An Introduction to Particle Methods with Financial Applications

René Carmona, Pierre Del Moral, Peng Hu and Nadia Oudjane

Abstract The aim of this article is to give a general introduction to the theory of interacting particle methods, and an overview of its applications to computational finance. We survey the main techniques and results on interacting particle systems and explain how they can be applied to the numerical solution of a variety of financial applications such as pricing complex path dependent European options, computing sensitivities, pricing American options or numerically solving partially observed control and estimation problems.

1 Introduction

The growing field of Feynman-Kac expectations and related particle models is one of the most active contact points between probability theory and practical applications. The particle simulation techniques they suggest are also called sequential Monte Carlo methods in Bayesian statistics, and particle or genetic type filters in advanced signal processing. They are used to approximate a flow of probability measures with an increasing level of complexity. This class of probabilistic mod-

René Carmona
Bendheim Center for Finance, Dpt of Operations Research & Financial Engineering, Princeton University, Princeton, NJ 08544-5263, USA e-mail: rcarmona@princeton.edu

Pierre Del Moral
INRIA Bordeaux-Sud Ouest Center & Bordeaux Mathematical Institute Universit Bordeaux I, 351, cours de la Libération 33405 Talence cedex, France e-mail: pierre.del-moral@inria.fr

Peng Hu
INRIA Bordeaux-Sud Ouest Center & Bordeaux Mathematical Institute Universit Bordeaux I, 351, cours de la Libération 33405 Talence cedex, France e-mail: peng.hu@inria.fr

Nadia Oudjane
EDF R&D, Clamart, France, Université Paris 13 and FiME (Finance for Energy Market Research Centre (Dauphine, CREST, EDF R&D)), e-mail: nadia.oudjane@edf.fr

els includes conditional distributions of signals with respect to noisy and partial observations, non absorption probabilities in Feynman-Kac-Schrödinger models, Boltzmann-Gibbs measures, as well as conditional distributions of stochastic processes in critical regimes. For a thorough discussion on the application domains of interacting particle algorithms, we refer the reader to the first rigorous study of particle filters [30], the review article [40], the monograph [31], and the references therein.

Recently, these interacting particle techniques have been applied in several areas of finance. For instance, using the rare event interpretation of particle methods, R. Carmona, J. P. Fouque and D. Vestal proposed in [16] an interacting particle algorithm for the computation of the probabilities of simultaneous defaults in large credit portfolios. These developments for credit risk computation were then improved in the subsequent paper [15] by R. Carmona and S. Crépey, and by P. Del Moral and F. Patras in [44].

Following the particle filtering approach which is already widely used to estimate hidden Markov models, V. Genon-Catalot, T. Jeantheau and C. Laredo [60] introduced particle methods for the estimation of stochastic volatility models. This approach has been applied for filtering nonlinear and non-Gaussian models by R. Casarin [19], R. Casarin and C. Trecroci [20]. More recently, M. S. Johannes, N. G. Polson and J.R. Stroud [66] used a similar approach to filter latent variables such as the jump times and sizes in jump diffusion price models. Particle techniques can also be used for stochastic optimization as demonstrated by S. Ben Hamida and R. Cont who provide in [6] a new calibration algorithm allowing for the existence of multiple global minima. Finally, in [42], interacting particle methods were used to estimate backward conditional expectation for American option pricing.

In this review paper, we survey the main ideas behind the *particle technology* alluded to above, with illustrations from recent applications in computational finance. We tried to provide a synthetic picture of particle solutions to some estimation problems arising in mathematical finance. We adopted an informal style of presentation, focusing on the ideas rather than on their detailed rigorous mathematical justification.

The article is organized as follows. In the following section, we highlight the natural link between option prices and Feynman-Kac formula. Then in the third section, the main principles and results related to particle methods are recalled. Finally, we dedicate the last sections of this article to the application of these particle techniques to some specific financial problems: credit risk analysis, sensitivity computation, American option pricing and control and estimation of partially observed models.

2 Option prices and Feynman-Kac formula

The numerical pricing of European-style options has been extensively studied in the mathematical finance literature. It would be foolish to try to cover this subject in the

present paper. We refer the reader to I. Karatzas and S.E. Shreve's book [67], and the more focused account by Y. Achdou and O. Pironneau [2] for a sample of texts relevant to the present discussion. European option pricing is a standard numerical problem in finance, well suited to our interpretation of option prices in terms of Feynman-Kac formula.

2.1 Discrete time models

We first consider discrete time models (often called multi-period models by economists). Option prices are often given by Feynman-Kac formulas of the form

$$Q_{p,n}(f_n)(X_p) := \mathbb{E}\left(f_n(X_n) \prod_{p\le q<n} G_q(X_q) \,\middle|\, X_p \right), \tag{1}$$

with the terminal condition $Q_{n,n}(f_n)(x) = f_n(X_n)$. Here p is the time at which the price is computed, and n is the time of maturity of the option, $f = (f_n)_n$ is a space-time function, i.e. a function of time n and a variable in the state space E_n at time n of the possible values of the underlying interest $X = (X_n)_n$, $f_n(x)$ giving the payoff of the option at time n if the underlying interest has the value x, i.e. $X_n = x$. We will assume throughout that $X = (X_n)_n$ is a Markov chain, and we will denote by $M = (M_{p,n})_{p,n}$ its transition probability

$$M_{p,n}(x,dy) := \mathbb{P}(X_n \in dy | X_p = x), \qquad x \in E_p. \tag{2}$$

We assume that the state space E_n of the chain at time n can change with n. We shall use the simpler notation M when the Markov chain is time homogeneous, in which case $M_{p,n} = M_{n-p} = M^{n-p}$, and the state space E does not change with time. The major ingredients in the above equation are the Markov chain X and the space-time potential $G = (G_n)_n$, given for each n, by a non-negative measurable function on E_n. The chain X is usually constructed from random factors evolving in time and price series $S^j = (S^j_n)_n$ which gives the time evolution of the risky asset prices. We give some simple example below. To conform with the terminology of the particle models used to understand theoretically and implement numerically the Feynman-Kac formulas of the above type, we shall sometimes call X_n a *particle* at time n. Depending on the application under consideration, the role of the potential G_n will be to capture the discounting necessary in the computation of the price, or some constraints (like barriers) present in the indenture of the option, or the risk premium in the form of a pricing kernel, or even to force the particle to visit some parts of the space-time domain where rare events occur, in which case its financial interpretation will not be possible.

We give some simple examples to illustrate the versatility of formula (1).

European barrier option.

In this case, we assume that there is only one underlying stock whose time evolution is given by a Markov chain $S = (S_n)_n$, that $X_n = S_n$ is the price of this underlying stock at time n, that f_n gives the payoff function if the maturity is n, and that $K > 0$ is the strike of the option. If we assume that stochastic interest rates are given by a non-negative space-time function $r = (r_n)_n$ of the chain, and if we denote by $A = (A_n)$ the sequence of barrier sets A_p, then the price of the barrier option is given by the Feynman-Kac formula (1) with

$$f_n(X_n) = (X_n - K)_+ \quad \text{and} \quad G_q(X_q) = 1\!\!1_{A_q}(X_q)\, e^{-r_q(X_q)} \; . \tag{3}$$

Asian option

This example is important because it allows us to illustrate the use of the Feynman-Kac formula (1) when the chain X evolves on *path space*. Indeed, if we assume that $S = (S_n)_n$ is a Markov chain in a state space E giving the time evolution of the stock price on which the Asian option is written, at each time n we define X_n as the path from time $p = 0$ up to the current time $p = n$ of the underlying Markov chain. In other words:

$$X_n := (S_0, \ldots, S_n) \in E_n := E^{n+1}$$

and the payoff of the option can be written in the form

$$f_n(X_n) = (H_n(X_n) - K)_+ \; , \tag{4}$$

where $K > 0$ is the strike of the option and where, in the case of the one dimensional fixed strike Asian option ($E = \mathbb{R}$):

$$H_n(X_n) = \frac{1}{n+1} \sum_{p=0}^{n} S_p \; . \tag{5}$$

Notice that this formalism for the Asian option includes the case of plain European options if we take $H_n(X_n) = f_n(S_n)$. Notice also that, if we choose $K = 0$ in (4) and

$$H_n(X_n) = \frac{1}{n+1} \sum_{p=0}^{n} S_p - S_n \; ,$$

then we have the *floating strike* Asian option with a null price at the origin. Many other payoff functions on path space can be considered, including geometric means, *better-off* or *worse-off* lookback options related to the maximum or the minimum values of the historical asset prices.

Remark 1. Notice that Importance Sampling models can also be encapsulated in the Feynman-Kac formula (1). These stochastic sampling methods are *simple* change of

probability measures. They are often used in rare event simulation to make events with small occurrence probability less rare [31, 37].

2.2 Continuous time models

In continuous time finance, the stochastic factors and the underlying stock prices are often given by diffusion models, and the reference Markov chain sequence S or X often results from a discretization procedure, such as those given by Euler or Milshtein schemes. For instance, let us suppose we are given an $\mathbb{R}^d$-valued Itô stochastic differential equation

$$dS_t^c = b(S_t^c)\, dt + \sigma(S_t^c)\, dW_t \ , \tag{6}$$

with some initial random vector $S_0^c \in \mathbb{R}^d$ with distribution $\eta_0 = \mathrm{Law}(S_0^c)$. Here, $W = (W_t)_{t\geq 0}$ is a standard d-dimensional Wiener process, and for any $x \in \mathbb{R}^d$, $\sigma(x) = (\sigma_{i,j}(x))_{1\leq i,j\leq d}$ is a $d \times d$ symmetric nonnegative definite matrix, and $b(x) = (b_i(x))_{1\leq i\leq d}$ a d-dimensional vector. The Euler discretization scheme over the regular time subdivision (also called time grid) $(t_n)_{n\geq 0}$, with the mesh $(t_n - t_{n-1}) = \Delta > 0$ is given by

$$S_n - S_{n-1} = b\,(S_{n-1})\ \Delta + \sigma\,(S_{n-1})\ \left(W_{t_n} - W_{t_{n-1}}\right)\ . \tag{7}$$

The elementary Markov transition

$$M(x,dy) := \mathbb{P}\,(S_n \in dy | S_{n-1} = x)$$

(the time subscripts are not needed because of the time homogeneity of the chain) can alternatively be defined in the integral form on bounded test functions as below

$$M(f)(x) := \int M(x,dy)\ f(y) = \mathbb{E}\left(f\left(x + b(x)\Delta + \sigma(x)\sqrt{\Delta}\ Y\right)\right)\ , \tag{8}$$

where $Y = (Y^i)_{1\leq i\leq d}$ is a sequence of independent and centred Gaussian random variables with unit variance.

In the same vein, suppose that the evolution of the underlying prices is given by a jump type Markov process S^c which evolves between jumps times T_n as in (6) the jump times T_n being defined in terms of a sequence $(e_n)_{n\geq 1}$ of independent and identically exponentially distributed random variables with unit parameter by the following recursion

$$T_n = \inf\left\{ t \geq T_{n-1}\ :\ \int_{T_{n-1}}^{t} \lambda\,(S_u)\ du \geq e_n \right\}\ , \tag{9}$$

with $T_0 = 0$ and some non negative function λ. At the time T_n of a jump, the process jumps from $S^c_{T_n-}$ to a new location $S^c_{T_n}$ randomly chosen with distribution $P(S^c_{T_n-}, dy)$ where $P(x,dy)$ is a given Markovian transition kernel.

A discrete time approximation model S_n is defined as above by replacing the transition M in (8), by the Markov transition MJ such that

$$(MJ)(x,dz) := \int M(x,dy)\, J(y,dz)\ ,$$

with the geometric jump type Markov transition

$$J(y,dz) = e^{-\lambda(y)\Delta}\ \delta_y(dz) + \left(1 - e^{-\lambda(y)\Delta}\right)\ P(y,dz)\ .$$

If we revisit the example of the barrier option for the sake of illustration, for time homogeneous barrier regions $A_n = A$, and non-negative stochastic interest rates $(R(S_t))_{t\geq 0}$ given by a function R on $\mathbb{R}$, if we set $r_n(x) = R_{t_n}(x)\Delta$ and $X = S$ in (3), then formula (1) gives a Δ-approximation of the continuous time model

$$\mathbb{E}\left(f_{t_n}(S^c_{t_n})\ 1\!\!1_{T\geq t_n}\ \exp\left\{-\int_{t_p}^{t_n} R_s(S^c_u)du\right\}\Big|S^c_{t_p} = x\right)\ ,$$

where T stands for the first time the process S gets out of the barrier region A.

3 Interacting particle approximations

In this section, we present a brief introduction to interacting particle methods as they pertain to the computation of the Feynman-Kac expectations discussed in the previous section. These advanced stochastic techniques are becoming increasingly popular in economics as well as in finance. A detailed survey to this field can be found in [27, 39].

3.1 Feynman-Kac semigroups

First, we notice that the integral operators $Q_{p,n}$ defined in (1) can be interpreted as the linear semigroup associated with the flow of non negative measures γ_n whose values on test functions f_n are given by:

$$\gamma_n(f_n) := \int \gamma_n(dx)\ f_n(x) = \mathbb{E}\left(f_n(X_n)\ \prod_{0\leq q<n} G_q(X_q)\right)\ . \tag{10}$$

The operators $Q_{p,n}$ were defined in (1) through their action on functions. Letting them act on measures by duality we get:

$$\gamma_n(dy) = (\gamma_p Q_{p,n})(dy) := \int\ \gamma_p(dx)\ Q_{p,n}(x,dy)\ ,$$

and for $0 \leq p \leq q \leq n$ we have the semigroup property

$$Q_{p,n}(x,dz) = (Q_{p,q}Q_{q,n})(x,dz) := \int Q_{p,q}(x,dy)\; Q_{q,n}(y,dz)\;.$$

Using these formulas in numerical implementations requires extensive calculations due to the fact that the total mass of the measures γ_n obtained by choosing the constant function $f_n(x) \equiv 1$ in (10) is very costly to compute with a reasonable precision. To illustrate this assertion, let us suppose that $G_p = 1\!\!1_A$, for any $p \leq n$. Then, the total mass $\gamma_n(1\!\!1)$ coincides with the probability that the trajectories of the Markov chain X stay in the set A for all times:

$$\gamma_n(1\!\!1) = \mathbb{E}\left(\prod_{0\leq q<n} G_q(X_q)\right) = \mathbb{P}\left(X_p \in A \;\; 0 \leq p < n\right),$$

which is, in most cases, difficult to compute. One natural way to resolve this estimation problem is to work with the normalized distributions η_n defined by:

$$\eta_n(f_n) := \gamma_n(f_n)/\gamma_n(1\!\!1)\;. \tag{11}$$

which should be a reasonable alternative since the original unnormalized measures can be recovered from the normalized ones with the following easily checked multiplicative formula:

$$\gamma_n(f_n) = \eta_n(f_n) \times \prod_{0\leq p<n} \eta_p(G_p)\;. \tag{12}$$

The second key observation is that the normalized distributions η_n satisfy the following recursive equation giving a nonlinear transition in η_{n-1}:

$$\eta_n(dy) = \left(\eta_{n-1}K_{n,\eta_{n-1}}\right)(dy) = \int K_{n,\eta_{n-1}}(x,dy)\eta_{n-1}(dx)\;, \tag{13}$$

where for each probability measure η on E_{n-1}, the Markovian transition kernel $K_{n,\eta}$ on E_{n-1} is defined by

$$K_{n,\eta}(x,dz) = \int \mathscr{S}_{n-1,\eta}(x,dy)M_{n-1,n}(y,dz)\;, \tag{14}$$

where in the above displayed formula, $\mathscr{S}_{n-1,\eta}$ is the selection-jump type Markov transition defined by

$$\mathscr{S}_{n-1,\eta}(x,dy) = G_{n-1}(x)\;\delta_x(y) + (1-G_{n-1}(x))\;\Psi_{G_{n-1}}(\eta)(dy)\;, \tag{15}$$

with the Boltzmann-Gibbs transformation

$$\Psi_g(\eta)(dy) = \frac{g(y)}{\eta(g)}\;\eta(dy)\;. \tag{16}$$

Remark 2. It is instructive, and in fact crucial given the use of the above result in the next subsection, to understand the effect of this Boltzmann-Gibbs transformation (16) in the case of point measures. Indeed, in this case:

$$\eta = \sum_{i=1}^{N} \alpha_i \delta_{x_i} \;\hookrightarrow\; \psi_g(\eta) = \sum_{i=1}^{N} \beta_i \delta_{x_i}$$

where the new weights β_i are given by:

$$\beta_i = \frac{\alpha_i g(x_i)}{\sum_{j=1}^{N} \alpha_j g(x_j)}, \qquad i = 1, \cdots, N$$

which shows that de facto, the Boltzmann-Gibbs transformation is a *resampling with replacement* of the x_i's according to the weights $\alpha_i g(x_i)$ given by the original weights and the function g. This interpretation will be extremely important for Monte Carlo implementation purposes.

