UNCERTAIN VOLATILITY MODEL: A MONTE-CARLO APPROACH

JULIEN GUYON AND PIERRE HENRY-LABORDÈRE

Abstract. The uncertain volatility model has long ago attracted the attention of practitioners as it provides worst-case pricing scenario for the sell-side. The valuation of a financial derivative based on this model requires solving a fully non-linear PDE. One can rely on finite difference schemes only when the number of variables (that is, underlyings and path-dependent variables) is small - in practice no more than three. In all other cases, numerical valuation seems out of reach. In this paper, we outline two accurate, easy-to-implement Monte-Carlo-like methods which hardly depend on dimensionality. The first method requires a parameterization of the optimal covariance matrix and consists in a series of backward low-dimensional optimizations. The second method relies heavily on a recently established connection between second-order backward stochastic differential equations and non-linear second-order parabolic PDEs. Both methods are illustrated by numerical experiments.

1. Introduction

The uncertain volatility model (in short UVM), developed by Avellaneda *et al.* (1995), assumes that a risky asset follows a controlled diffusion under a risk-neutral measure

$$dX_t = \sigma_t X_t dW_t$$

The control (σ_t), representing the uncertain volatility process of the asset X, is valued in the interval $[\underline{\sigma}, \overline{\sigma}]$. It is also assumed to be adapted to the Brownian filtration, that is, it is not allowed to look into the future. The UVM selects one volatility at each time such that the value of the option under consideration is maximized. In this way, it provides a worst-case scenario for the sell-side. The valuation of an option can be written as the solution (in the viscosity sense) of an Hamilton-Jacobi-Bellman (HJB) equation with a control on the diffusion coefficient (Pham (2005)). This leads to a *fully non-linear* second-order PDE, the so-called Black-Scholes-Barenblatt PDE (in short BSB):

$$\partial_t u(t,x) + \frac{1}{2} x^2 \Sigma\left(\partial_x^2 u(t,x)\right)^2 \partial_x^2 u(t,x) = 0, \qquad (t,x) \in [0,T) \times \mathbb{R}_+^* \tag{1}$$

with some terminal condition $u(T,x) = g(x)$, and $\Sigma(\Gamma) = \underline{\sigma} 1_{\Gamma<0} + \overline{\sigma} 1_{\Gamma \geq 0}$ (Avellaneda *et al.* (1995)).

This PDE is called *fully* non-linear because the non-linearity affects the second order space derivative $\partial_x^2 u$, the so-called "gamma." In practice, this highly non-linear PDE cannot be solved analytically and we must resort to a finite-difference scheme. This method suffers from the curse

Received by the editors April 2009.

Key words and phrases. Uncertain volatility model, optimization of non-smooth multivariate function, backward stochastic differential equation, Monte-Carlo simulation, regression, Malliavin.

of dimensionality when the number of variables - assets or path-dependent variables[1] - is large.[2] A probabilistic representation of the BSB equation leading to a Monte-Carlo algorithm would solve this problem. Unfortunately, the classical link between Monte-Carlo and finite-difference methods as stated by the Feynman-Kac formula only holds for *linear* second-order parabolic PDEs.

In this paper, we suggest two Monte Carlo approaches for pricing options in the UVM that are almost immune to dimensionality. Both approaches consist in (a) estimating the optimal covariance matrix and, (b) computing a lower bound of the price using this (necessarily suboptimal) estimate. By the BSB PDE, the optimal covariance matrix at date t is a function of the gammas of the product at that time. Note, however, that the gammas are often very poorly estimated by the naive method consisting in doubly differentiating a continuation value proxy with respect to the asset values. For instance, the Longstaff-Schwartz algorithm (Longstaff and Schwartz (2001)) is designed to produce good estimates of the value of the product in the future, not to build good proxys for the gammas in the future.

The first approach requires that the user specify a parameterization of the optimal covariance matrix and consists in a series of backward optimizations on the underlying parameters. It is instructive to compare such an approach with the Monte Carlo valuation of American options suggested by Longstaff and Schwartz: the parameterization of the optimal covariance matrix is analogous to the choice of regressors; the backward optimizations are similar to the successive least-squares procedures. Note that the parameterization step, as well as the choice of regressors in the Longstaff-Schwartz method, requires human intelligence, namely an accurate analysis of the payoff. Nevertheless, we will explain how to build tools helping us to choose good parameterizations.

