mathbb{E} \left[\left| \int_{\tau_1}^{\tau_2} \sigma(s, X_s^{1,N}, Y_s^{1,N}) dW_s \right|^2 \right] \right). \end{aligned}$$

On the one hand, we have $\left| \int_{\tau_1}^{\tau_2} b(s, X_s^{1,N}, Y_s^{1,N}) ds \right|^2 \leq \delta \int_{\tau_1}^{\tau_2} C(1 + |Y_s^{1,N}|^2) ds$ for some constant $C > 0$ by Hypothesis 2. On the other hand, Burkholder-Davis-Gundy's inequality gives $\mathbb{E} \left[\left| \int_{\tau_1}^{\tau_2} \sigma(s, X_s^{1,N}, Y_s^{1,N}) dW_s \right|^2 \right] \leq C \mathbb{E} \left[\int_{\tau_1}^{\tau_2} 1 + |Y_s^{1,N}|^2 ds \right]$. Since $\mathbb{E}[\sup_{t \leq T} |Y_t^{1,N}|^2] < \infty$, we get $P_2 \leq \frac{C}{\eta^2} \delta$. Thus, combining the upper bounds on P_1 and P_2 , Aldous' criterion is satisfied for $\delta := \frac{\varepsilon}{\frac{\lambda f}{f} + \frac{C}{\eta^2}}$.

References

- F. Abergel and R. Tachet. A nonlinear partial integro-differential equation from mathematical finance. *Discrete Contin. Dyn. Syst.*, 27(3):907–917, 2010. doi: 10.3934/dcds.2010.27.907.
- C. Alexander and L. M. Nogueira. Stochastic local volatility. *Proc. of the 2nd IASTED International Conference*, pages 136–141, 2004.
- A. Alfonsi. High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Math. Comp.*, 79(269):209–237, 2010. doi: 10.1090/S0025-5718-09-02252-2.
- A. Alfonsi. An introduction to the multiname modelling in credit risk. *Recent Advancements in the Theory and Practice of Credit Derivatives*, Eds T. Bielecki, D. Brigo, F. Patras, Bloomberg Press., 2011.
- M. Arnsdorf and I. Halperin. BSLP: Markovian bivariate spread-loss model for portfolio credit derivatives. *J. Comput. Finance*, 12(2):77–107, 2008.

- R. Cont and A. Minca. Recovering portfolio default intensities implied by CDO quotes. *Mathematical Finance*, 23(1):94–121, 2013. doi: 10.1111/j.1467-9965.2011.00491.x.
- B. Dupire. Pricing with a smile. *Risk*, (7):18–20, 1994.
- J. Guyon and P. Henry-Labordère. The Smile Calibration Problem Solved. *SSRN eLibrary*, 2011.
- I. Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an Itô differential. *Probab. Theory Relat. Fields*, 71(4):501–516, 1986. doi: 10.1007/BF00699039.
- J.-P. Lepeltier and B. Marchal. Problème des martingales et équations différentielles stochastiques associées à un opérateur intégro-différentiel. *Ann. Inst. H. Poincaré Sect. B (N.S.)*, 12(1):43–103, 1976.
- A. V. Lopatin and T. Misirpashaev. Two-dimensional Markovian model for dynamics of aggregate credit loss. In *Econometrics and risk management*, volume 22 of *Adv. Econom.*, pages 243–274. Emerald/JAI, Bingley, 2008. doi: 10.1016/S0731-9053(08)22010-4.
- S. Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, volume 1627 of *Lecture Notes in Math.*, pages 42–95. Springer, Berlin, 1996. doi: 10.1007/BFb0093177.
- V. Piterbarg. Markovian Projection Method for Volatility Calibration. *SSRN eLibrary*, 2006.
- P. Schönbucher. Portfolio losses and the term structure of loss transition rates: A new methodology for the pricing of portfolio credit derivatives. *Working Paper*, 2005.
- A.-S. Sznitman. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991. doi: 10.1007/BFb0085169.