Remark 3. Formulas (13) and (14) show that the passage from η_{n-1} to η_n is done in two steps. The individual particles $x \in E_{n-1}$ distributed as η_{n-1}, are first moved into dy according to the transition $\mathscr{S}_{n-1,\eta_{n-1}}(x,dy)$. This is a *selection* since (15) says that the particle remains at x with probability $G_{n-1}(x)$, and with probability $1 - G_{n-1}(x)$ it is chosen at random (independently of its current position $x \in E_{n-1}$) according to the distribution $\Psi_{\eta_{n-1}}(\eta_{n-1})(dy)$. The resulting particles $y \in E_{n-1}$ are then *mutated* into particle $z \in E_n$ according to the transition $M_{n-1,n}$ of the original Markov chain $X = (X_n)_n$. The interpretation of the selection step will be crystal clear when we implement it in for probability distributions with finite supports which we will interpret as empirical distributions of particle systems.

Remark 4. Note that the above interpretation is not limited to $[0,1]$-valued potential functions G as long as G is non-negative and bounded, and as long as we replace G_n by $\varepsilon_n G_n$ in (15) and (16) with ε_n such that $\varepsilon_n G_n \in [0,1]$.

3.2 Interacting particle methodologies

We now revisit the measure flows of the previous subsection in the case of point measures given by the empirical distribution of a fixed but large number N of *particles*. Let $\underline{\xi} := (\xi_n)_{n\geq 0}$ be a Markov chain with product E_n^N as state space at time n. So at each time n, ξ_n is an N-tuple $\xi_n := (\xi_n^i)_{1\leq i\leq N}$. We assume that the transition probability of this chain is given by:

$$\mathbb{P}\left(\xi_n \in dx^1 \times \cdots \times x^N \,|\xi_{n-1}\right) = \prod_{1\leq i\leq N} K_{n,\eta_{n-1}^N}(\xi_{n-1}^i, dx^i)\,, \tag{17}$$

where η_{n-1}^N denotes the empirical measure of the components of ξ_{n-1}:

$$\eta_{n-1}^N := \frac{1}{N}\sum_{i=1}^N \delta_{\xi_{n-1}^i} .$$

We assume that the initial law η_0^N is a product distribution of the form $\eta_0^N = (\eta_0 \times \cdots \times \eta_0)$, or in other words that the initial system $\xi_0 = (\xi_0^i)_{1\le i\le N}$ consists of N independent and identically distributed random variables ξ_0^i with common law η_0. The transition mechanism of the chain $\underline{\xi}$ depends only upon the empirical distribution of the components of its state, not the actual values of these components. Indeed, given the empirical distribution η_{n-1}^N of the ξ_{n-1}^i's, these ξ_{n-1}^i's evolve independently of each other, each ξ_{n-1}^i moving according to the transition kernel K_{n,η_{n-1}^N}. So the interaction between the N particles is highly symmetric, and only through the empirical distribution of the particles. For this reasons, the name *mean field* particle system is used, still as a reference to the particle physics models for which they were introduced.

An interacting particle implementation of the measure flow introduced in the previous subsection is done via the flow of measures $(\eta_n^N)_n$ viewed as an approximation of the flow $(\eta_n)_n$. The rationale behind this approximation is that since η_n^N is the empirical distribution of N independent random variables with distributions $K_{n,\eta_{n-1}^N}(\xi_{n-1}^i, x)$, we expect that when η_{n-1}^N is a good approximation of η_{n-1} then in view of (17), η_n^N should be a good approximation of η_n. We define the approximation error (which is stochastic because of the randomness of the particles ξ_n^i) in terms of a sequence of centered random fields V_n^N defined by:

$$V^N = \sqrt{N}(\eta_n - \eta_n^N) = \sqrt{N}(\eta_{n-1}^N K_{n,\eta_{n-1}^N} - \eta_{n-1}K_{n,\eta_{n-1}}). \tag{18}$$

Then, under rather weak regularity conditions, one can prove that $(V_n^N)_{n\ge 0}$ converges in law as $N \to \infty$, toward a sequence of independent centered Gaussian fields $\underline{V} = (V_n)_{n\ge 0}$ with a variance function that can be explicitly expressed in terms of the Markov transitions $K_{n,\eta_{n-1}}$. This can be checked by induction [30] on the time parameter, or using martingale decompositions in terms of local sampling random fields [31, 40].

Using formula (12) as rationale, the unnormalized measures γ_n are approximated by the unbiased particle (unnormalized) measures γ_n^N defined by their actions on test functions by:

$$\gamma_n^N(f_n) = \eta_n^N(f_n) \times \prod_{0\le p<n} \eta_p^N(G_p),$$

and the weak consistency results

$$\lim_{N\to\infty} \gamma_n^N(f_n) = \gamma_n(f_n)$$

for each fixed test function f_n is proven by an elementary argument.

The stochastic perturbation analysis discussed above is developed in some details in [31, 43, 40, 46], and in the recent book [34]. Under some appropriate regularity

conditions on the flow of measures η_n, for any bounded measurable function f, any time horizon n, any $N \geq 1$, and any λ, the probability to have any of the following estimates is greater that $1-2e^{-\lambda}$

$$\left|\eta_n^N(f)-\eta_n(f)\right| \leq (1+\sqrt{2\lambda})\, c/\sqrt{N}$$

and

$$\left|1-\gamma_n^N(1\!\!1)/\gamma_n(1\!\!1)\right| \leq n\left(c_1\left(1+2(\lambda+\sqrt{\lambda})\right)/N+\sqrt{c_2\lambda/N}\right) ,$$

for some constants $c, c_1, c_2 < \infty$, whose values do not depend upon time.

By construction, the flow of Feynman-Kac measures evolves according to the two-step updating/prediction transitions,

$$\eta_n \xrightarrow{\mathscr{S}_{n,\eta_n}} \widehat{\eta}_n = \eta_n \mathscr{S}_{n,\eta_n} = \Psi_{G_n}(\eta_n) \xrightarrow{M_{n+1}} \eta_{n+1} = \widehat{\eta}_n M_{n+1} . \quad (19)$$

In the corresponding N-mean field particle model, this pair of recursions is replaced by a two-step selection/mutation transition in product spaces

$$\xi_n \in E^N \xrightarrow{selection} \widehat{\xi}_n \in E^N \xrightarrow{mutation} \xi_{n+1} \in E^N . \quad (20)$$

The genetic type evolution of the system is summarized by the following synthetic diagram:

$$\begin{bmatrix} \xi_n^1 \\ \vdots \\ \xi_n^i \\ \vdots \\ \xi_n^N \end{bmatrix} \xrightarrow{\mathscr{S}_{n,\eta_n^N}} \begin{bmatrix} \widehat{\xi}_n^1 \xrightarrow{M_{n+1}} \xi_{n+1}^1 \\ \vdots \\ \widehat{\xi}_n^i \xrightarrow{M_{n+1}} \xi_{n+1}^i \\ \vdots \\ \widehat{\xi}_n^N \xrightarrow{M_{n+1}} \xi_{n+1}^N \end{bmatrix}$$

with the selection Markov transition:

$$\mathscr{S}_{n,\eta_n^N}(\xi_n^i,x) := G_n(\xi_n^i)\; 1\!\!1_{\xi_n^i}(x) + \left(1-G_n(\xi_n^i)\right) \sum_{1\leq j\leq N} \frac{G_n(\xi_n^j)}{\sum_{1\leq k\leq N} G_n(\xi_n^k)}\, 1\!\!1_{\xi_n^j}(x) . \quad (21)$$

For general non necessarily $[0,1]$-valued potential functions G, we replace the acceptance rate $G_n(\xi_n^i)$ by $G_n(\xi_n^i)/\max_j G_n(\xi_n^j)$.

3.3 Path space models

We now work in the path space set up introduced earlier in our discussion of the Asian option example. In other words, we assume that the reference Markov chains X and the potential function G in (10) are defined on path spaces:

$$X_n := (S_0, \cdots, S_n) \in E_n = E^{n+1} \quad \text{and} \quad G_n(X_n) := G_n(S_0, \ldots, S_n) \ . \tag{22}$$

3.3.1 Genealogical tree based algorithms

The abstract Feynman-Kac formulae discussed above are more general than it may appear. They can be used to analyze path spaces models, including historical processes or transition space models, as well as finite excursion models. These stochastic models also encapsulate quenched Feynman-Kac models with respect to some parameter, island type coarse grained particle algorithms, Brownian type bridges and linear Gaussian Markov chains conditioned on starting and end points. For n extensive discussion on these path space models, we refer the interested reader to Section 2.4, Section 2.6, and Chapters 11-12 in the monograph [31], as well as Section 2.6 of the lecture notes [43], and Section 3.4 and Section 7.3 of the present article.

In the situation of this subsection, γ_n is a measure on E^{n+1} defined by

$$\gamma_n(f_n) = \mathbb{E}\left(f_n(S_0, \ldots, S_n) \prod_{0 \le q < n} G_q(S_0, \ldots, S_q) \right) \ .$$

Its mean field particle approximation is defined as before, but now, a particle at time n is a path of length $n+1$. The selection transition consists in selecting a path-particle with high potential value, while the mutation transition simply consists in extending the path with an elementary move according to the auxiliary process $X'_n = S_n$, with Markov transitions M'_n on the state space E. When the potential functions only depend upon the terminal value of the paths

$$G_n(X_n) := G'_n(S_n) \ ,$$

for some G'_n which we sometimes call *fitness function*, we can check that the path particle model gives the evolution of the genealogical tree model associated with the time evolution of the individuals ξ^i_n evolving with M'_n-mutations and G'_n-selections. In this situation, if

$$\xi^i_n := (\xi^i_{0,n}, \xi^i_{1,n}, \ldots, \xi^i_{n,n})$$

stands for the i-th ancestral line of the current individual $\xi^i_{n,n}$ after the n-th mutation, then for any function f_n on E_n, we have that

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} f_n\left(\xi^i_{0,n}, \xi^i_{1,n}, \ldots, \xi^i_{n,n}\right) = \frac{\mathbb{E}\left(f_n(S_0, \ldots, S_n) \prod_{0 \le p < n} G'_p(S_p)\right)}{\mathbb{E}\left(\prod_{0 \le p < n} G'_p(S_p)\right)} \ . \tag{23}$$

In addition, we also have the unbiased unnormalized estimates in the sense that:

$$\begin{aligned}&\frac{1}{N}\sum_{i=1}^{N} f_n\left(\xi^i_{0,n},\xi^i_{1,n},\ldots,\xi^i_{n,n}\right)\times\prod_{0\le p<n}\frac{1}{N}\sum_{i=1}^{N}G'_p(\xi^i_{p,p})\\ &\simeq_{N\uparrow\infty}\mathbb{E}\left(f_n(S_0,\ldots,S_n)\ \textstyle\prod_{0\le p<n}G'_p(S_p)\right)\ .\end{aligned}\tag{24}$$

If we look at what this particle algorithm gives in the case of the one dimensional fixed strike Asian option (5) with stochastic interest rates $r_p(X'_p)>0$ and time homogeneous barrier set $A_p=A$, we have the unbiased estimates:

$$\mathbb{E}\left(\left(\frac{1}{n+1}\sum_{p=0}^{n}S_p-K\right)_+ 1\!\!1_{T\ge n}\ \exp\left\{\sum_{0\le q<n}r_q(S_q)\right\}\right)$$

$$\simeq_{N\uparrow\infty}\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{n+1}\sum_{p=0}^{n}\xi^i_{p,n}-K\right)_+\times\prod_{0\le p<n}\frac{1}{N}\sum_{i=1}^{N}e^{-r_p(\xi^i_{p,p})}1\!\!1_A(\xi^i_{p,p})\ ,$$

where T stands for the first exit time of the process S outside the barrier A. The approximation of European barrier call option prices with strike $K>0$, stochastic interest rates $r_p(S_p)$, and time homogeneous barrier set A is even simpler. It is given by the unbiased estimates:

$$\mathbb{E}\left((S_n-K)_+\ 1\!\!1_{T\ge n}\ \exp\left\{\textstyle\sum_{0\le q<n}r_q(S_q)\right\}\right)$$

$$\simeq_{N\uparrow\infty}\frac{1}{N}\sum_{i=1}^{N}\left(\xi^i_{n,n}-K\right)_+\times\prod_{0\le p<n}\frac{1}{N}\sum_{i=1}^{N}e^{-r_p(\xi^i_{p,p})}1\!\!1_A(\xi^i_{p,p})\ .$$

If we use as before the notations η^N_n and η_n for the occupation measures of the ancestral lines and its limiting measures defined in (24), using the concentration analysis of mean field particle models developed in [46], the following exponential estimate was proved in [43]. Under some natural regularity conditions on the flow of the n-th time marginal measures, for any bounded measurable function f_n on path space, time horizon n, $N\ge 1$, and λ, the following estimate:

$$\left|\eta^N_n(f_n)-\eta_n(f_n)\right|\le\left(c_1\ (n+1)\left(1+2(\lambda+\sqrt{\lambda})\right)/N+c_2\ \sqrt{\lambda(n+1)/N}\right)\ ,$$

holds with probability greater that $1-2e^{-\lambda}$. Here c_1 and c_2 are finite constants whose values do not depend on the time parameter.

3.3.2 Backward Markov chain model

To distinguish path space measures and their finite time marginals, we denote by Γ_n and $\mathbb{Q}_n$ the measures on path space defined for any function F_n on E_n by

$$\Gamma_n(F_n) = \mathbb{E}\left(F_n(S_0,\ldots,S_n)\prod_{0\le q<n} G_q(S_q)\right) \quad \text{and} \quad \mathbb{Q}_n(F_n) = \Gamma_n(F_n)/\Gamma_n(1\!\!1)\ , \tag{25}$$

for some Markov chain $S=(S_n)_n$ on the state space E with initial distribution η_0, and some space time potential $G=(G_n)_n$. We also denote by γ_n and η_n the n-th marginal measure defined for any function f_n on E by

$$\gamma_n(f_n) = \mathbb{E}\left(f_n(S_n)\prod_{0\le q<n} G_q(S_q)\right) \quad \text{and} \quad \eta_n(F_n) = \gamma_n(F_n)/\gamma_n(1\!\!1)\ .$$

We observe that

$$\Gamma_n(d(s_0,\ldots,s_n)) := \left\{\prod_{0\le q<n} G_q(s_q)\right\}\ \mathbb{P}_n(ds_0\times\cdots\times ds_n)\ , \tag{26}$$

with the probability measure $\mathbb{P}_n$ on the path space E_n defined by

$$\mathbb{P}_n(ds_0\times\cdots\times ds_n) = \eta_0(ds_0)M_1(s_0,ds_1)\ldots M_n(s_{n-1},ds_n)\ .$$

We further assume that the Markov transitions $M_n(s,ds')$ of the reference Markov chain S has a density $H_n(s,s')$ with respect to some measure $\lambda_n(ds')$:

$$M_n(s,ds') = H_n(s,s')\ \lambda_n(ds')\ .$$

In this case, one easily derives the following backward representation:

$$\begin{array}{l}\mathbb{Q}_n(d(s_0,\ldots,s_n))\\ := \eta_n(ds_n)\times\ \mathbb{M}_{n,\eta_{n-1}}(s_n,ds_{n-1})\ \cdots\ \mathbb{M}_{2,\eta_1}(s_2,ds_1)\times\mathbb{M}_{1,\eta_0}(s_1,ds_0)\ ,\end{array} \tag{27}$$

with the time reversal Markov transitions $\mathbb{M}_{n,\eta_{n-1}}(s_n,ds_{n-1})$ defined by

$$\mathbb{M}_{n,\eta_{n-1}}(s_n,ds_{n-1}) := \frac{\eta_{n-1}(ds_{n-1})G_{n-1}(s_{n-1})H_n(s_{n-1},s_n)}{\eta_{n-1}\left(G_{n-1}H_n(\bullet,s_n)\right)}\ .$$

We refer the interested reader to the article [36] for a detailed discussion on these Markov transitions. Mimicking formula (27) an alternative particle approximation of the measures $\mathbb{Q}_n$ by the following estimates

$$\begin{array}{l}\mathbb{Q}_n^N(d(s_0,\ldots,s_n))\\ := \eta_n^N(ds_n)\times\mathbb{M}_{n,\eta_{n-1}^N}(s_n,ds_{n-1})\cdots\mathbb{M}_{2,\eta_1^N}(s_2,ds_1)\times\mathbb{M}_{1,\eta_0^N}(s_1,ds_0)\\ \to_{N\uparrow\infty}\mathbb{Q}_n(d(s_0,\ldots,s_n))\end{array} \tag{28}$$

and the unbiased unnormalized estimates

$$
\begin{aligned}
&\Gamma_n^N(d(s_0,\ldots,s_n)) \\
&:= \gamma_n^N(1\!\!1) \times \mathbb{Q}_n^N(d(s_0,\ldots,s_n)) \\
&= \gamma_n^N(ds_n) \times \mathbb{M}_{n,\eta_{n-1}^N}(s_n,ds_{n-1}) \cdots \mathbb{M}_{2,\eta_1^N}(s_2,ds_1) \times \mathbb{M}_{1,\eta_0^N}(s_1,ds_0) \\
&\rightarrow_{N\uparrow\infty} \mathbb{Q}_n(d(s_0,\ldots,s_n)) \; .
\end{aligned}
\tag{29}
$$