The second method makes use of *backward* stochastic differential equations (in short BSDEs). Recently, first-order BSDEs (1-BSDEs) have attracted attention in the mathematical finance community, see El Karoui *et al.* (1997) for references and a review of applications. One of the key features of 1-BSDEs is that they provide a probabilistic representation of solutions of non-linear parabolic PDEs, generalizing the Feynman-Kac formula. However, the corresponding PDEs are not allowed to be non-linear in the second-order derivative and are therefore connected to HJB equations with no control on the diffusion term. In Cheridito *et al.* (2007), the authors provide a stochastic representation for solutions of fully non-linear parabolic PDEs by introducing a new class of BSDEs, the so-called second-order BSDEs (in short 2-BSDEs). In this paper, we will apply this tool to obtain our second Monte-Carlo implementation of the UVM. Note that in a recent preprint (Fahim *et al.* (2009)), this tool has been used for the numerical computation of the solution of the mean curvature flow in dimensions two and three and for two and five-dimensional HJB equations arising in the theory of portfolio optimization.

Both methods cannot be said to be strictly impervious to dimensionality, because as the dimension grows, it becomes harder to guess an efficient parameterization in the first method, and guessing an efficient regression basis becomes harder as well in the second method. This is exactly the same problem one faces when using the Longstaff-Schwartz algorithm to price American options. However, it often happens in finance that despite the high number of variables, the price and gamma of an option (like the optimal exercise strategy of an American option) depend essentially on only a few variables, like the smallest of d assets, or the smallest and second smallest of d assets.

[1]In this paper, we consider path-dependent variables whose values can change only at discrete dates. For instance, we exclude continuously computed realized variance, but we allow realized variance computed on a monthly basis.

[2]However, see von Petersdorff *et al.* (2004) for sparse matrix methods.

As a consequence, our two methods hardly depend on the dimensionality and can be very efficient even in very high dimension, exactly like the Longstaff-Schwartz method.

As far as the authors know, it is the first time that Monte-Carlo implementations of the UVM, which are unavoidable in high dimension, are suggested in the literature. Such methods dramatically enlarge the range of options that can be priced - and hedged - under uncertain volatility. After the seminal works by Avellaneda *et al.* (1995) and Lyons (1995), Avellaneda and Buff (1999) and Buff (2002) have considered the pricing in the UVM of a basket of options, including barrier and American options, written on a single asset. Still in the single-asset case: Smith (2002) has also studied the pricing of American options; Martini (1999) has established a link between the UVM and American options; Leblanc and Martini (2000) have focused on the case where $\overline{\sigma} = +\infty$; Forsyth *et al.* (2001) have developed a fully implicit PDE method to price discretely observed barrier options; and Pooley *et al.* (2003) have looked at the convergence properties of some numerical schemes. Pooley *et al.* (2003) also looked at some numerical details peculiar to the two-asset case. Meyer (2006) has applied the BSB PDE to static hedging. Zhang and Wang (2009) have recently developed a fitted finite volume method to numerically solve the BSB PDE.

The paper is organized as follows. In Section 2, we recall the main results about the UVM and we set our notations. Section 3 describes the parametric method, first in the general multi-asset case, then with a special focus on the single-asset and two-asset cases. As 1-BSDEs seem to be quite unheard of in the practitioner community, we present in Section 4 a short accessible introduction stating the main definitions and results, and probabilistic representations of non-linear PDEs. This leads us to 2-BSDEs as introduced in Cheridito *et al.* (2007). In Section 5, we use 2-BSDEs to build our second Monte-Carlo approach. Eventually, our two Monte-Carlo techniques are illustrated by numerical examples in the last section.

2. Uncertain Volatility Model

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space equipped with a d-dimensional Brownian motion W. We denote $(\mathcal{F}_t)$ the natural filtration of W. The market is assumed to be made of d assets, the values of which are represented by a d-dimensional positive local Itô $(\mathcal{F}_t, \mathbb{P})$-martingale X. Let us consider an option delivering some payoff H_T at maturity T. H_T is some function of the assets path $(X_t, 0 \leq t \leq T)$. The time-t value V_t of the option in the UVM is the solution of a maximization problem:

$$(2) \qquad V_t = \sup_{[t,T]} \mathbb{E}[H_T|\mathcal{F}_t], \quad dX_t^\alpha = \sigma_t^\alpha X_t^\alpha dW_t^\alpha, \quad dW_t^\alpha dW_t^\beta = \rho_t^{\alpha\beta} dt, \quad 1 \leq \alpha < \beta \leq d$$