Notice also that the computation of sums with respect to these particle measures are reduced to summations over the particles locations ξ_n^i. It is therefore natural to identify a population of individual $(\xi_n^1,\ldots,\xi_n^N)$ at time n to a specific ordering of the set $\{1,\ldots,N\}$ of indexes. In this case, the occupation measures and the functions are identified with the following row and column vectors

$$
\eta_n^N := \left[\frac{1}{N},\ldots,\frac{1}{N}\right] \quad \text{and} \quad \mathrm{f}_n := \begin{bmatrix} f_n(\xi_n^{1\!\!1}) \\ \vdots \\ f_n(\xi_n^N) \end{bmatrix}
$$

and the matrices $\mathbb{M}_{n,\eta_{n-1}^N}$ by the $N \times N$ matrices

$$
\mathbb{M}_{n,\eta_{n-1}^N} := \begin{bmatrix} \mathbb{M}_{n,\eta_{n-1}^N}(\xi_n^{1\!\!1},\xi_{n-1}^{1\!\!1}) & \cdots & \mathbb{M}_{n,\eta_{n-1}^N}(\xi_n^{1\!\!1},\xi_{n-1}^N) \\ \vdots & \vdots & \vdots \\ \mathbb{M}_{n,\eta_{n-1}^N}(\xi_n^N,\xi_{n-1}^{1\!\!1}) & \cdots & \mathbb{M}_{n,\eta_{n-1}^N}(\xi_n^N,\xi_{n-1}^N) \end{bmatrix} ,
$$

with the (i,j)-entry $\mathbb{M}_{n,\eta_{n-1}^N}(\xi_n^i,\xi_{n-1}^j)$ defined by:

$$
\mathbb{M}_{n,\eta_{n-1}^N}(\xi_n^i,\xi_{n-1}^j) = \frac{G_{n-1}(\xi_{n-1}^j) H_n(\xi_{n-1}^j,\xi_n^i)}{\sum_{k=1}^N G_{n-1}(\xi_{n-1}^k) H_n(\xi_{n-1}^k,\xi_n^i)} \; .
$$

For instance, the $\mathbb{Q}_n$-integration of normalized additive linear functionals of the form

$$
\overline{F}_n(s_0,\ldots,s_n) = \frac{1}{n+1} \sum_{0\le p\le n} f_p(s_p) \tag{30}
$$

is given in the particle matrix approximation model by:

$$
\mathbb{Q}_n^N(\overline{F}_n) = \frac{1}{n+1} \sum_{0\le p\le n} \eta_n^N \mathbb{M}_{n,\eta_{n-1}^N} \mathbb{M}_{n-1,\eta_{n-2}^N} \ldots \mathbb{M}_{p+1,\eta_p^N}(f_p) \; . \tag{31}
$$

Several non asymptotic convergence estimates have been developed in [36], distinguishing the bias error

$$
\sup_{n\ge 0} \left| \mathbb{E}\Big(\mathbb{Q}_n^N(\overline{F}_n)\Big) - \mathbb{Q}_n^N(\overline{F}_n) \right| \le \frac{c}{N} \; ,
$$

and the mean quadratic error

$$\mathbb{E}\left|\mathbb{Q}_n^N(\overline{F}_n)-\mathbb{Q}_n(\overline{F}_n)\right|^2 \leq \frac{c}{N}\left(\frac{1}{n+1}+\frac{1}{N}\right) , \quad \text{for all } n \geq 0 ,$$

where c is some finite constant that does not depend on the time parameter n. Thus, for any large time horizon $n \geq N$, the upper bound on the mean square error given in the above right hand side is of the order $1/N^2$. More recently, a different estimate was proven in [43] using the concentration methodology developed in [46]. Under some appropriate regularity conditions on the flow of the n-th time marginal measures, for any sequence of bounded measurable functions f_n, any time horizon n, any $N \geq 1$, and any λ, the estimate

$$\left|\mathbb{Q}_n^N(\overline{F}_n)-\mathbb{Q}_n(\overline{F}_n)\right| \leq \left(c_1\left(1+2(\lambda+\sqrt{\lambda})\right)/N + c_2\sqrt{\lambda/(N(n+1))}\right) ,$$

for some constants $c_1, c_2 < \infty$ whose values do not depend upon n, holds with probability greater that $1-2e^{-\lambda}$.

Additive functionals of the form (30) arise in many applications in finance. For instance, in the context of continuous Asian option, this approach could allow to improve seriously the trade-off between the bias induced by the discrete approximation of the continuous integral payoff and the variance of the Monte Carlo method approximating the expectation. We refer to [58] for a survey of numerical methods for this type of options. As an example we consider the case of continuous Asian options with payoff functions of the following general form

$$F_T\left((S_t^c)_{0\leq t\leq T}\right) = \left(\frac{1}{T}\int_0^T f(u,S_u^c)\,du - K\right)^+ .$$

The strategy coming out of the above discussion suggests to first estimate the payoff by the following arithmetic average

$$\overline{F}_n(S_{t_0}^c,\cdots,S_{t_n}^c) = \frac{1}{n+1}\sum_{0\leq p\leq n} f(t_p,S_{t_p}) ,$$

with a time discretization $t_{p+1}-t_p = T/(n+1)$, inducing an approximation error of order $1/n$. Then the backward Markov chain scheme (31) can be used to estimate the expectation with $N \geq n$ particles, inducing the same order of approximation error $1/n$. Of course, usual variance reduction techniques as control variate can be applied in addition to that approach.

This type of additive functionals will be used in Section 5 in our discussion of sensitivity measure computations.

3.4 Parallel island particle models

Island genetic models are powerful parallel computational techniques used to speed up interacting genetic search algorithms. These coarse grained parallel procedures are very popular in genetic algorithms literature (see for instance [26, 96],[74], and the references therein).

In our context, we run in parallel several genetic type interacting particle algorithms on a collection of islands. At a geometric stochastic rate, the populations between islands interact according to some selective migration processes. The island selection mechanism is defined in terms of the averaged fitness of the individuals in the island population.

To define these island particle models more precisely, we observe that the unbiased properties of the unnormalized Feynman-Kac measures γ_n^N can be rewritten as follows

$$\mathbb{E}\left(f_n(S_n)\prod_{0\leq q<n}G_q(S_q)\right)=\mathbb{E}\left(\mathscr{F}_n(\mathscr{X}_n)\prod_{0\leq p<n}\mathscr{G}_p(\mathscr{X}_p)\right), \tag{32}$$

with the Markov chain $\mathscr{X}_n=(\xi_n^i)_{1\leq i\leq N}$ on the product spaces $\mathscr{E}_n=E_n^N$, an the empirical functionals $\mathscr{F}_n$, and $\mathscr{G}_n$ defined by

$$\mathscr{F}_n(\mathscr{X}_n)=\eta_n^N(f_n)=\frac{1}{N}\sum_{i=1}^N f_n(\xi_n^i)\quad\text{and}\quad\mathscr{G}_n(\mathscr{X}_n)=\eta_n^N(G_n)=\frac{1}{N}\sum_{i=1}^N G_n(\xi_n^i)\,.$$

Now, it is important to notice that the r.h.s. term in formula (32) has exactly the same mathematical form as the Feynman-Kac measures γ_n introduced in (10). Thus, applying the particle methodologies developed in Section 3.2 to these models, we define an N-interacting island particle model with a mutation and a selection transition on the space of islands $\mathscr{E}_n$.

During the mutation stage, the population in each island evolve independently one another according to the genetic type Markov transitions of the chain $\mathscr{X}_n$. In other words, we run in parallel the selection mutation transitions of N genetic particle models (20). During the selection stage, we evaluate the $\mathscr{G}_n$-potential value of each island. As in (21), at a geometric rate we select the island populations using the empirical potential function $\mathscr{G}_n$.

We observe that the island version of the acceptance ratio in the selection transition (21) discussed in the end of Section 3.2 tends to 1, as the number of individuals in each island tends to infinity. In other words, the independence degree between the islands increases with respect to the size of their populations. As proposed in the recent article [1], an alternative sampling approach is to use an independent Metropolis-Hasting model with a target measure defined by the r.h.s. term in formula (32) (up to some normalizing constant). One again, the unbiased property (32) ensures that the limiting target measure coincides with the desired Feynman-Kac measures η_n, as well as the measure $\mathbb{Q}_n$ for the path space version of these island

models. These island particle models will be used in Section 7.3 dedicated to fixed parameter estimation in Hidden Markov chain models.

4 Application in credit risk analysis

The simulation of credit events with remarkably small probabilities is a key issue for regulatory and risk management purposes, as well as for the pricing of credit derivatives. The main variance reduction technique used in Monte Carlo computations of rare events is importance sampling. However in general multi-name credit models, desirable changes of measure favoring sample paths realizing rare events are highly unlikely to lead to explicit formula. In this case importance sampling is no longer an option. A natural alternative is then interacting particles methods.

Though interacting particle systems are known to provide very efficient variance reductions in Monte Carlo approximations of rare events, these algorithms have only appeared recently in the credit risk literature with for instance the articles of Carmona, Fouque and Vestal [16] and Carmona and Crepey [15]. In Chapter 21 of [12], the authors provide an overview of the main techniques and results of the application of interacting particle systems to credit risk analysis. We also refer to [44] for some recent applications of these techniques in the financial risk area. All these results show the strengths of IPS based Monte Carlo computations of small default probabilities, especially when other methods fail. A systematic comparison with importance sampling is provided in [15].

4.1 Change of measure for rare events and Feynman-Kac formula

We consider a Markov chain $S = (S_n)_{0 \leq n \leq T}$ representing at each time n, d correlated risky sources $S_n = (S_n^1, \ldots, S_n^d) \in E$. We are interested in understanding the asymptotic behavior of probabilities of rare events of the form $\{V_T(S_T) \geq K\}$ or more generally $\{V_T(S_0, \ldots, S_T) \geq K\}$, where V_T is some real positive function whose value can be thought of as a risk measure.

To compute $\mathbb{P}(V_T(S_T) \geq K)$, standard Monte Carlo simulations usually fail, because of the difficulty to ensure that enough simulation samples realize the rare event. A partial remedy amounts to providing a reasonably tight upper-bound based on large deviations ideas. Indeed, for any $\lambda \geq 0$, we have:

$$\mathbb{P}(V_T(S_T) \geq K) = (\mathbb{E}\left(\mathbf{1}_{V_T(S_T) \geq K} e^{\lambda V_T(S_T)} e^{-\lambda V_T(S_T)}\right) \leq e^{-\lambda K} \mathbb{E}\left(\mathbf{1}_{V_T(S_T) \geq K} e^{\lambda V_T(S_T)}\right),$$

and if we denote by $\mathbb{E}^{(\lambda)}$ the expectation under the probability $\mathbb{P}^{(\lambda)}$ defined by

$$d\mathbb{P}^{(\lambda)} \propto e^{\lambda V_T(S_T)} d\mathbb{P}\,,$$

we have

$$\mathbb{E}\left(\mathbf{1}_{V_T(S_T)\geq K}e^{\lambda V_T(S_T)}\right)=\mathbb{E}^{(\lambda)}\left(\mathbf{1}_{V_T(S_T)\geq K}\right)\mathbb{E}\left(e^{\lambda V_T(S_T)}\right),$$

and using the fact that:

$$\mathbb{E}^{(\lambda)}\left(\mathbf{1}_{V_T(S_T)\geq K}\right)\leq 1,$$

we get:

$$\mathbb{P}(V_T(S_T)\geq K)\leq e^{-\sup_{\lambda\geq 0}(\lambda K-\Lambda(\lambda))}\,, \tag{33}$$

where $\Lambda(\lambda)$ is defined by Fenchel transformation as $\log(\mathbb{E}(\lambda V_T(S_T)))$.

From the above argument we see that we can approximate the desired probability by searching a proper λ. This large deviation type approach is widely used, but in the form of (33), it requires extensive calculations in order to obtain a reasonable approximation of the desired probability.

Del Moral and Garnier provide in [37] a zero-bias estimate with interacting particle systems. The idea is to construct a genealogical tree based model as mentioned in Section 3.3.1 instead of the large deviation type inequality used above.

Using again the same change of measure from $\mathbb{P}$ to $\mathbb{P}^{(\lambda)}$, we remark that the target probability

$$\mathbb{P}(V_T(S_T)\geq K)=\mathbb{E}\left(\mathbf{1}_{V_T(S_T)\geq K}e^{\lambda V_T(S_T)}e^{-\lambda V_T(S_T)}\right)$$

can be written as

$$\mathbb{E}^{(\lambda)}\left(\mathbf{1}_{V_T(S_T)\geq K}e^{-\lambda V_T(S_T)}\right)\mathbb{E}\left(e^{\lambda V_T(S_T)}\right)=\mathbb{E}^{(\lambda)}\left(f_T(S_T)\right)\mathbb{E}\left(e^{\lambda V_T(S_T)}\right)\,,$$

with $f_T(S_T):=\mathbf{1}_{V_T(S_T)\geq K}e^{-\lambda V_T(S_T)}$. It is also important to notice that, with the convention $V_0=0$, we have the following decomposition

$$e^{\lambda V_T(S_T)}\equiv\prod_{p=1}^{T}e^{\lambda(V_p(S_p)-V_{p-1}(S_{p-1}))}\,.$$

By using the notation $\mathscr{X}_k=(S_k,S_{k+1})$ for $0\leq k<T$, the above product can be defined as

$$\prod_{p=1}^{T}G_{p-1}(\mathscr{X}_{p-1})\,,\quad\text{where}\quad G_{p-1}(\mathscr{X}_{p-1}):=e^{\lambda(V_p(S_p)-V_{p-1}(S_{p-1}))}\,.$$

Using the notation $F_T(\mathscr{X}_T)=f_T(S_T)$, we see that we need to simulate the same formulae as in Section 3.3.1.

$$\mathbb{E}^{(\lambda)}\left(f_T(S_T)\right)=\frac{\mathbb{E}\left(F_T(\mathscr{X}_T)\prod_{p=1}^{T}G_p(\mathscr{X}_P)\right)}{\mathbb{E}\left(\prod_{p=1}^{T}G_p(\mathscr{X}_P)\right)}:=\eta_T(F_T)\,.$$

The general discussion of the previous section shows that these quantities can be approximated efficiently by interacting particle systems.

In practice, we are interested in the conditional law $\mathscr{L}\Big((S_0,\ldots,S_T)|V_T(S_T)\geq K\Big)$. By modifying the function F_T in the above analysis, the same framework can be applied directly. Particle models are very flexible, but the choice of the space-time potential function on path space can become very tricky and the performance of the algorithm can deteriorate with a poor choice of this potential. The particular choice

$$G_p(s_0,\ldots,s_p)=\mu e^{\lambda\left(V_p(s_0,\ldots,s_p)-V_{p-1}(s_0,\ldots,s_{p-1})\right)} \tag{34}$$

was proposed and analyzed in [37] where μ is chosen so that $G_p\leq 1$ and λ can be fine-tuned to the given rare event set.

4.2 On the choice of the potential functions

As mentioned earlier, the choice of suitable space-time potential functions G is a key ingredient in the ability of interacting particle systems to tackle rare events problems. In the recent work of Carmona, Fouque and Vestal [16], the authors propose a choice of the potential functions that departs from the one given above in (34). Their construction illustrates the flexibility of the particle methods regarding the crucial point of choice of the potential functions. In the case of large credit portfolios, typically with $d=125$, we write the dynamics of the various assets values as a Markov chain S_n^i, with time $n=1,\cdots,T$ and $i=1,\ldots,d$ associated to

$$G_n=\exp\left(-\alpha\sum_{i=1}^{d}\log\frac{\min\limits_{0\leq l\leq n}S_l^i}{\min\limits_{0\leq l\leq n-1}S_l^i}\right)\ ,$$

where the parameter α has to be fine-tuned to the particular class of rare events of interest. Numerical performance of this technique is dicussed in [16] where examples are provided under a structural model with stochastic volatility. The authors demonstrate the efficiency of this method, especially in situations where importance sampling is not possible or numerically unstable.

In a similar vein, a fast algorithm without requirement of fine-tuned parameters has been recently developed for multiple defaults models by setting the potential function

$$G_p(x)=1-\mathbf{1}_{\{c\}}(x)\ , \tag{35}$$

where the c stands for a cemetery state under a multilevel splitting approach introduced in Chapter 12 in [31]. Let $S=(S_n)_{0\leq n\leq T}$ be a Markov chain on a sequence of state spaces $E=(E_n)_{0\leq n\leq T}$ and X_n (resp. $F_n=E_0\times\cdots\times E_n$) the corresponding path space Markov chain (resp. sequence of path state spaces). We assume that a sequence of subsets $U_1,\ldots,U_T,\ U_p\in F_p$ is fixed. We are typically interested in the probability $1-\mathbb{P}(X_1\notin U_1,\ldots,X_T\notin U_T)$ that the trajectory does enter at least one of these subsets. The key idea is to introduce a series of intermediate events interpo-

lating between the series of the full state space $E_1, \ldots, E_T$ of the path space and the target rare event series $U_1, \ldots, U_T$. Then we assume that such a series is given:

$$\forall\, p \leq T,\ U_p = U_p^{(k)} \subset U_p^{(k-1)} \subset \cdots \subset U_p^{(1)} \subset U_p^{(0)} = F_p\ .$$

Then the state $\{c\}$ appearing in (35) is defined in the construction of a new Markov chain in constant state space $F := F_0 \cup \cdots \cup F_T \cup \{c\}$. With a series of stopping times:

$$\tau_j := (T+1) \wedge \inf\{p, X_p \in U_p^j\}$$

with the convention that $X_{T+1} := c$. Then the process $Z_0 := X_0,\ Z_1 := X_{\tau_1}, \ldots, Z_k := X_{\tau_k}$ is a Markov chain on F. In this context, the potential functions (35) consist in, roughly speaking, killing the trajectories at some point of the recursion of the particle algorithm when they reach some of the intermediate rare event sets associated to c.