where $\sup_{[t,T]}$ means that the supremum is taken over all $(\mathcal{F}_s)$-adapted processes $(\xi_s)_{t \leq s \leq T} \equiv ((\sigma_s^\alpha, \rho_s^{\alpha\beta})_{1 \leq \alpha < \beta \leq d})_{t \leq s \leq T}$ such that for all $s \in [t, T]$, ξ_s belongs to some compact domain D. The domain D must be such that for all $\xi = (\sigma^\alpha, \rho^{\alpha\beta})_{1 \leq \alpha < \beta \leq d} \in D$, the covariance matrix $(\rho^{\alpha\beta}\sigma^\alpha\sigma^\beta)_{1 \leq \alpha, \beta \leq d}$, with $\rho^{\beta\alpha} = \rho^{\alpha\beta}$ and $\rho^{\alpha\alpha} = 1$, is non-negative. We will consider domains D of the form $D = [\underline{\sigma}, \overline{\sigma}]$ when $d = 1$, and $D = [\underline{\sigma}^1, \overline{\sigma}^1] \times [\underline{\sigma}^2, \overline{\sigma}^2] \times [\underline{\rho}, \overline{\rho}]$ when $d = 2$.

Pricing vanilla options. Let us first consider vanilla payoffs $H_T = g(X_T)$. The payoff function g is assumed continuous with quadratic growth. From the HJB principle, we obtain

Theorem 2.1 (Black-Scholes-Barenblatt PDE, Pham (2005)). *We have $V_t = u(t, X_t)$ where $u(\cdot,\cdot)$ is the unique (viscosity) solution with quadratic growth of the non-linear PDE*

$$(3) \qquad \partial_t u(t,x) + f(x, \nabla_x^2 u(t,x)) = 0, \qquad (t,x) \in [0,T) \times (\mathbb{R}_+^*)^d$$

with the terminal condition $u(T,x) = g(x)$ and the Hamiltonian

$$f(X,\Gamma) = \frac{1}{2} \max_{(\sigma^\alpha,\rho^{\alpha\beta})_{1\leq\alpha<\beta\leq d}\in D} \sum_{\alpha,\beta=1}^{d} \rho^{\alpha\beta}\sigma^\alpha\sigma^\beta X^\alpha X^\beta \Gamma^{\alpha\beta} \tag{4}$$

Note that if $g \in \mathcal{C}^3((\mathbb{R}_+^*)^d)$ and $g, \nabla_x g$ have quadratic growth, then $u \in \mathcal{C}^{1,2}([0,T]\times(\mathbb{R}_+^*)^d)$.

Remark 2.2 ($d=1$). For $d=1$, the above PDE reduces to (1). If we consider a convex payoff, u coincides with the Black-Scholes price with the upper volatility $\overline{\sigma}$.

Remark 2.3 (Transaction cost). It is interesting to note the similarity of the Black-Scholes-Barenblatt PDE with the pricing equation resulting from the Leland (1985) transaction cost model sketched here in one dimension (this analogy between the Black-Scholes-Barenblatt PDE and Leland's model has also been noticed in Forsyth (2001)):

$$\Sigma(\Gamma)^2 = \sigma^2 + 2\sqrt{\frac{2}{\pi}}\frac{k\sigma}{\sqrt{\delta t}}\mathrm{sign}(\Gamma)$$

If Γ shares are bought or sold at a price S, the transaction cost is $k|\Gamma|S$ where k is a constant depending on the individual investor; δt is the re-hedging frequency. In the general case with d assets, the Hamiltonian reads (Avellaneda and Kampen (2003))

$$f(X,\Gamma) = \frac{1}{2}\sum_{\alpha,\beta=1}^{d} \rho^{\alpha\beta}\sigma^\alpha\sigma^\beta X^\alpha X^\beta\Gamma^{\alpha\beta} + \sqrt{\frac{2}{\pi\delta t}}\sum_{\alpha=1}^{d} k_\alpha X^\alpha \sqrt{\sum_{\beta,\gamma=1}^{d}\rho^{\beta\gamma}\sigma^\beta\sigma^\gamma X^\beta X^\gamma \Gamma^{\alpha\beta}\Gamma^{\alpha\gamma}}$$