5 Sensitivity computation

Partial derivatives of financial option values allow traders to determine how sensitive the values of options are to small changes in the set of parameters on which they depend, such as the volatility parameter, the risk free stochastic interest rates or prices of assets related to the option. The computation of these *sensitivities*, often called *Greeks* (because they are traditionally denoted by Greek letters) is a central problem in computational finance that must be addressed for risk analysis applications. Besides, in the specific case of sensitivities with respect to assets prices, (called *delta* and *gamma* for the first and second order derivatives) the practical issue is even more crucial since they are the basic ingredients of dynamic hedging strategies.

There are mainly three approaches to compute sensitivities. We refer to the survey paper of Kohatsu-Higa and Montero [68], for a detailed presentation and comparison of those methods. The most natural and simple approach to compute sensitivities is the usual finite difference method. It is easily implemented but known to necessitate large computing budgets (requiring for instance two option calculations in the case of a first order sensitivity) and unstable with a subtle trade-off between bias and variance. We focus here on the two alternative approaches introduced in the pioneering paper of Broadie and Glasserman [13], namely the *likelihood ratio* method and the *pathwise*, or *tangent process* method. In this section, these techniques are presented in terms Feynman-Kac formula, showing in some specific examples how particle methods can be used.

5.1 Likelihood ratio: application to dynamic parameter derivatives

This technique introduced in [13] requires that the underlying interest on which a European option is written admits a sufficiently regular density with respect to Lebesgue measure, also known as state price density. The main idea is to interchange differentiation and integration and whenever the derivative with respect to a variable not appearing in the payoff function, to apply the differentiation on the density of the distribution. The advantage of this approach is that it does not require any regularity assumption on the payoff function, allowing for kinks and discontinuities. This approach has been generalized by Fournié, Lasry, Lebuchoux, Lions and Touzi [56] to path space using Malliavin integration-by-parts argument, allowing for a wide class *Greek weights*.

In this subsection, we focus on the computation of the sensitivity of an option to dynamic parameters related to the risky asset evolution or to the risk free rate variations.

We let $\theta \in \mathbb{R}^d$ be a parameter that may represent the volatility of some asset price movements, or any other kinetic parameter. We assume that the evolution of the risky asset price $S_k^{(\theta)}$ associated to some value of the parameter θ, is given by a one-step probability transition of the form

$$M_k^{(\theta)}(s,ds') := \text{Proba}\left(S_k^{(\theta)} \in ds' | S_{k-1}^{(\theta)} = s\right) = H_k^{(\theta)}(s,s')\ \lambda_k(ds')\ ,$$

for some positive density functions $H_k^{(\theta)}(s,s')$ and some reference measure λ_k. We also consider a collection of functions $G_k^{(\theta)}(s) = e^{-r_k^{(\theta)}(s)}$ that depend on θ. We also assume that the gradient and the Hessian of the logarithms of these functions with respect to the parameter θ are well defined.

In this situation, following the Feynman-Kac representation (1) or (25), a general form of the option price on path space is provided by

$$\Gamma_n^{\theta}(F_n) = \mathbb{E}\left(F_n(S_0^{(\theta)},\ldots,S_n^{(\theta)}) \prod_{0\leq p<n} G_p^{(\theta)}\left(S_p^{(\theta)}\right)\right)\ . \tag{36}$$

For each value of the parameter θ, we denote by $\Gamma_n^{(\theta,N)}$ the N-particle approximation measures associated with a given value of the parameter θ and defined in (29). Simple derivations, show that the first order derivative of the option value with respect to θ is given by

$$\begin{aligned}\nabla\Gamma_n^{(\theta)}(F_n) &= \Gamma_n^{(\theta)}(F_n\Lambda_n^{(\theta)})\\ \nabla^2\Gamma_n^{(\theta)}(F_n) &= \Gamma_n^{(\theta)}\left[F_n(\nabla\mathbb{L}_n^{(\theta)})'(\nabla\mathbb{L}_n^{(\theta)}) + F_n\nabla^2\mathbb{L}_n^{(\theta)}\right]\end{aligned}$$

with $\Lambda_n^{(\theta)} := \nabla\mathbb{L}_n^{(\theta)}$ and

$$\mathbb{L}_n^{(\theta)}(x_0,\ldots,x_n) := \sum_{p=1}^{n} \log\left(G_{p-1}^{(\theta)}(s_{p-1})H_p^{(\theta)}(s_{p-1},s_p)\right) .$$

These quantities can be approximated by the unbiased particle models

$$\begin{aligned}
&\nabla_N \Gamma_n^{(\theta)}(F_n) := \Gamma_n^{(\theta,N)}(F_n\Lambda_n^{(\theta)}) \\
&\nabla_N^2 \Gamma_n^{(\theta)}(F_n) \;=\; \Gamma_n^{(\theta,N)}\left[F_n(\nabla\mathbb{L}_n^{(\theta)})'(\nabla\mathbb{L}_n^{(\theta)}) + F_n\nabla^2\mathbb{L}_n^{(\theta)}\right] .
\end{aligned}$$

We illustrate the above discussion with the computation of the Vega of the option, i.e. the sensitivity to changes in the diffusion volatility coefficient of the stochastic equation (7), with $d = 1$. We suppose $X_n^{(\theta)} = S_n^{(\theta)}$ satisfies equation

$$S_n^{(\theta)} - S_{n-1}^{(\theta)} = b\left(S_{n-1}^{(\theta)}\right)\,\Delta + \left[\sigma\left(S_{n-1}^{(\theta)}\right) + \theta\;\sigma'\left(S_{n-1}^{(\theta)}\right)\right]\;\left(W_{t_n} - W_{t_{n-1}}\right) ,$$

for some function σ' such that $\sigma + \theta\;\sigma' > 0$ for any $\theta \in [0,1]$. In this situation, we have

$$\begin{aligned}
&\frac{\partial}{\partial\theta}\sum_{p=1}^{n}\log\left(H_p^{(\theta)}(s_{p-1},s_p)\right) \\
&= \sum_{p=1}^{n}\frac{\sigma'(s_{p-1})}{\sigma(s_{p-1})+\theta\sigma'(s_{p-1})}\;\left[\left(\frac{(s_p-s_{p-1})-b(s_{p-1})\Delta}{(\sigma(s_{p-1})+\theta\sigma'(s_{p-1}))\sqrt{\Delta}}\right)^2 - 1\right] .
\end{aligned}$$

To compute the rho of the option, i.e. the sensitivity to changes in the drift of the stochastic equation (7), with $d = 1$, we assume that $X_n^{(\theta)}$ satisfies equation

$$S_n^{(\theta)} - S_{n-1}^{(\theta)} = \left[b\left(S_{n-1}^{(\theta)}\right) + \theta b'\left(S_{n-1}^{(\theta)}\right)\right]\;\Delta + \sigma\left(S_{n-1}^{(\theta)}\right)\;\left(W_{t_n} - W_{t_{n-1}}\right) ,$$

for some function b'. In this situation, we have

$$\begin{aligned}
&\frac{\partial}{\partial\theta}\sum_{p=1}^{n}\log\left(H_p^{(\theta)}(s_{p-1},s_p)\right) \\
&= \sum_{p=1}^{n}\left[(s_p-s_{p-1}) - \left[b(s_{p-1})+\theta b'(s_{p-1})\right]\Delta\right]\times b(s_{p-1})/\sigma^2(s_{p-1}) .
\end{aligned}$$

Finally, if we assume that changes in the stochastic interest rates are given by the space-time potential function

$$G_n(x) = \exp\left(-\left[r_n(x) + \theta r_n'(x)\right]\right) ,$$

for some non negative functions r_n and r_n', then we have:

$$\frac{\partial}{\partial\theta}\sum_{0\leq p<n}\log\left(G_p^{(\theta)}(s_p)\right)=-\sum_{0\leq p<n}r'_p(s_p)\ .$$

We illustrate these particle models with an European option associated with a risky asset $S_n^{(\theta)}=S_n$ whose values do not depend on θ, and payoff function $f_n(S_n)$. In this situation, the option price is given by the formula

$$\gamma_n^{(\theta)}(f_n)=\mathbb{E}\left(f_n(S_n)\ \exp\left\{-\sum_{0\leq q<n}r_q^{(\theta)}(S_q)\right\}\right)\ .$$

Then using the backward Markov chain model developed in Section 3.3.2, we obtain the following unbiased particle matrix approximation for the sensitivity with respect to the interest rate r:

$$\begin{aligned}\nabla\gamma_n^{(\theta)}(f_n) &= -\sum_{0\leq p<n}\gamma_n^{(\theta)}\left(f_n\,\mathbb{M}_{n,\eta_{n-1}^{(\theta)}}\ldots\mathbb{M}_{p+1,\eta_p^{(\theta)}}\left(\nabla r_p^{(\theta)}\right)\right)\\ &\simeq_{N\uparrow\infty}-\sum_{0\leq p<n}\gamma_n^{(\theta,N)}\left(f_n\,\mathbb{M}_{n,\eta_{n-1}^{(\theta,N)}}\ldots\mathbb{M}_{p+1,\eta_p^{(\theta,N)}}\left(\nabla r_p^{(\theta)}\right)\right)\ .\end{aligned}$$

5.2 Tangent process: application to initial state derivatives

We review the tangent process approach introduced by Broadie and Glasserman in [13], and focus on the computation of the sensitivity of an option price to perturbations of the initial value of the underlying asset price – this sensitivity is usually called the *delta* of the option – which is in general more complex than in the case of the sensitivity measures with respect to the dynamic parameters parameters. Efficient numerical schemes for the implementation of the method we are about to discuss can be found in Giles and Glasserman [64].

To simplify our presentation, we only consider European-style options with smooth payoff functions $f-(f_n)_n$.

As before, the strategy is to interchange the differentiation and expectation operations. However, in the present situation. this requires regularity of the payoff function, so discontinuous payoff profiles will have to be regularized using Gaussian kernel convolution type techniques, or any related smoothing method. For instance, we can approximate the call option (3) by the following smoothed payoff profile

$$f_\varepsilon(x)=\frac{1}{2}\left[(x-K)+\sqrt{(K-x)^2+\varepsilon}\right]\to_{\varepsilon\downarrow 0}f(x)=(x-K)_+\ .$$

We assume that the stochastic dynamics of the underlying stock price $S=(S_n)_n$ are given by an induction equation of the type:

$$S_{n+1}:=F_n(S_n)=(F_n\circ F_{n-1}\circ\cdots\circ F_0)(S_0) \tag{37}$$

starting at some random state S_0, where the random functions F_n are of the form of the form

$$F_n(x) = \mathscr{F}_n(x, W_n) \ , \tag{38}$$

for functions $\mathscr{F}_n$

$$\mathscr{F}_n \ : \mathbb{R}^{d+d'} \ni (x,w) \hookrightarrow \mathscr{F}_n(x,w) \in \mathbb{R}^d$$

and some independent random variables W_n taking values in $\mathbb{R}^{d'}$, with $d' \geq 1$. We also assume that these random variables are also independent of S_0. Under these assumptions, the prices of European options are given by the semigroup of the Markov chain S defined for any regular function f and initial state x by

$$P_{n+1}(f)(x) := \mathbb{E}\left(f(S_{n+1}) \mid S_0 = x\right) = \mathbb{E}\left(f(S_{n+1}(x))\right) \ ,$$

with the random flows $(S_n(\cdot))_{n\geq 0}$ defined for any $n \geq 0$ and $x \in E$ by:

$$S_{n+1}(x) = F_n(S_n(x)) \ ,$$

with the initial condition $S_0(x) = x$. By the chain rule, for any $1 \leq i, j \leq d$ and any $x \in \mathbb{R}^d$ we have

$$\frac{\partial S^i_{n+1}}{\partial x^j}(x) = \sum_{1\leq k\leq d} \frac{\partial F^i_n}{\partial x^k}(S_n(x)) \ \frac{\partial S^k_n}{\partial x^j}(x) \ . \tag{39}$$

Interchanging derivations and expectations in the definition of the semigroup we get:

$$\frac{\partial P_{n+1}(f)}{\partial x^j}(x) = \mathbb{E}\left(\sum_{1\leq i\leq d} \frac{\partial f}{\partial x^i}(S_{n+1}(x)) \ \frac{\partial S^i_{n+1}}{\partial x^j}(x)\right) \ . \tag{40}$$

Let us denote by $V_n = (V_n^{(i,j)})_{1\leq i,j\leq d}$ and $A_n = (A_n^{(i,j)})_{1\leq i,j\leq d}$ the random $d \times d$ matrices whose entries are given by:

$$V_n^{(i,j)}(x) = \frac{\partial S^i_n}{\partial x^j}(x)$$

and

$$A_n^{(i,j)}(x) = \frac{\partial F^i_n}{\partial x^j}(x) = \frac{\partial \mathscr{F}^i_n(\bullet, W_n)}{\partial x^j}(x) := \mathscr{A}_n^{(i,j)}(x, W_n) \ .$$

With this notation in hand, equation (39) can be rewritten in terms of the following random matrix formulae

$$\begin{aligned} V_{n+1}(x) &= A_n(S_n(x)) \ V_n(x) \\ &= A_n(S_n(x)) A_{n-1}(S_{n-1}(x)) \cdots A_1(S_1(x)) A_0(x) := \prod_{p=0}^{n} A_p(S_p(x)) \ , \end{aligned} \tag{41}$$

with a product $\prod_{p=0}^n A_p$ of non commutative random elements A_p taken in the order $A_n, A_{n-1}, \ldots, A_0$. Using equation (40) with the payoff function $f = f_{n+1}$, we get:

$$\begin{aligned}\nabla P_{n+1}(f_{n+1})(x) &= \mathbb{E}\left(\nabla f_{n+1}(S_{n+1}(x))\ V_{n+1}(x)\right) \\ &= \mathbb{E}\left(\nabla f_{n+1}(S_{n+1}) \prod_{0\leq p\leq n} A_p(S_p) \mid S_0 = x\right) , \end{aligned} \tag{42}$$

which is, except for the fact that we are dealing with products of non-commuting random matrices, of the form of the Feynman-Kac formulas studied in this paper.

For one dimensional models of the form

$$S_{n+1} = S_n + b(S_n)\ \Delta + \sigma(S_n)\ \sqrt{\Delta}\ W_n , \tag{43}$$

with a sequence of independent and and identically distributed mean zero Gaussian random variables W_n, it is readily checked that

$$A_n(x) = \mathscr{A}_n(x, W_n) = \left(1 + \frac{\partial b}{\partial x}(x)\ \Delta + \frac{\partial \sigma}{\partial x}(x)\ \sqrt{\Delta}\ W_n\right)$$

and therefore

$$\begin{aligned} V_{n+1}(x) &= \prod_{p=0}^{n}\left(1 + \frac{\partial b}{\partial x}(S_p)\ \Delta + \frac{\partial \sigma}{\partial x}(S_p)\ \sqrt{\Delta}\ W_p\right) \\ &\simeq_{\Delta\downarrow 0} \exp \sum_{0\leq p\leq n}\left(\frac{\partial b}{\partial x}(S_p)\ \Delta + \frac{\partial \sigma}{\partial x}(S_p)\ \sqrt{\Delta}\ W_p\right) . \end{aligned}$$

As already mentioned, for non smooth payoff functions we can use the following Gaussian regularization kernel

$$P_{n+1,\varepsilon}(f_{n+1})(x) := \mathbb{E}\left(f_{n+1}(S_{n+1}(x) + \varepsilon Y)\right) \simeq_{\varepsilon\downarrow 0} P_{n+1,\varepsilon}(f_{n+1})(x) , \tag{44}$$

for some auxiliary Gaussian variable, independent of S_n and W_n. In this case, we have the following formula

$$\frac{\partial}{\partial x} P_{n+1,\varepsilon}(f_{n+1})(x) = \mathbb{E}\left(\varepsilon^{-1}\left[f_{n+1}(S_{n+1}(x) + \varepsilon Y) - f_{n+1}(S_{n+1}(x))\right]\ Y\ V_{n+1}(x)\right) .$$

In the particular case $d = 1$, the particle interpretation developed in Section 3.2 applies directly. W

Remark 5. As an aside, we also mention that these expansions are closely related to the time discretization of the stochastic integrals arising in exponential weights of the Feynman-Kac interpretation of the Kushner-Stratonovitch filtering equation [29]. In this interpretation, the particle interpretations of the Feynman-Kac formulae (42) coincide with the particle filters developed in the last referenced article.