Pricing with path-dependent variables. When the price of an option depends on path-dependent variables whose values can change only at discrete dates, one solves PDE (3) between two such dates t_l and t_{l-1} for fixed values of the path-dependent variables A, and defines

$$u(t_{l-1}^-, X, A) = u(t_{l-1}^+, X, \varphi(A, X))$$

with the function φ linking the past and new values of the path-dependent variables. For instance, if the option value depends on a monthly-computed realized variance,

$$A_t^1 = \sum_{\{l|t_l\leq t\}}\left(\ln\frac{X_{t_l}}{X_{t_{l-1}}}\right)^2, \qquad A_t^2 = X_{\sup_{\{l|t_l\leq t\}} t_l}, \qquad \text{and} \qquad \varphi(X,A) = A^1 + \left(\ln\frac{X}{A^2}\right)^2$$

Remark 2.4 (On path-dependent variables whose values can change continuously). It is for the sake of simplicity that, in this paper, we consider path-dependent variables whose values can change only at discrete dates. What happens in the case where the price of an option depends on path-dependent variables whose values can change continuously, is that the Hamiltonian f may not involve only the gammas. For instance, in the single-asset case, if the price $u(t,x,v)$ of an option depends on the continuously compounded realized variance v, the Hamiltonian reads

$$f(x, \partial_x^2 u, \partial_v u) = \max_{\underline{\sigma}\leq\sigma\leq\overline{\sigma}} \sigma^2\left(\frac{1}{2}x^2\partial_x^2 u + \partial_v u\right),$$

i.e., the optimal volatility is either $\underline{\sigma}$ or $\overline{\sigma}$, depending on the sign, not of the gamma $\partial_x^2 u$, but of $\frac{1}{2}x^2\partial_x^2 u + \partial_v u$. Our Monte-Carlo approaches are easily adapted to such a case, see Remarks 3.3 and 6.1.

Computation of the Hamiltonian. Equation (4) is a constrained programming problem. In the particular case where $d = 2$, which we will consider in our numerical experiments, the Hamiltonian reads

$$f(X,\Gamma) = \max_{(\sigma^1,\sigma^2,\rho)\in D} \left(\frac{1}{2}(\sigma^1)^2(X^1)^2\Gamma^{11} + \frac{1}{2}(\sigma^2)^2(X^2)^2\Gamma^{22} + \rho\sigma^1\sigma^2X^1X^2\Gamma^{12}\right) \tag{5}$$

and has a closed-form solution when $D = [\underline{\sigma}^1, \overline{\sigma}^1] \times [\underline{\sigma}^2, \overline{\sigma}^2] \times [\underline{\rho}, \overline{\rho}]$. Indeed, the optimal correlation is bang-bang: it is either $\underline{\rho}$ or $\overline{\rho}$, depending on the sign of the cross gamma Γ^{12}: $R(\Gamma^{12}) = \underline{\rho}1_{\Gamma^{12}<0} + \overline{\rho}1_{\Gamma^{12}\geq 0}$. Then, the problem

$$\max_{\sigma^1,\sigma^2} \phi(\sigma^1,\sigma^2), \qquad \phi(\sigma^1,\sigma^2) = \left(\frac{1}{2}(\sigma^1)^2(X^1)^2\Gamma^{11} + \frac{1}{2}(\sigma^2)^2(X^2)^2\Gamma^{22} + R(\Gamma^{12})\sigma^1\sigma^2X^1X^2\Gamma^{12}\right)$$

is a two-dimensional quadratic form maximization under double inequality constraint, which has an explicit solution because one can first freeze σ^2 and maximize in the σ^1 variable to get

$$\sigma^{*1}(\sigma^2) = 1_{\Gamma^{11}<0}\left(\left(\left(-\frac{R(\Gamma^{12})\sigma^2X^2\Gamma^{12}}{X^1\Gamma^{11}}\right)\vee\underline{\sigma}^1\right)\wedge\overline{\sigma}^1\right) + 1_{\Gamma^{11}\geq 0}\begin{cases}\underline{\sigma}^1 & \text{if } \phi(\underline{\sigma}^1,\sigma^2) > \phi(\overline{\sigma}^1,\sigma^2)\\ \overline{\sigma}^1 & \text{otherwise}\end{cases}$$

and then compute the maximum μ of $\psi(\sigma^2) = \phi(\sigma^{*1}(\sigma^2),\sigma^2)$. Now, $\mu = 1_{\Gamma^{11}\geq 0}(\mu_1 \vee \mu_2) + 1_{\Gamma^{11}<0}(\mu_1 \vee \mu_2 \vee \mu_3)$ where

$$\begin{aligned}
\mu_1 &= \max_{\underline{\sigma}^2\leq\sigma^2\leq\overline{\sigma}^2} \phi(\underline{\sigma}^1,\sigma^2)\\
\mu_2 &= \max_{\underline{\sigma}^2\leq\sigma^2\leq\overline{\sigma}^2} \phi(\overline{\sigma}^1,\sigma^2)\\
\mu_3 &= \max_{\underline{\sigma}^2\leq\sigma^2\leq\overline{\sigma}^2,\quad \underline{\sigma}^1\leq -\frac{R(\Gamma^{12})\sigma^2X^2\Gamma^{12}}{X^1\Gamma^{11}}\leq\overline{\sigma}^1} \phi\left(-\frac{R(\Gamma^{12})\sigma^2X^2\Gamma^{12}}{X^1\Gamma^{11}},\sigma^2\right)
\end{aligned}$$

and each μ_i is itself the maximum of a polynomial of degree 2 under a double inequality constraint, the solution of which has a closed form. Note that, contrary to the single-asset case, the optimal volatility $(\Sigma^1(X,\Gamma),\Sigma^2(X,\Gamma))$ may take values out of the boundary. However, if $\underline{\rho} = \overline{\rho} = 0$, it necessarily hits the boundary.

In the general case with d assets, we can first maximize in ρ_{ij} as above. Equation (4) then becomes a constrained quadratic programming problem that can be solved efficiently using Lemke's algorithm, see Henry-Labordère (1995).

In practice, PDE (3) is not solvable and we must rely on a finite-difference scheme. But standard finite-difference schemes can only be implemented when the number of variables - assets or path-dependent variables - is no more than three[3]. To get rid of this curse of dimensionality, we suggest two Monte Carlo approaches. In the next section, we first focus on the parametric one.

3. The parametric approach

3.1. **The ideas behind the algorithm.** The two main ingredients in the parametric approach are the following.

First, we cut the optimization problem (2) into a series of smaller and simpler ones. To this end, let us discretize the time interval $[0,T]$, say $[0,T] = \cup_{i=1}^n [t_{i-1}, t_i]$. Assume that the price of the option at time t in the UVM is a function u of t, X_t and some path-dependent variables that

[3]However, see von Petersdorff *et al.* (2004) for sparse matrix methods.

we sum up under the notation A_t. For instance, a component of the vector A_t may represent realized variance, running maximum, average value, or some past value of an asset. The dynamic programming principle then reads

$$u(t_i, X_{t_i}, A_{t_i}) = \sup_{(\xi_t, t\in[t_i,t_{i+1}])} \mathbb{E}[u(t_{i+1}, X_{t_{i+1}}, A_{t_{i+1}})|\mathcal{F}_{t_i}],$$
$$dX_t^\alpha = \sigma_t^\alpha X_t^\alpha dW_t^\alpha, \qquad dW_t^\alpha dW_t^\beta = \rho_t^{\alpha\beta} dt, \qquad 1 \le \alpha < \beta \le d$$

with $u(T, X_T, A_T) = H_T$, the option payoff, and where $\xi_t \equiv (\sigma_t^\alpha, \rho_t^{\alpha\beta})_{1\le\alpha<\beta\le d}$. As a consequence, we can proceed backwards and split the original maximization problem (2), written for $t = 0$, into n consecutive smaller ones:

$$V_{t_i} = \sup_{(\xi_t, t\in[t_i,t_{i+1}])} \mathbb{E}[H_T|\mathcal{F}_{t_i}], \qquad dX_t^\alpha = \sigma_t^\alpha X_t^\alpha dW_t^\alpha, \qquad dW_t^\alpha dW_t^\beta = \rho_t^{\alpha\beta} dt, \qquad 1 \le \alpha < \beta \le d$$

where, for $t \in [t_{i+1}, T]$, $\xi_t \equiv (\sigma_t^\alpha, \rho_t^{\alpha\beta})_{1\le\alpha<\beta\le d}$ are the volatilities and correlations that are the solution of the previously solved maximization problem. We are thus left with n small consecutive optimization problems to solve, namely $(P_{n-1}), (P_{n-2}), \dots, (P_0)$, where:

> (P_i): find a D-valued deterministic function a_{t_i} of X and A so as to maximize $\mathbb{E}[H_T|\mathcal{F}_{t_i}]$, with $\xi_t = a_{t_i}(X_t, A_t)$ for $t \in [t_i, t_{i+1})$, and with the already optimized $\xi_t = a^*_{t_j}(X_t, A_t)$ for $t \in [t_j, t_{j+1})$, $j \ge i + 1$.