Before getting into multi-dimensional models, let us pause for a while to discuss the connexions of the above methodology with the existing literature. Firstly, we observe that the Gaussian regularization formula (44) can be interpreted as the addition of an extra Gaussian move. This suggests that we can alternatively use the last

transition to regularize the model.

$$P_{n+1}(f_{n+1})(x) = \mathbb{E}\left(\mathbb{E}\Big(f_{n+1}\big(S_{n+1}(x)\big)\Big|S_n(x)\Big)\right) .$$

Letting $H_{n+1}(x_n, x_{n+1})$ be the density of the Markov transition $S_n = x_n \rightsquigarrow S_{n+1}$ with respect to the Lebesgue measure, arguing as above we find that

$$\frac{\partial}{\partial x} P_{n+1}(f_{n+1})(x) = \mathbb{E}\left(f_{n+1}(S_{n+1}(x))\; dH_{n+1}(S_n(x), S_{n+1}(x))\; V_n(x)\right) ,$$

with the weight function

$$\begin{aligned} & dH_{n+1}(x_n, x_{n+1}) \\ &= \frac{\partial}{\partial x_n} \log H_{n+1}(x_n, x_{n+1}) \\ &= \left(\left(\tfrac{(x_{n+1}-x_n)-b(x_n)\Delta}{\sigma(x_n)\sqrt{\Delta}}\right)^2 - 1\right) \frac{\partial}{\partial x} \log \sigma(x_n) - \left(\tfrac{(x_{n+1}-x_n)-b(x_n)\Delta}{\sigma(x_n)\sqrt{\Delta}}\right) \frac{1 + \frac{\partial b}{\partial x}(x_n)\Delta}{\sigma(x_n)\sqrt{\Delta}} . \end{aligned}$$

These formulae and the corresponding conventional weighted Monte Carlo approximations have been recently proposed by N. Chen and P. Glasserman [22] as an alternative to the Malliavin calculus computation of the Greeks introduced by E. Fournié, J.M. Lasry, J. Lebuchoux, P.L. Lions, and N. Touzi in their groundbreaking articles [56, 57]. If $P_{s,t}$ denotes the semigroup associated with the (continuous time) diffusion equation (6) (recall that $d = 1$ in the present discussion):

$$P_{s,t}(f)(S_s) = \mathbb{E}\left(f(S_t) \mid S_s\right) ,$$

one easily checks that, for any $0 \le s \le t$ it holds:

$$P_{s,t}(f)(S_s) = P_{0,t}(f)(S_0) + \int_0^s \frac{\partial P_{r,t}(f)}{\partial x}(S_r)\; \sigma(S_r)\; dW_r ,$$

and if we set $s = t$ in the above equation, then we find that

$$\mathbb{E}\left[f(S_t(x)) \int_0^t \frac{\partial S_s}{\partial x}(x)\; \sigma^{-1}(S_s(x)))\; dW_s\right] = \mathbb{E}\left[\int_0^t \frac{\partial P_{s,t}(f)}{\partial x}(S_s(x))\; \frac{\partial S_s}{\partial x}(x)\; ds\right] ,$$

whenever σ is a smooth positive function bounded away from 0. Recalling that

$$\frac{\partial}{\partial x} P_{0,t}(f)(x) = \frac{\partial}{\partial x}\mathbb{E}[P_{s,t}(f)(S_s(x))] = \mathbb{E}\left[\frac{\partial P_{s,t}(f)}{\partial x}(S_s(x))\; \frac{\partial S_s}{\partial x}(x)\right] ,$$

we arrive at a *Malliavin formulation* of the semigroup derivatives

$$\frac{\partial}{\partial x} P_{0,t}(f)(x) = \mathbb{E}\left[f(S_t(x))\; \frac{1}{t} \int_0^t \sigma^{-1}(S_s(x))\; \frac{\partial S_s}{\partial x}(x)\; dW_s\right] .$$

A rigorous derivation of the above equations is provided in [56, 57]. We also refer the reader to the contribution of B. Bouchard and X. Warin in the present volume.

The Euler time discretization scheme justifies using the discrete time approximate model:

$$S_{(n+1)\Delta} - S_{n\Delta} = b(S_{n\Delta})\ \Delta + \sigma(S_{n\Delta})\ \sqrt{\Delta}\ Y_n\ , \tag{45}$$

for a sequence of independent mean zero Gaussian random variables Y_n. We thus have the approximation model

$$\frac{\partial}{\partial x} P_{0,(n+1)\Delta}(f)(x) \simeq_{\Delta\downarrow 0} \frac{1}{(n+1)\sqrt{\Delta}} \sum_{0\leq p\leq n} \mathbb{E}\left(f(S_{(n+1)\Delta}(x))\mathscr{Z}_p(x)\right)\ , \tag{46}$$

with the random weights

$$\mathscr{Z}_p(x) := \varphi\left(S_{p\Delta}(x), Y_p\right) \prod_{0\leq q<p} G_q(S_{q\Delta}(x), Y_q)$$

$$\varphi(x,y) = \sigma^{-1}(x)\ y \qquad \text{and} \quad G_q(x,y) = 1 + \frac{\partial b}{\partial x}(x)\ \Delta + \frac{\partial \sigma}{\partial x}(x)\ \sqrt{\Delta}\ y\ .$$

The ratio $1/\sqrt{\Delta}$ in the right hand side of (46) may induce numerical degeneracies. One way to overcome this problem and *to remove* this term from the numerical scheme is to use the following formula

$$\mathbb{E}\left(f(S_{(n+1)\Delta}(x))\mathscr{Z}_p(x)\right) = \mathbb{E}\left(\Upsilon_{p+1,n+1}(f)\left[S_{p\Delta}(x), Y_p\right] \times \mathscr{Z}_p(x)\right)\ ,$$

with the function

$$\begin{aligned}
&\Upsilon_{p+1,n+1}(f)[x,y] \\
&= P_{(p+1)\Delta,(n+1)\Delta}(f)\left(x + b(x)\Delta + \sigma(x)\sqrt{\Delta}y\right) - P_{(p+1)\Delta,(n+1)\Delta}(f)\left(x + b(x)\Delta\right)\ .
\end{aligned}$$

Under some appropriate regularity conditions, we notice that

$$\begin{aligned}
&\Upsilon_{p+1,n+1}(f)[x,y] \\
&\simeq_{\Delta\downarrow 0} P_{p\Delta,(n+1)\Delta}(f)\left(x + b(x)\Delta + \sigma(x)\sqrt{\Delta}y\right) - P_{p\Delta,(n+1)\Delta}(f)\left(x + b(x)\Delta\right) \\
&\simeq_{\Delta\downarrow 0} \frac{\partial P_{p\Delta,(n+1)\Delta}(f)}{\partial x}(x)\ \ \sigma(x)\sqrt{\Delta}\ y\ ,
\end{aligned}$$

which implies that

$$\begin{aligned}
&\frac{\partial}{\partial x} P_{0,(n+1)\Delta}(f)(x) \\
&\simeq_{\Delta\downarrow 0} \frac{1}{(n+1)} \sum_{0\leq p\leq n} \mathbb{E}\left(\frac{\partial P_{p\Delta,(n+1)\Delta}(f)}{\partial x}\left(S_{p\Delta}(x)\right)\ Y_p^2 \prod_{0\leq q<p} G_q(S_{q\Delta}(x), Y_q)\right)\ .
\end{aligned}$$

In higher dimensions, the calculations are more involved. To analyze these models, we design a Feynman-Kac interpretation of the distributions of product of random matrices.

Using the notation $\|\bullet\|$ for a fixed norm in $\mathbb{R}^d$, we assume that for any state U_0 in the unit sphere $\mathscr{S}^{d-1}$, we have

$$\left\| \left[\prod_{0\leq p\leq n} A_p(S_p) \right] U_0 \right\| > 0 \ .$$

In this situation, we have the multiplicative formulae

$$\left[\nabla f_{n+1}(S_{n+1}) \prod_{0\leq p\leq n} A_p(S_p) \right] U_0 = [\nabla f_{n+1}(S_{n+1})\, U_{n+1}] \prod_{0\leq p\leq n} \left\| A_p(S_p)\, U_p \right\| \ ,$$

with the well defined $\mathscr{S}^{d-1}$-valued Markov chain defined by

$$U_{n+1} = \frac{A_n(S_n)U_n}{\|A_n(S_n)U_n\|} = \frac{\left[\prod_{0\leq p\leq n} A_p(S_p)\right]\, U_0}{\left\|\left[\prod_{0\leq p\leq n} A_p(S_p)\right]\, U_0\right\|} \ .$$

If we choose $U_0 = u_0$, then we obtain the following Feynman-Kac interpretation of the gradient of a semigroup

$$\nabla P_{n+1}(f_{n+1})(x)\, u_0 = \mathbb{E}\left(F_{n+1}(\mathscr{X}_{n+1}) \prod_{0\leq p\leq n} G_p\,(\mathscr{X}_p) \right) \ .$$

In the above display, $\mathscr{X}_n$ is the *multivariate* Markov chain sequence

$$\mathscr{X}_n := (S_n, U_n, W_n)$$

and the functions F_{n+1} and G_n are defined by

$$F_{n+1}(x,u,w) := \nabla f_{n+1}(x)\, u \quad \text{and} \quad G_n\,(x,u,w) := \|\mathscr{A}_n(x,w)\, u\| \ .$$

In physics literature, the mean field particle approximations of these non commutative Feynman-Kac models are often referred as *Resampled Monte Carlo methods* [94].

6 American-style option pricing

6.1 Description of the model

Optimal stopping problems are at the heart of the theory of stochastic control. Their importance in quantitative finance is due to the large number of financial instruments with American exercises, sometimes called Bermudan exercises in the framework of discrete time models.

In this section, n stands for a fixed final time horizon, and for each $k \in \{0,\ldots,n\}$, we let $\mathscr{F}_k$ denote the set of events known at time k and $\mathscr{T}_k$ the set of $\mathscr{F}_k$ - stopping times τ taking values in $\{k,\ldots,n\}$. These stopping times are used to model the decision by the holder of the option to exercise it at a given time of his or her choice. The payoff is given by an adapted (*no crystal ball* can be used in this model!) stochastic process $Z = (Z_k)_{0\leq k\leq n}$. For each $k \in \{0,\ldots,n\}$, Z_k represents the reward to the holder for exercising the option at time k. To recast the problem in the framework used so far, we assume that the filtration $\mathscr{F} = (\mathscr{F}_k)_{0\leq k\leq n}$ is generated by a Markov chain $X = (X_k)_{0\leq k\leq n}$ in some measurable state space E, and that $Z_k = F_k(X_0,\ldots,X_k)$ for some known deterministic functions F_k on E^{k+1}. As usual, we shall use the notation $M = (M_k)_k$ to denote the transition probability of the Markov chain X.

The Snell envelope of $(Z_k)_{0\leq k\leq n}$, is the stochastic process $(U_k)_{0\leq k\leq n}$ defined for any $0 \leq k < n$ by the following backward equation

$$U_k = Z_k \vee \mathbb{E}(U_{k+1}|(X_0,\ldots,X_k)) \ ,$$

with the terminal condition $U_n = Z_n$. The main property of this stochastic process is that

$$U_k = \sup_{\tau\in\mathscr{T}_k} \mathbb{E}(Z_\tau|(X_0,\ldots,X_k)) = \mathbb{E}(Z_{\tau_k^*}|(X_0,\ldots,X_k)) \tag{47}$$

$$\text{with} \qquad \tau_k^* = \min\{k \leq l \leq n \ : \ U_l = Z_l\} \in \mathscr{T}_k \ .$$

Notice that $U_k \geq Z_k$, for any $0 \leq k \leq n$ and τ_k^* is given by the following backward formula

$$\tau_k^* = k\ \mathbf{1}_{Z_k\geq U_k} + \tau_{k+1}^*\ \mathbf{1}_{Z_k<U_k} \quad \text{with} \quad \tau_n^* = n \ .$$

To get one step further, we let $\eta_0 = \text{Law}(X_0)$ be the initial distribution on E, and we denote by $M_k(x,y)$ the elementary Markov transition of the chain X_k from E into itself.

To be more specific we also assume that

$$Z_k = F_k(X_0,\ldots,X_k) := f_k(X_k) \prod_{0\leq l<k} G_l(X_l), \qquad 0 \leq k \leq n,$$

for some non negative space-time functions $f = (f_k)_k$ and $G = (G_k)_k$ on $\{0,1,\cdots,n\} \times E$. In this situation, the Snell envelope process is given in terms of deterministic functions u_k through

$$U_k = u_k(X_k) \prod_{0\leq p<k} G_p(X_p),$$

where the functions u_k are given inductively by the backward functional equation

$$u_k = \mathscr{H}_{k+1}(u_{k+1}), \qquad 0 \leq k < n, \tag{48}$$

with the terminal value $u_n = f_n$, and the functional transformations

$$\mathscr{H}_{k+1}(u_{k+1}) := f_k \vee Q_{k+1}(u_{k+1}) \quad \text{with} \quad Q_{k+1}(x,dy) = G_k(x)\ M_{k+1}(x,dy) \ .$$

In the above displayed formula, $Q_{k+1}(u_{k+1})$ stands for the measurable function on E defined for any $x_k \in E$ by the conditional expectation:

$$Q_{k+1}(u_{k+1})(x) = G_k(x)\,\mathbb{E}\left(u_{k+1}(X_{k+1})|X_k = x\right) = \int Q_{k+1}(x,dy)\,u_{k+1}(y)\ .$$

For a detailed derivation of these formulae, we refer the interested reader to the article [41].

We let $\mathscr{H}_{k,l} = \mathscr{H}_{k+1} \circ \mathscr{H}_{k+1,l}$, with $k \leq l \leq n$, be the nonlinear semigroups associated with the backward equations (48). We use the convention $\mathscr{H}_{k,k} = I$, the identity operator, so that $u_k = \mathscr{H}_{k,l}(u_l)$, for any $k \leq l \leq n$.

If for any given sequence of bounded integral operators $(Q_k)_k$ from some state space E into itself, we denote by $Q_{k,l}$ the iterated composition operator defined by

$$Q_{k,l} := Q_{k+1}Q_{k+2}\cdots Q_l,$$

for any $k \leq l$, with the convention $Q_{k,k} = I$, then one can check that a necessary and sufficient condition for the existence of the Snell envelope $(u_k)_{0\leq k\leq n}$ is that $Q_{k,l}f_l(x) < \infty$ for any $1 \leq k \leq l \leq n$, and any state $x \in E$. To check this claim, we simply notice that

$$f_k \leq u_k \leq f_k + Q_{k+1}u_{k+1} \qquad \forall\ 1 \leq k \leq n$$

implies that

$$f_k \leq u_k \leq \sum_{k\leq l\leq n} Q_{k,l}f_l \qquad \forall\ 1 \leq k \leq n\ . \tag{49}$$

From the readily proved Lipschitz property $|\mathscr{H}_k(u) - \mathscr{H}_k(v)| \leq Q_{k+1}\left(|u-v|\right)$, for any functions u, v on E, we also have that

$$\left|\mathscr{H}_{k,l}(u) - \mathscr{H}_{k,l}(v)\right| \leq Q_{k,l}\left(|u-v|\right)\ , \tag{50}$$

for any functions u, v on E, and any $k \leq l \leq n$.

6.2 A perturbation analysis

Even if it may look innocent at first, solving numerically the recursion (48) often requires extensive calculations. The major issue is to compute the conditional expectations $M_{k+1}(u_{k+1})$ on the whole state space E, at every time step $0 \leq k < n$.

For Markov chain models taking values in some finite state spaces, the above expectations can be computed by systematic backward inspection of the realization tree that lists all possible outcomes and every transition of the chain. For excessively large state spaces, or more general situations, we need to resort to approximation strategies.

Over the last two decades, several approximation methodologies have been proposed, including Longstaff-Schwartz's functional regression style methods [18, 23, 72, 92], refined singular values decomposition strategies [9], Monte Carlo simulation methods [14, 45, 41, 42, 71], and the quantization grid technology developed by Pagès and his co-authors [76, 77, 78, 79, 80].

Most of the numerical approximation schemes amount to replacing the pair $(f_k, Q_k)_{0\leq k\leq n}$ by some approximation model $(\widehat{f}_k, \widehat{Q}_k)_{0\leq k\leq n}$ on some possibly reduced finite subsets $\widehat{E} \subset E$. We let $\widehat{u}_k$ be the Snell envelope on $\widehat{E}_k$ associated with the functions $\widehat{f}_k$ and the sequence of transition operators $\widehat{M}_k$ from $\widehat{E}$ into itself.

$$\widehat{u}_k = \widehat{\mathscr{H}_{k+1}}(\widehat{u}_{k+1}) := \widehat{f}_k \vee \widehat{Q}_{k+1}(\widehat{u}_{k+1}) \; . \tag{51}$$

Let also $\widehat{\mathscr{H}_{k,l}} = \widehat{\mathscr{H}_{k+1}} \circ \widehat{\mathscr{H}_{k+1,l}}$ with $k \leq l < n$ be the nonlinear semigroups associated with the backward equations (51) so that $\widehat{u}_k = \widehat{\mathscr{H}_{k,l}}(\widehat{u}_l)$ for any $k \leq l \leq n$. Using the elementary inequality $|a \vee a' - b \vee b'| \leq |a-b| + |a'-b'|$ which is valid for any $a, a', b, b' \in \mathbb{R}$, for any $0 \leq k < n$ and for any functions u on E_{k+1} one readily obtains the local approximation inequality

$$\left| \mathscr{H}_{k+1}(u) - \widehat{\mathscr{H}_{k+1}}(u) \right| \leq |f_k - \widehat{f}_k| + |(Q_{k+1} - \widehat{Q}_{k+1})(u)| \; . \tag{52}$$

To transfer these local estimates to the semigroups $\mathscr{H}_{k,l}$ and $\widehat{\mathscr{H}_{k,l}}$ we use a perturbation analysis. The difference between the approximate and the exact Snell envelope can be written as a telescoping sum

$$u_k - \widehat{u}_k = \sum_{l=k}^{n} \left[\widehat{\mathscr{H}_{k,l}}(\mathscr{H}_{l+1}(u_{l+1})) - \widehat{\mathscr{H}_{k,l}}(\widehat{\mathscr{H}_{l+1}}(u_{l+1})) \right] \; ,$$

setting for simplicity $\mathscr{H}_{n+1}(u_{n+1}) = u_n$ and $\widehat{\mathscr{H}_{n+1}}(u_{n+1}) = \widehat{u}_n$, for $l = n$. Combining the Lipschitz property (50) of the semigroup $\widehat{\mathscr{H}_{k,l}}$ with the local estimate (52), one gets the final estimates:

$$|u_k - \widehat{u}_k| \leq \sum_{l=k}^{n} \widehat{Q}_{k,l} |f_l - \widehat{f}_l| + \sum_{l=k}^{n-1} \widehat{Q}_{k,l} |(Q_{l+1} - \widehat{Q}_{l+1}) u_{l+1}| \; .$$

The perturbation analysis of nonlinear semigroups discussed above is a natural and fundamental tool for the analysis of the Snell envelope approximations. It can be used sequentially, and without further work, to obtain non asymptotic estimates for models combining several levels of approximations. In the same vein, and whenever possible, it can also be used as a technical tool to reduce the analysis of Snell approximation models on compact state spaces or even on finite but possibly large quantization trees or Monte Carlo grids. This perturbation analysis is clearly not new, it has been used with success in [32, 38, 63, 93] in the context of nonlinear filtering semigroups and particle approximation models. In the context of optimal stopping problems and numerical quantization schemes, these techniques were also

used for instance in the papers of Egloff [51] and Gobet, Lemor and Warin [65] or Pagès [79]. To the best of our knowledge, the general and abstract formulation given above has first been presented in the recent article [42].

6.3 Particle approximations

In this subsection, we focus on a type of Monte Carlo importance sampling scheme which is a version called *average density* of the Stochastic Mesh schemes proposed by Broadie and Glasserman in [14]. The formulation of this algorithm in terms of interacting particles was crucial to derive precise convergence results in [42].

We let η_n be the normalized Feynman-Kac measures defined in (11). By (19), we have that

$$\eta_{k+1} = \Psi_{G_k}(\eta_k) M_{k+1} .$$

Now, we assume that the Markov transitions M_k have a density H_k with respect to some reference measure λ_k

$$M_{k+1}(x,dy) = H_k(x,y)\, \lambda_k(dy) .$$

Under this assumption, we can rewrite $Q_{k+1}(u_{k+1})(x)$ as follows

$$Q_{k+1}(u_{k+1})(x) = \eta_k(G_k) \int \eta_{k+1}(y)\, \frac{G_k(x) H_{k+1}(x,y)}{\int \eta_k(dz) G_k(z) H_{k+1}(z,y)}\, u_{k+1}(y) , \qquad (53)$$

and as before, we let

$$\eta_n^N := \frac{1}{N} \sum_{i=1}^N \delta_{\xi_n^i} \to_{N\to\infty} \eta_n$$

be the particle approximation of the measures η_n defined in Section 3.2 . We denote by $\widehat{Q}_{k+1}$ the matrix obtained by replacing the measures η_k by their N-particle approximations:

$$\begin{aligned} \widehat{Q}_{k+1}(f)(x) &:= \eta_k^N(G_k) \int \eta_{k+1}^N(y)\, \frac{G_k(x) H_{k+1}(x,y)}{\int \eta_k^N(dz) G_k(z) H_{k+1}(z,y)}\, f(y) \\ &= \eta_k^N(G_k) \sum_{j=1}^N \frac{G_k(x) H_{k+1}(x,\xi_{k+1}^j)}{\sum_{j'=1}^N G_k(\xi_k^{j'}) H_{k+1}(\xi_k^{j'},\xi_{k+1}^j)}\, f(\xi_{k+1}^j) , \end{aligned}$$

for any test function f on E. Notice that these expressions are easily computed (with computational cost N^2) at any state ξ_k^i of the k-th population when the values $f(\xi_{k+1}^j)$ of the function f are known

$$\widehat{Q}_{k+1}(f)(\xi_k^i) = \eta_k^N(G_k) \sum_{j=1}^N \frac{G_k(\xi_k^i) H_{k+1}(\xi_k^i,\xi_{k+1}^j)}{\sum_{j'=1}^N G_k(\xi_k^{j'}) H_{k+1}(\xi_k^{j'},\xi_{k+1}^j)}\, f(\xi_{k+1}^j) .$$

By (51) the corresponding backward particle approximation of the Snell envelope is given by the following equations, for $i = 1, \cdots N$,

$$\widehat{u}_k(\xi_k^i) = f_k(\xi_k^i) \vee \left(\eta_k^N(G_k) \sum_{j=1}^N \frac{G_k(\xi_k^i) H_{k+1}(\xi_k^i, \xi_{k+1}^j)}{\sum_{j'=1}^N G_k(\xi_k^{j'}) H_{k+1}(\xi_k^{j'}, \xi_{k+1}^j)} \widehat{u}_{k+1}(\xi_{k+1}^j) \right) .$$

Also notice that the values $\widehat{u}_k(x)$ on any state x can be computed using the formula

$$\widehat{u}_k(x) = f_k(x) \vee \left(\eta_k^N(G_k) \sum_{j=1}^N \frac{G_k(x) H_{k+1}(x, \xi_{k+1}^j)}{\sum_{j'=1}^N G_k(\xi_k^{j'}) H_{k+1}(\xi_k^{j'}, \xi_{k+1}^j)} \widehat{u}_{k+1}(\xi_{k+1}^j) \right) .$$

For a thorough discussion on these particle models, their convergence analysis, and a variety of related approximation grid type models, we refer the reader to the pair of articles [41, 42]. In particular, this formalization allows to prove that the L_p mean error induced by this version of Stochastic Mesh approximation vanishes, under mild assumptions, with a rate $1/\sqrt{N}$. Also, a new Monte Carlo approximation scheme is proposed in [42] using simulations of a genealogical tree with neutral selections and mutations associated with a discrete-space Markov chain approximating the price dynamics. The main advantage of this new scheme is the fact that the computational effort of the algorithm is linear in the number of sampled points, as opposed to quadratic as for the Stochastic Mesh scheme.

7 Pricing models with partial observation models

Managing large portfolios and pricing financial instruments under partial observations are quite common problems in quantitative finance. See for instance the series of articles [69, 82, 84, 85], and references therein. The case of stochastic volatility models is the epitome of these situations: one can *more or less* observe stock prices but not the evolution of the stochastic volatility.

7.1 Abstract formulation and particle approximation

We work in discrete time and we recast the dynamical financial model in the framework of hidden Markov models. The basic object is a pair process $(X,Y) = ((X_n,Y_n))_n$ forming a Markov chain on some product space $E^X \times E^Y$ with elementary transitions given

$$\mathbb{P}((X_n,Y_n) \in d(x,y) \mid (X_{n-1},Y_{n-1})) = M_n(X_{n-1},dx) \times g_n(x,y)\ \lambda_n(dy)\ , \qquad (54)$$

for some positive likelihood function g_n, and some reference probability measure λ_n on E^Y. According to our setup throughout the paper, the marginal process $X = (X_n)_n$

is also assumed to be a Markov chain and as usual, we denote by M_n its transition probability. We can think of X_n as a vector of prices and random factors (instantaneous volatility could be one of them), and Y_n a vector of observations of quantities derived from the components of X_n. We can also consider X_n as a stochastic volatility model, and Y_n the stock price observations. For the sake of definiteness, we choose to illustrate the particle methods on a pricing problem, so we assume that we are given a European payoff function $f_n(X_n,Y_n)$ for each time $n \geq 0$. The price of the contingent claim is given at time $p \leq n$ by:

$$V_{p,n}(f_n) := \mathbb{E}\left[f_n(X_n,Y_n) \prod_{p\leq q<n} G'_q(X_q,Y_q) \;\middle|\; (Y_0,\ldots,Y_p) \right] .$$

for some non negative functions G'_p related to barrier sets or stochastic interest rates, as explained in Section 3. It is important to observe that the conditional expectations

$$U_{p,n}(f_n)(x,y) := \mathbb{E}\left[f_n(X_n,Y_n) \prod_{p\leq q<n} G'_q(X_q,Y_q) \;\middle|\; (X_p,Y_p) = (x,y) \right]$$

have the same form as the Feynman-Kac definitions of the measures introduced in (10), with the reference Markov chain (X_q,Y_q), from the initial time $q = p$, starting from $(X_p,Y_p) = (x,y)$ at time p. For any starting point $(X_p,Y_p) = (x,y)$, these unnormalized distributions can be approximated by running an N-particle model on $(E^X \times E^Y)$, with selection potential functions G'_q. We denote by $U^N_{p,n}(f_n)(x,y)$ the corresponding *unbiased* particle approximation. Fix an observation sequence $Y = y$, and consider the Feynman-Kac models (10) associated with the likelihood potential functions:

$$G_p(x) := g_p(x,y_p) \qquad 0 \leq p \leq n \; .$$

To emphasize the dependence of the Feynman-Kac measures on the observation sequence, we use the notations

$$\eta_{n+1}^{[y_0,\ldots,y_n]} \quad \text{and} \quad \gamma_{n+1}^{[y_0,\ldots,y_n]} \tag{55}$$

for the normalized and unnormalized measures associated with the series of observations $Y_p = y_p$, for $0 \leq p \leq n$. These conditional distributions can be approximated using an N-particle model on E^X, with selection potential functions G_q. We denote by

$$\eta_{n+1}^{([y_0,\ldots,y_n],N)} := \frac{1}{N}\sum_{i=1}^{N} \delta_{\xi_{n+1}^{([Y_0,\ldots,Y_{n-1}],i)}}$$

the empirical measures providing the particle approximation. Notice that

$$\text{Law}(X_n \mid Y_p = y_p,\ 0 \leq p < n) = \eta_n^{[y_0,\ldots,y_{n-1}]} \tag{56}$$

and by the Bayes rule

$$\text{Law}(X_n \mid Y_p = y_p,\ 0 \le p \le n) := \Psi_{g_n(\bullet,y_n)}\left(\eta_n^{[y_0,\ldots,y_{n-1}]}\right) = \widehat{\eta}_n^{[y_0,\ldots,y_n]} \ , \qquad (57)$$

and by construction, we have:

$$V_{p,n}(f_n) = \int \Psi_{g_p(\bullet,Y_p)}\left(\eta_n^{[Y_0,\ldots,Y_{n-1}]}\right)(dx)\ U_{p,n}(f_n)(x,Y_p) \ ,$$

and these quantities can be approximated combining the particle estimates defined above. Indeed, we have that

$$V_{p,n}(f_n) \simeq_{N\uparrow\infty} V_{p,n}^N(f_n) \ ,$$

with

$$\begin{aligned} V_{p,n}^N(f_n) &:= \int \Psi_{g_p(\bullet,Y_p)}\left(\eta_n^{([Y_0,\ldots,Y_{n-1}],N)}\right)(dx)\ U_{p,n}^N(f_n)(x,Y_p) \\ &= \sum_{i=1}^N \frac{g_p(\xi_p^{([Y_0,\ldots,Y_{p-1}],i)},Y_p)}{\sum_{j=1}^N g_p(\xi_p^{([Y_0,\ldots,Y_{p-1}],j)},Y_p)}\ U_{p,n}^N(f_n)(\xi_p^{([Y_0,\ldots,Y_{p-1}],i)},Y_p) \ . \end{aligned}$$

7.2 Optimal stopping with partial observation

We work with the setup of a pair (X_n,Y_n) Markov chain model introduced in the previous section. According to our discussion in Section 6.1, the Snell envelop associated with an American option with finite maturity n, payoffs $Z_k = f_k(X_k,Y_k)$ is given by

$$U_k := \sup_{\tau\in\mathscr{T}_k^Y} \mathbb{E}(f_\tau(X_\tau,Y_\tau)|(Y_0,\ldots,Y_k)) \ ,$$

where $\mathscr{T}_k^Y$ stands for the set of all $\mathscr{F}_k^Y$ - stopping times τ taking values in $\{k,\ldots,n\}$, where the filtration is know given by the sigma fields $\mathscr{F}_k^Y$ generated by the observation sequence Y_p, from $p=0$ up to the time k. We denote by $\eta_n^{[y_0,\ldots,y_{n-1}]}$ and $\widehat{\eta}_n^{[y_0,\ldots,y_n]}$ the conditional distributions defined in (56) and (57). With these notations, for any $0\le k\le n$ we have that

$$\begin{aligned} \mathbb{E}(f_\tau(X_\tau,Y_\tau)|(Y_0,\ldots,Y_k)) &= \mathbb{E}\left(\sum_{p=k}^n \mathbf{1}_{\tau=p}\ \mathbb{E}\left(f_p(X_p,Y_p) \mid (Y_0,\ldots,Y_p)\right) \mid (Y_0,\ldots,Y_k)\right) \\ &= \mathbb{E}\left(F_\tau\left(Y_\tau,\widehat{\eta}_\tau^{[Y_0,\ldots,Y_\tau]}\right) \mid (Y_0,\ldots,Y_k)\right) \ , \qquad (58) \end{aligned}$$

with the conditional payoff function

$$F_p\left(Y_p,\widehat{\eta}_p^{[Y_0,\ldots,Y_p]}\right) = \int \widehat{\eta}_p^{[Y_0,\ldots,Y_p]}(dx)\ f_p(x,Y_p) \ .$$

It is well known that

$$\mathscr{X}_p := \left(X_p, Y_p, \widehat{\eta}_p^{[Y_0,\ldots,Y_p]}\right)$$

is a Markov chain with elementary transitions defined by, for any integrable function F on product space $E^X \times E^Y \times \mathscr{P}(E^X)$,

$$\mathbb{E}\left[F\left(X_p, Y_p, \widehat{\eta}_p^{[Y_0,\ldots,Y_p]}\right) \;\middle|\; \left(X_{p-1}, Y_{p-1}, \widehat{\eta}_{p-1}^{[Y_0,\ldots,Y_{p-1}]}\right) = (x,y,\mu)\right]$$

$$= \int\int \lambda_p(dy_p)\, M_p(x,dx_p)\, g_p(x_p,y_p)\, F\left(x_p, y_p, \Psi_{g_p(\bullet,y_p)}(\mu M_p)\right) .$$

A proof of this assertion can be found in any textbook on advanced stochastic filtering. For instance, the book of W. Runggaldier and L. Stettner [90] provides a detailed treatment of discrete time partially observed models, their non linear filtering, and related partially observed control problems.

Roughly speaking, using Bayesian notation, we have

$$\begin{aligned}\eta_p^{[y_0,\ldots,y_{p-1}]}(dx_p) &= dp_p(x_p \mid (y_0,\ldots,y_{p-1})) \\ &= \int dp_p(x_p \mid x_{p-1}) \times p_n(x_{p-1} \mid (y_0,\ldots,y_{p-1})) \\ &= \widehat{\eta}_{p-1}^{[y_0,\ldots,y_{p-1}]} M_p(dx_p)\end{aligned}$$

and

$$\Psi_{g_p(\bullet,y_p)}\left(\widehat{\eta}_{p-1}^{[y_0,\ldots,y_{p-1}]} M_p\right)(dx_p)$$

$$= \frac{p(y_p|x_p)}{\int p_p(y_p \mid x'_p)\, dp_p(x'_p \mid (y_0,\ldots,y_{p-1}))}\, dp_p(x_p \mid (y_0,\ldots,y_{p-1}))$$

$$= dp_p(x_p \mid (y_0,\ldots,y_{p-1},y_p)) ,$$

from which we can prove that

$$\begin{aligned}\mu M_p(g_p(\bullet,y_p)) &= \int p_p(y_p \mid x_p)\, dp_p(x_p \mid (y_0,\ldots,y_{p-1})) \\ &= p_p(y_p \mid (y_0,\ldots,y_{p-1}))\end{aligned}$$

and

$$\Psi_{g_p(\bullet,y_p)}(\mu M_p) = \widehat{\eta}_p^{[y_0,\ldots,y_p]} ,$$

as long as $\mu = \widehat{\eta}_{p-1}^{[y_0,\ldots,y_{p-1}]}$ $\left(\Rightarrow \mu M_p = \eta_p^{[y_0,\ldots,y_{p-1}]}\right)$.

The above discussion suggests the following interpretation. We can rewrite (58) as the Snell envelop of a fully observed *augmented* Markov chain:

$$\mathbb{E}(f_\tau(X_\tau,Y_\tau)|(Y_0,\ldots,Y_k)) = \mathbb{E}\left(F_\tau\left(\mathscr{X}_\tau\right) \mid (\mathscr{X}_0,\ldots,\mathscr{X}_k)\right) .$$

This Markov chain $\mathscr{X}_n$ takes values in an infinite dimensional state space, and it can rarely be sampled without some addition level of approximation. Therefore, most of the grid or Monte Carlo simulation based techniques for solving these models require the introduction of a specific grid approximation of conditional distributions, or judicious approximation sampling schemes. The particle methodology advocated in this paper provides a natural strategy. Using the particle approximations discussed in Section 3.2, we can replace the chain $\mathscr{X}_n$ by the N-particle approximation defined by

$$\mathscr{X}_n^N := \left(Y_p, \widehat{\eta}_p^{([Y_0,\ldots,Y_p],N)}\right) ,$$

where

$$\widehat{\eta}_p^{([Y_0,\ldots,Y_p],N)} := \Psi_{g_p(\cdot,Y_p)}\left(\widehat{\eta}_{p-1}^{([Y_0,\ldots,Y_{p-1},N)]}\right)$$

stands for the updated measure associated with the particle scheme associated with the likelihood selection functions $g_p(\cdot,Y_p)$. The corresponding N-particle approximation of the Snell envelop is now given by

$$\mathbb{E}(f_\tau(X_\tau,Y_\tau)|(Y_0,\ldots,Y_k)) \simeq_{N\uparrow\infty} \mathbb{E}\left(F_\tau\left(\mathscr{X}_\tau^N\right) \mid (\mathscr{X}_0^N,\ldots,\mathscr{X}_k^N)\right) .$$

In this setup, the approximated optimal stopping problem requires the computation of the quantities

$$U_k^N := \sup_{\tau\in\mathscr{T}_k^N} \mathbb{E}\left(F_\tau\left(\mathscr{X}_\tau^N\right) \mid (\mathscr{X}_0^N,\ldots,\mathscr{X}_k^N)\right) ,$$

where $\mathscr{T}_k^N$ stands for the set of $\mathscr{F}_k^{\mathscr{X}_k^N}$ - stopping times τ taking values in $\{k,\ldots,n\}$, where the filtration is formed by the sigma fields generated by the Markov chain random variables $\mathscr{X}_k^N$, from $p=0$ up to the current time k.

We close this section with an alternative representation in terms of the unnormalized filters $\gamma_n^{[Y_0,\ldots,Y_{n-1}]}$ defined in (55). We let $\mathbb{P}_n$ be the probability distribution of a Markov chain $(X_p,Y_p)_{0\leq p\leq n}$ defined in (54), and $\mathbb{P}_n^{(0)}$ the probability distribution of the Markov chain $(X_p,Y_p)_{0\leq p\leq n}$ with independent random observations Y_p with distribution λ_p with $p\leq n$, also assume to be *independent of the chain* $(X_p)_{0\leq p\leq n}$,. By construction, $\mathbb{P}_n$ is absolutely continuous with respect to $\mathbb{P}_n^{(0)}$, and its Radon-Nykodym derivative is given by:

$$\frac{d\mathbb{P}_n}{d\mathbb{P}_n^{(0)}} = \prod_{0\leq p\leq n} g_p(X_p,Y_p) .$$

Now, for any $\tau\in\mathscr{T}_0^Y$ we observe that

$$\mathbb{E}(f_\tau(X_\tau,Y_\tau)) = \mathbb{E}^{(0)}\left(\sum_{p=0}^n \mathbf{1}_{\tau=p}\,\mathbb{E}^{(0)}\left(f_p(X_p,Y_p)\prod_{0\leq q\leq p} g_q(X_q,Y_q) \mid (Y_0,\ldots,Y_p)\right)\right)$$

and

$$\begin{aligned}
&\mathbb{E}^{(0)}\left(f_p(X_p,Y_p)\ \textstyle\prod_{0\le q\le p} g_q(X_q,Y_q)\mid (Y_0,\ldots,Y_p)\right)\\
&=\gamma_p^{[Y_0,\ldots,Y_{p-1}]}\left(f_p(\bullet,Y_p)g_p(\cdot,Y_p)\right)\\
&=\eta_p^{[Y_0,\ldots,Y_{p-1}]}\left(f_p(\bullet,Y_p)g_p(\cdot,Y_p)\right)\times\textstyle\prod_{0\le q<n}\eta_q^{[Y_0,\ldots,Y_{q-1}]}\left(g_q(\cdot,Y_q)\right)\ .
\end{aligned}$$

The last assertion is a direct consequence of the multiplicative formula (12) for unnormalized Feynman-Kac measures. Arguing as above, we introduce the Markov chain

$$\mathscr{X}_n:=\left(Y_n,\eta_n^{[Y_0,\ldots,Y_{n-1}]}\right)$$

the payoff and the potential functions

$$F_n(\mathscr{X}_n):=\eta_n^{[Y_0,\ldots,Y_{n-1}]}\left(f_n(\bullet,Y_n)g_n(\bullet,Y_n)\right)\quad\text{and}\quad\mathscr{G}_n(\mathscr{X}_n):=\eta_n^{[Y_0,\ldots,Y_{n-1}]}\left(g_n(\bullet,Y_n)\right)\ .$$

By construction, we have

$$\mathbb{E}\left(f_\tau(X_\tau,Y_\tau)\right)=\mathbb{E}^{(0)}\left(F_\tau(\mathscr{X}_\tau)\prod_{0\le p<\tau}\mathscr{G}_p(\mathscr{X}_p)\right)\ .$$

We have now reduced the optimal stopping problem with partial observations to a conventional optimal stopping problem of a measure valued Markov chain $\mathscr{X}_n$ with *stochastic potential functions* $\mathscr{G}_p(\mathscr{X}_p)$, and independent random observations sequences. Once more, using the particle approximation models discussed in Section 3.2, we can replace the chain $\mathscr{X}_n$ by the N-particle approximation model defined by

$$\mathscr{X}_n^N:=\left(Y_p,\eta_p^{([Y_0,\ldots,Y_p],N)}\right)\ .$$

Here again, we have turned a complex optimal stopping problem under partial observations into an almost equivalent optimal stopping problem of an *easy to sample* Markov chain sequence of the same form as the one discussed in Section 6.1. These particle transformations can also be used for more general stochastic control problems with partial observations. We refer the reader to [8, 90] for a more thorough discussion on this subject.

7.3 Parameter estimation in hidden Markov chain models

In many economic and financial applications, the parameters are unknown and must be estimated from partial and noisy observations. This situation is typical of hidden Markov chain problems which arise in a variety of domains, ranging from signal processing, medical Bayesian inference, communication and information theory. For an overview of some of the problems occurring in finance and econometrics we refer the reader to [4, 24, 25, 28, 61], to mention only a few.

As in (54) these models are framed in terms of a signal-observation type pair Markov chain $(X_n,Y_n)_n$ with a collection of transition probabilities $M_{\theta,n}$ and likeli-

hood functions $g_{\theta,n}$ that depend upon the realization of a random parameter $\Theta = \theta$ taking values in some state space S, equipped with a probability measure μ. We also denote by $\eta_{\theta,0}$ the conditional distribution of X_0 given $\Theta = \theta$.

The example we have in mind is the quintessential calibration problem in partially observed models arising in computational finance. One way to set up a stochastic volatility model as a filtering problem in discrete time is to choose $(X_n, Y_n) = (\sigma_n, S_n)$. In this case $X_n = \sigma_n$ represents the instantaneous stochastic volatility, and the observation $Y_n = S_n$ is given by the price of the asset. In most practical applications, the evolution of these quantities is given by a parametric model of the form:

$$\begin{aligned}\sigma_k &= F^1_{\theta,n}(\sigma_{n-1}, S_{n-1}, W^1_n)\\ S_k &= F^2_{\theta,n}(\sigma_{n-1}, S_{n-1}, W^2_n),\end{aligned}$$

where $F^1_{\theta,n}$ and $F^2_{\theta,n}$ are functions depending upon some unknown parameter θ. The objective is to compute the conditional distribution $\mathrm{Law}(\theta|S_0,\ldots,S_n)$ of θ given the observations of the price. To be more specific, we can precise our illustration by choosing the popular Heston's stochastic volatility model. In our framework, this model is given by:

$$\begin{aligned}F^1_{\theta,n}(\sigma_{n-1}, S_{n-1}, W^1_n) &= (ab + (1-a)\sigma_{n-1})\Delta t + c\sqrt{\sigma_{n-1}}\Delta W^1_n\\ F^2_{\theta,n}(\sigma_{n-1}, S_{n-1}, W^2_n) &= S_{k-1}(1+d)\Delta t + S_{k-1}\sqrt{\sigma_{n-1}}\Delta W^2_n,\end{aligned}$$

where $\theta = (a,b,c,d)$ is the collection of parameters to calibrate.

Using the notations of Section 3.3.2, the conditional distribution of the random path $(X_0,\ldots,X_n)$, given $\Theta = \theta$, and the sequence $(Y_0,\ldots,Y_{n-1}) = (y_0,\ldots,y_{n-1})$ of observations is given by the Feynman-Kac measures

$$\mathbb{Q}_{\theta,n}(dx_0 \times \cdots \times dx_n) = \frac{1}{\mathscr{Z}_n(\theta)} \left\{ \prod_{0\le q<n} G_{\theta,q}(x_q) \right\} \mathbb{P}_{\theta,n}(dx_0 \times \cdots \times dx_n)\ ,$$

with the potential functions $G_{\theta,q}(x_q) = g_{\theta,n}(x_q, y_q)$, and the conditional distribution $\mathbb{P}_{\theta,n}$ of the random path $(X_0,\ldots,X_n)$ given $\Theta = \theta$. As in (12), the normalizing constants $\mathscr{Z}_n(\theta)$ are given by the multiplicative formula

$$\mathscr{Z}_n(\theta) = \prod_{0\le p<n} \eta^{[y_0,\ldots,y_{p-1}]}_{\theta,p}(G_{\theta,p})\ ,$$

with the p-th marginal distributions $\eta^{[y_0,\ldots,y_{p-1}]}_{\theta,p}$ of the measure $\mathbb{Q}_{\theta,n}$ i.e. the conditional distribution of the random variable X_p given $\Theta = \theta$, and the sequence $(Y_0,\ldots,Y_{p-1}) = (y_0,\ldots,y_{p-1})$ of observations. In the Bayesian literature, the normalizing constants $\mathscr{Z}_{n+1}(\theta)$ are often called the likelihood functions of the parameter θ, given the observation data $(y_0,\ldots,y_n)$, and they are denoted by $p(y_0,\ldots,y_n \mid \theta)$

to emphasize that they are given by the conditional density of the observations given the unknown parameter.

In this section, the observation sequence $Y = y$ is fixed, so in order to streamline the notations we suppress the superscript $^{[y_0,\ldots,y_n]}$and write $\eta_{\theta,p}$ and $\widehat{\eta}_{\theta,p} = \Psi_{G_{\theta,n}}\left(\eta_{\theta,p}\right)$ for the one step predictor $\eta_{\theta,p}^{[y_0,\ldots,y_{p-1}]}$ and the optimal filter $\widehat{\eta}_{\theta,p}^{[y_0,\ldots,y_p]}$.

From the previous discussion, it should be clear that the conditional distributions of the parameter Θ with respect to the sequence of observations $(Y_0,\ldots,Y_{n-1}) = (y_0,\ldots,y_{n-1})$ is given by the measures

$$\mu_n(d\theta) := \frac{1}{\mathscr{Z}_n}\left(\prod_{0\le p<n} h_p(\theta)\right)\mu(d\theta) \quad \text{with the functions} \quad h_p(\theta) = \eta_{\theta,p}(G_{\theta,p}) \tag{59}$$

for some normalizing constant $\mathscr{Z}_n$. In the Bayesian literature, the likelihood functions $h_p(\theta)$ are often denoted by $p(y_p \mid (y_0,\ldots,y_{p-1}),\theta)$. In some instances, such as classical linear-Gaussian models for example, the local likelihood functions $h_p(\theta)$ can be computed explicitly in terms of Gaussian densities and optimal one-step predictors given by the Kalman recursions. In this case, we can use a dedicated Monte Carlo Markov Chain model (MCMC for short) algorithm to sample from the Boltzmann-Gibbs measures (59). One can also turn this MCMC algorithm into an interacting MCMC model. This is done by letting K_n be a MCMC transition with target measure $\mu_n = \mu_n K_n$. By definition of the Boltzmann-Gibbs transformation (16), we readily see that

$$\mu_{n+1} = \Psi_{h_n}(\mu_n) \Rightarrow \mu_{n+1} = \Psi_{h_n}(\mu_n)K_{n+1},$$

which shows that μ_n is given by the normalized Feynman-Kac measure defined for any measurable function f on S, by the following equation

$$\mu_n(f) \propto \mathbb{E}\left(f(\Theta_n)\prod_{0\le p<n} h_p(\Theta_p)\right),$$

where Θ_n is a Markov chain on S with initial distribution $\mu_0 = \mu$ and Markov transitions K_n. The interacting particle approximation

$$\mu_n^N = \frac{1}{N}\sum_{1\le i\le N}\delta_{\Theta_n^i}$$

of the measures μ_n (and their normalizing constants) is a genetic type particle model on the product space S^N

$$\Theta_n = \left(\Theta_n^i\right)_{1\le i\le N} \in S^N \xrightarrow{\quad selection \quad} \widehat{\Theta}_n = \left(\widehat{\Theta}_n^i\right)_{1\le i\le N} \in S^N \xrightarrow{\quad mutation \quad} \Theta_{n+1} \in S^N. \tag{60}$$

The mutation transitions are given by the MCMC transitions K_n, and the selection transitions are obtained from the selection potential functions h_n. The complete con-

ditional distribution of the random sequence $(\Theta,(X_0,\ldots,X_n))$ given the sequence of observations $(Y_0,\ldots,Y_{n-1})=(y_0,\ldots,y_{n-1})$ is given by the Feynman-Kac measures

$$\mu_n(d\theta)\times\mathbb{Q}_{\theta,n}(d(x_0,\ldots,x_n))\simeq_{N\uparrow\infty}\mu_n^N(d\theta)\times\mathbb{Q}_{\theta,n}^N(d(x_0,\ldots,x_n))\ .$$

The measures $\mathbb{Q}_{\theta,n}^N$ appearing in the above right hand side stand for the particle backward model defined in (28). Alternatively, we can also use the genealogical tree approximation discussed in (3.3.1).

For linear-Gaussian models, we emphasize that the measure $\mathbb{Q}_{\theta,n}$ can be computed explicitly. More precisely, the backward Markov chain formula (27) can be computed using the updating transition of the Kalman filter, with the Gaussian likelihood density function $H_{\theta,n+1}$ of the transition $M_{\theta,n+1}$. In this case, (27) is the backward product of the Gaussian transitions given below

$$\mathbb{M}_{\theta,n+1,\eta_{\theta,n}}(x_{n+1},dx_n):=\frac{H_{\theta,n+1}(x_n,x_{n+1})}{\widehat{\eta}_{\theta,n}\left(H_{\theta,n+1}(\bullet,x_{n+1})\right)}\ \widehat{\eta}_{\theta,n}(dx_n)\ .$$

When the local likelihood functions h_n are not known, we need to add another approximation level. To this end, we also consider the probability distribution $P(\theta,d\xi)$ of the N-particle model

$$\xi_\theta:=\left(\xi_{\theta,0},\xi_{\theta,1},\ldots,\xi_{\theta,T}\right)\ ,$$

on the interval $[0,T]$, with mutation transitions $M_{\theta,n}$, and potential selection functions $G_{\theta,n}$, with $n\leq T$. We fix a large time horizon T, and for any $0\leq n\leq T$, we set

$$\overline{\mu}_n(d(\xi,\theta))=\frac{1}{\overline{\mathscr{Z}}_n}\left\{\prod_{0\leq p<n}\overline{h}_p(\xi,\theta)\right\}\overline{\mu}(d(\xi,\theta))\ , \tag{61}$$

for some normalizing constants $\overline{\mathscr{Z}}_n$, the reference measure $\overline{\mu}$ being given by

$$\overline{\mu}(d(\xi,\theta))=\mu(d\theta)\ P(\theta,d\xi),$$

and the potential functions h_n on the product space $\left(\left(\prod_{0\leq p\leq T}E_p^N\right)\times S\right)$ defined by

$$\overline{h}_n(\xi,\theta)=\frac{1}{N}\sum_{1\leq i\leq N}G_{\theta,n}(\xi_{\theta,n}^i)=\eta_{\theta,n}^N\left(G_{\theta,n}\right)\in(0,\infty)\ .$$

Firstly, we observe that these target measures have the same form as the Boltzmann-Gibbs measures (59). Thus, they can be sampled using the MCMC or the interacting MCMC methodologies discused above. For a detailed discussion these types of sophisticated serial MCMC methodologies, we refer the reader to the recent article [1].

More interestingly, using the unbiased property of the unnormalized particle models presented in (32), we clearly have $\overline{\mathscr{Z}}_n=\mathscr{Z}_n$ and

$$\int P(\theta, d\xi) \left\{ \prod_{0\le p<n} \overline{h}_p(\xi,\theta) \right\} = \mathbb{E}\left(\prod_{0\le p<n} \eta^N_{\theta,p}\left(G_{\theta,p}\right) \right) = \prod_{0\le p<n} \eta_{\theta,p}(G_{\theta,p}) \ ,$$

from which we conclude that the Θ-marginal of $\overline{\mu}_n$ coincides with the desired target measure

$$\left(\overline{\mu}_n \circ \Theta^{-1}\right)(d\theta) = \mu_n(d\theta) = \frac{1}{\mathscr{Z}_n} \left\{ \prod_{0\le p<n} \eta_{\theta,p}(G_{\theta,p}) \right\} \nu(d\theta) \ .$$

References

1. C. Andrieu, A. Doucet, and R. Holenstein. Particle Markov chain Monte Carlo methods. *Journal Royal Statistical Society B*, vol. 72, no. 3, pp. 269-342 (2010).
2. Y. Achdou and O. Pironneau. Computational methods for option pricing. *SIAM, Frontiers in Applied Mathematics series*, (2005).
3. O. Bardou, S. Bouthemy and G. Pagès. Optimal quantization for the pricing of swing options. *Applied Mathematical Finance*, vol. 16, Issue 2, pp. 183 - 217 (2009).
4. R. Bahr and S. Hamori. *Hidden Markov models : Applications to finance and economics*. Advanced studies in theoretical and applied econometrics. vol. 40, Kluwer Academic Publishers (2004).
5. C. Bender and J. Steiner. Least-squares Monte Carlo for backward SDEs. *Numerical Methods in finance, Springer* (2011).
6. S. Ben Hamida and R. Cont. Recovering Volatility from Option Prices by Evolutionary Optimization. *Journal of Computational Finance*, Vol. 8, No. 4 (2005).
7. T. Björk, M. H A Davis and C. Landèn. Optimal investment under partial information *Mathematical Methods of Operations Research* (2010).
8. V. S. Borkar. Controlled diffusion processes. *Probability Surveys*, Vol. 2, pp. 213-244 (2005).
9. B. Bouchard and X. Warin. Monte-Carlo valuation of American options: facts and new algorithms to improve existing methods. *Numerical Methods in finance, Springer* (2011).
10. J. P. Bouchaud and M. Potters. Back to basics: historical option pricing revisited. *Philosophical Transactions: Mathematical, Physical & Engineering Sciences*, 357(1758), pp. 2019-2028 (2002).
11. P. P. Boyle and D. Emanuel. Discretely adjusted option hedges. *Journal of Financial Economics*, vol. 8, pp. 259-282 (1980).
12. D. Brigo, T. Bielecki and F. Patras. *Credit Risk Frontiers*. Wiley–Bloomberg Press (2011).
13. M. Broadie and P. Glasserman. Estimating security prices using simulation. *Management Science*, 42, pp. 269–285 (1996).
14. M. Broadie and P. Glasserman. A Stochastic Mesh Method for Pricing High-Dimensional American Options. *Journal of Computational Finance*, vol. 7, pp. 35–72, (2004).
15. R. Carmona and S. Crépey. Importance Sampling and Interacting Particle Systems forthe Estimation of Markovian Credit Portfolios Loss Distribution. *International Journal of Theoretical and Applied Finance*, vol. 13, No. 4 (2010) 577 – 602.
16. R. Carmona, J.-P. Fouque and D. Vestal. Interacting Particle Systems for the Computation of Rare Credit Portfolio Losses. *Finance and Stochastics*, vol. 13, no. 4, 2009 pp. 613-633 (2009).
17. R. Carmona and N. Touzi. Optimal Multiple Stopping and Valuation of Swing Options. *Mathematical Finance*, Vol. 18, Issue 2, pp. 239-268 (2008).

18. J. F. Carrière. Valuation of the Early-Exercise Price for Options using Simulations and Nonparametric Regression. *Insurance : Mathematics and Economics*, 19, 19-30 (1996).
19. R. Casarin. Simulation Methods for Nonlinear and Non-Gaussian Models in Finance. *Premio SIE, Rivista Italiana degli Economisti*, vol. 2, pp. 341-345 (2005).
20. R. Casarin and C. Trecroci. Business Cycle and Stock Market Volatility: A Particle Filter Approach, *Cahier du CEREMADE N. 0610*, University Paris Dauphine (2006).
21. A. Cerny. Dynamic programming and mean variance hedging in discrete time. *Tanaka Business School Discussion Papers* TBS/DP04/15 London: Tanaka Business School (2004).
22. N. Chen and P. Glasserman. Malliavin Greeks without Malliavin calculus. Stochastic Processes and their Applications 117, pp. 1689 – 1723 (2007).
23. E. Clément, D. Lamberton and P. Protter. An analysis of a least squares regression method for American option pricing. *Finance and Stochastics*, 6, 449-472, (2002).
24. R. Cont and P. Tankov. Non-parametric calibration of jump-diffusion option pricing models. *Journal of computational finance*, vol. 7, no. 3, pp. 1–49, (2004).
25. R. Cont and P. Tankov. Retrieving Levy processes from option prices: regularization of ill-posed inverse problem. *SIAM Journal on Control and Optimization*, vol. 45, no. 1, pp. 1–25 (2007).
26. A.L. Corcoran and R.L. Wainwright. A parallel island model genetic algorithm for the multiprocessor scheduling problem. *Proceedings of the 1994 ACM/SIGAPP Symposium on Applied Computing*, March 6-8, pp. 483-487, ACM Press (1994).
27. D. D. Creal. A survey of sequential Monte Carlo methods for economics and finance. *To appear in Econometric Reviews* (2011).
28. S. Crépey. Calibration of the local volatility in a trinomial tree using Tikhonov regularization. *Inverse Problems*, vol. 19, pp. 91–127 (2003)
29. D. Crisan, P. Del Moral and T. Lyons. Interacting Particle Systems Approximations of the Kushner Stratonovitch Equation. *Advances in Applied Probability*, vol.31, no. 3, 819-838 (1999).
30. Del Moral P. Non Linear Filtering: Interacting Particle Solution. *Markov Processes and Related Fields*, Volume 2 Number 4, 555–580 (1996).
31. P. Del Moral. *Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications*, New York: Springer-Verlag (2004).
32. P. Del Moral, Measure Valued Processes and Interacting Particle Systems. Application to Nonlinear Filtering Problems. *Annals of Applied Probabability*, vol. 8, no. 2, pp. 1254-1278 (1998).
33. P. Del Moral, A. Doucet and A. Jasra. Sequential Monte Carlo samplers. *J. Royal Statist. Soc.* B, vol. 68, pp. 411–436 (2006).
34. P. Del Moral and A. Doucet. *Sequential Monte Carlo and Genetic particle models. Theory and Practice*, in preparation, Chapman & Hall (2011).
35. P. Del Moral, A. Doucet and A. Jasra. On Adaptive Resampling Procedures for Sequential Monte Carlo Methods. HAL-INRIA RR-6700-2008. To appear in Bernoulli (2011).
36. P. Del Moral, A. Doucet and S. S. Singh. A Backward Particle Interpretation of Feynman-Kac Formulae *HAL-INRIA RR-7019 (07-2009)*, *M2AN*, vol 44, no. 5, pp. 947–976 M2AN (sept. 2010).
37. P. Del Moral, J. Garnier. Genealogical Particle Analysis of Rare events. *Annals of Applied Probability*, vol. 15, no. 4, 2496–2534 (2005).
38. P. Del Moral and A. Guionnet. On the stability of interacting processes with applications to filtering and genetic algorithms, *Annales de l'Institut Henri Poincaré*, Vol. 37, No. 2, 155–194 (2001).
39. P. Del Moral and A. Jasra. SMC for option pricing. To appear in *Stochastic Analysis and Applications* Volume 29, Issue 2, pp. 292–316 (2011).
40. P. Del Moral and L. Miclo. *Branching and Interacting Particle Systems Approximations of Feynman-Kac Formulae with Applications to Non-Linear Filtering*. Séminaire de Probabilits XXXIV, Ed. J. Azma and M. Emery and M. Ledoux and M. Yor, Lecture Notes in Mathematics, Springer-Verlag Berlin, Vol. 1729, 1–145 (2000).

41. P. Del Moral, P. Hu and N. Oudjane. Snell envelope with path dependent multiplicative optimality criteria *HAL-INRIA*, RR-7360 (2010).
42. P. Del Moral, P. Hu, N. Oudjane and Br. Rémillard. On the Robustness of the Snell envelope. *HAL-INRIA*, RR-7303, to appear in *SIAM Journal on Financial Mathematics* (2011).
43. P. Del Moral, P. Hu and L. Wu. On the concentration properties of Interacting particle processes. *HAL-INRIA*, RR-7677, to appear in *Foundations and Trends in Machine Learning* (2011).
44. P. Del Moral and Fr. Patras. Interacting path systems for credit risk. *Credit Risk Frontiers*. D. Brigo, T. Bielecki, F. Patras Eds. *Wiley–Bloomberg Press*, (2011), 649–674. *Short announcement available as Interacting path systems for credit portfolios risk analysis. INRIA:RR-7196 (2010)*.
45. P. Del Moral, Br. Rémillard and S. Rubenthaler. Monte Carlo approximations of American options that preserve monotonicity and convexity. (2011).
46. P. Del Moral and E. Rio. Concentration Inequalities for Mean Field Particle Models. HAL-INRIA RR-6901, to appear in the *Annals of Applied Probability* (2011).
47. P. Del Moral, J. Jacod and P. Protter. The Monte Carlo method for filtering with discrete time observations. *Probability Theory and Related Fields*, vol. 120, no. 3, pp. 346–368 (2001).
48. A. Doucet, L. Montesano and A. Jasra. Optimal filtering for partially observed point processes. *Proceedings ICASSP* (2006).
49. A. Doucet, J. F. G. de Freitas and N. J. Gordon (eds.) *Sequential Monte Carlo Methods in Practice*. New York: Springer-Verlag, (2001).
50. P. Dybvig. Distributional analysis of portfolio choice. *The Journal of Business*, 61(3):369–393 (1988).
51. D. Egloff. Monte Carlo algorithms for optimal stopping and statistical learning. *Annals of Applied Probability* , 15, pp. 1–37 (2005).
52. N. El Karoui, S. Peng and M. C. Quenez. Backward stochastic differential equations in finance. *Mathematical Finance*, Vol. 7, no. 1, pp. 1–71 (1997).
53. J. J. Florentin. Optimal observability and optimal control. *J. of Electronics and Control*. Vol. 13, pp. 263–279 (1962).
54. I. Florescu and Fr. Viens. Stochastic Volatility: Option Pricing using a Multinomial Recombining Tree. *Applied Mathematical Finance*, vol. 15, Issue 2, pp.151–181 (2008).
55. J. P. Fouque, G. Papanicolaou and R. Sircar. Derivatives in financial markets with stochastic volatility. *Cambridge University Press* (2000).
56. E. Fournié, J. M. Lasry, J. Lebuchoux, P. L. Lions and N. Touzi. Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance and Stochastics*, vol. 3, pp. 391–412 (1999).
57. E. Fournié, J. M. Lasry, J. Lebuchoux and P. L. Lions. Applications of Malliavin calculus to Monte Carlo methods in finance. II. *Finance and Stochastics*, vol. 5, pp. 201–236 (2001).
58. M. C. Fu, D. B. Madan and T. Wang. Pricing continuous Asian options: a comparison of Monte Carlo and Laplace transform inversion methods. *Journal of Computational Finance*, vol. 2, pp. 49–74, (1998).
59. D. Gasbarra. Particle filters for counting process observations. http://www.rni.helsinki.fi/∼dag/newpart2.ps, Research report Helsinki University, (2001).
60. V. Genon-Catalot, Th. Jeantheau and C. Laredo. Conditional Likelihood Estimators for Hidden Markov Models and Stochastic Volatility Models. *Scandinavian Journal of Statistics*, Vol 30: 297–316 (2003).
61. R. Lagnado and S. Osher. A technique for calibrating derivative security pricing models: numerical solution of the inverse problem. *Journal of computational finance*, vol. 1, pp. 13–25 (1997).
62. F. Le Gland and N. Oudjane. A robustification approach to stability and to uniform particle approximation of nonlinear filters: the example of pseudo-mixing signals. *Stochastic Processes and their Applications*, vol. 106, no. 2, pp. 279–316 (2003).
63. Fr. Le Gland and N. Oudjane. Stability and uniform approximation of nonlinear filters using the Hilbert metric and application to particle filters. *Annals Applied Probability*, Vol. 14, no. 1, 144-187 (2004).

64. M. Giles and P. Glasserman, Smoking adjoints: fast Monte Carlo Greeks. *Risk*, pp. 92–96 (2006).
65. E. Gobet, J. P. Lemor and X. Warin. A regression-based Monte-Carlo method for backward stochastic differential equations. *Annals Applied Probability*, 15, pp. 2172–2202 (2005).
66. M. S. Johannes, N. G. Polson and J. R. Stroud. Optimal Filtering of Jump Diffusions: Extracting Latent States from Asset Prices. *Review of Financial Studies*, 22, Issue. 7, pp. 2759–2799 (2009).
67. I. Karatzas and S. Shreve. *Methods of Mathematical Finance*. Springer (1998).
68. A. Kohatsu-Higa and M. Montero. *Malliavin Calculus in Finance*. Handbook of Computational and Numerical Methods in Finance, Birkhauser, pp. 111–174 (2004).
69. J. M. Lasry and P. L. Lions. Contrôle stochastique avec informations partielles et applications à la Finance. Comptes Rendus de l'Académie des Sciences - Series I - Mathematics, vol. 328, issue 11, pp. 1003–1010 (1999).
70. A. E. B. Lim. Mean variance hedging when there are jumps. *SIAM J. Control Optim.*. vol. 44, no.5, pp. 1893–1922 (2005).
71. G. Liu and L. J. Hong. Revisit of stochastic mesh method for pricing American options. Operations Research Letters. 37(6), 411–414 (2009).
72. F. A. Longstaff and E. S. Schwartz. Valuing American options by simulation: A simple least-squares approach. *Review of Financial Studies* 14, pp. 113–147 (2001).
73. D. Lefèvre. An introduction to Utility Maximization with Partial Observation. *Finance*, vol. 23, (2002).
74. F. A. Matsen and J. Wakeley. Convergence to the Island-Model Coalescent Process in Populations With Restricted Migration. *Genetics*, vol. 172, pp. 701-708 (2006).
75. L. Mazliack. Approximation of a Partially Observable Stochastic Control Problem. *Markov Processes and Related Fields*, vol.5, pp. 477–486, (1999).
76. G. Pagès and B. Wilbertz. Optimal quantization methods for pricing American style options. *Numerical Methods in finance, Springer* (2011).
77. G. Pagès and H. Pham. Optimal quantization methods for nonlinear filtering with discrete-time observations. *Bernoulli*, vol.11, pp. 893–932 (2005).
78. G. Pagès, H. Pham and J. Printems. Optimal quantization methods and applications to numerical problems in finance. *in Handbook of Computational and Numerical Methods in Finance.* Ed. S. T. Rachev, *Birkhuser, Boston*, 2004, pp. 253–297 (2004).
79. G. Pagès. A space vector quantization method for numerical integration. *Journal of Comput. Appl. Math.*, 89, pp. 1–38 (1997).
80. G. Pagès and J. Printems. Functional quantization for numerics with an application to option pricing. *Monte Carlo Methods and Appl.*, 11(4), pp. 407–446 (2005).
81. N. Papageorgiou, Br. Rémillard and A. Hocquard . Replicating the properties of hedge fund returns. *Journal of Alternative Invesments*, vol. 11, pp. 8–38 (2008).
82. H. Pham, W. Runggaldier and A. Sellami. Approximation by quantization of the filter process and applications to optimal stopping problems under partial observation. *Monte Carlo Methods and Applications*. Volume 11, Issue 1, pp. 57–81 (2005).
83. H. Pham. Mean-variance hedging under partial observation. *International Journal of Theoretical and Applied Finance*, 4, pp. 263–284 (2001).
84. H. Pham and M. C. Quenez. Optimal portfolio in partially observed stochastic volatility models. *Annals of Applied Probability*, 11, pp. 210–238 (2001).
85. H. Pham, M. Corsi, and W. Runggaldier, Numerical Approximation by Quantization of Control Problems in Finance Under Partial Observations. *Handbook of Numerical Analysis*, vol. 15, pp. 325-360 (2009).
86. B. R. Rambharat and A. E. Brockwell. Sequential Monte Carlo pricing of American-style options under stochastic volatility models. *The Annals of Applied Statistics*, vol. 4, no. 1, pp. 222–265 (2010).
87. Br. Rémillard, A. Hocquard, H. Langlois and N. Papageorgiou. Optimal hedging of American options in discrete time. *Numerical Methods in finance, Springer* (2011).
88. Br. Rémillard and S. Rubenthaler. Optimal hedging in discrete and continuous time. Technical Report G-2009-77, Cahiers du Gerad (2009).

89. V. Rossi and J. P. Vila. Nonlinear filtering in discrete time : A particle convolution approach. Ann. I.SU.P., vol.50, no. 3, pp. 71–102 (2006).
90. W. Runggaldier and L. Stettner. Approximations of Discrete Time Partially Observed Control Problems Applied Mathematics Monographs CNR, Giardini Editori, Pisa (1994).
91. M. Schweizer. Variance-optimal hedging in discrete time. *Math. Oper. Res.*, vol. 20, no. 1, pp. 1–32 (1995).
92. J. N. Tsitsiklis and B. Van Roy. Regression Methods for Pricing Complex American-Style Options. *IEEE Transactions on Neural Networks*, Vol. 12, No. 4 (special issue on computational finance), pp. 694-703 (2001).
93. R. Van Handel. Uniform time average consistency of Monte Carlo particle filters, *Stoch. Proc. Appl.*, 119, pp. 3835–3861 (2009).
94. J. Vanneste. Estimating generalized Lyapunov exponents for products of random matrices. *Phys. Rev. E*, vol. 81, 036701 (2010).
95. F. Viens. Portfolio optimization under partially observed stochastic volatility. *Proceedings of the 8th international conference on advances in communication and control : telecommunications / signal processing*, pp. 3–12, (2001).
96. D. Whitley, S. Rana and R.B. Heckendorn. The island Model Genetic algorithm: On separability, population size and convergence. *CIT. Journal of computing and information technology.* vol. 7, no. 1, pp. 33-47 (1999).