The second ingredient concerns the resolution of problem (P_i). In this problem, the maximization set is huge: each point in this set is a deterministic function a_{t_i} of (X, A) taking values in D. In particular, a_{t_i} lives in an infinite-dimensional space. Even if we discretize the domain D (say, q possible values) and the domain of possible values for (X, A) (say, p possible values), enumerating the maximization set takes too much time. This is essentially because the set of functions of $\{1, \dots, p\}$ onto $\{1, \dots, q\}$ has cardinal q^p, a quantity which grows exponentially with p.

To overcome this issue, we *decide* to restrict the maximization domain to a parametized set of *relevant* functions we want to test in the optimization procedure. To be precise, we will only test a_{t_i}'s of the form $\lambda_{t_i}(\cdot; \theta)$, for θ in some parameter set $\Theta \subset \mathbb{R}^l$. The $\cdot$ stands for (X, A). A typical example is $\lambda_{t_i}(X, A; \theta) = \xi_1$ if $\gamma_{t_i}(X, A; \theta) \ge 0$, $= \xi_2$ otherwise, for some function γ_{t_i}, where ξ_1 and ξ_2 are two given points in the domain D. At each date t_i, we are thus left with a maximization over the low-dimensional parameter θ, a problem that can be solved much more quickly than the initial optimization over all possible functions a_{t_i}'s. A similar approach was used in Andersen (1999) to compute the fair value of Bermudan swaptions.

3.2. **The algorithm.** We can now build a Monte-Carlo algorithm as follows:

- Simulate N_1 replicas of X with some diffusion, say, for instance, the log-normal diffusion with some volatilities and correlations $\hat{\xi} = (\hat{\sigma}^\alpha, \hat{\rho}^{\alpha\beta})_{1\le\alpha<\beta\le d}$.
- For $i = n-1, n-2, \dots, 1, 0$, find a numerical solution θ_i^* of the maximization problem (Q_i):

$$\sup_{\theta_i\in\Theta} h(\theta_i), \quad h(\theta_i) = \frac{1}{N_1}\sum_{p=1}^{N_1} H_T^{(p)}, \quad dX_t^\alpha = \sigma_t^\alpha X_t^\alpha dW_t^\alpha, \quad dW_t^\alpha dW_t^\beta = \rho_t^{\alpha\beta} dt \tag{6}$$

where

$$\xi_t \equiv (\sigma_t^\alpha, \rho_t^{\alpha\beta})_{1\le\alpha<\beta\le d} = \begin{cases} \hat{\xi} & \text{if } t \in [0, t_i), \\ \lambda_{t_i}(X_t, A_t; \theta_i) & \text{if } t \in [t_i, t_{i+1}), \\ \lambda_{t_j}(X_t, A_t; \theta_j^*) & \text{if } t \in [t_j, t_{j+1}), j \ge i+1 \end{cases} \tag{7}$$

- Independently, simulate N_2 replicas of X using $\xi_t = \lambda_{t_i}(X_t, A_t; \theta_i^*)$ for $t \in [t_i, t_{i+1})$ and compute $\frac{1}{N_2}\sum_{p=1}^{N_2} H_T^{(p)}$.

Let us make a few comments on each of the three steps of this algorithm. In Step 1, we must pick some numerical volatilities and correlations $\hat{\xi} = (\hat{\sigma}^\alpha, \hat{\rho}^{\alpha\beta})_{1\le\alpha<\beta\le d}$. Subsection 5.4 deals with this issue.

In Step 2, solving (Q_i) requires that we compute $h(\theta_i)$ for many values of θ_i. Each $H_T^{(p)}$ depends on θ_i, because it is a function of path p whose volatility at date $t \in [t_i, t_{i+1})$ depends on θ_i. As a consequence, to compute each $h(\theta_i)$, we must resimulate the N_1 paths from t_i to T. No Brownian increments are drawn at Step 2: the resimulations only consist in multiplying the Brownian increments drawn at Step 1 by new volatilities.

In our view, Step 3 is crucial. We could have estimated the price by $h(\theta_0^*)$, the value function resulting from the last maximization problem (Q_0) in Step 2. Instead, in Step 3, we decide to run new independent paths which use the nearly optimal (but necessarily suboptimal) parameters θ_i^* resulting from Step 2. This ensures that the true price of the option in the UVM is *larger* than the output of our algorithm. For instance, we can run the algorithm for several parameterizations, numbers of paths, time steps, etc., and keep only the maximum of all these results as our estimator, since all the results are lower bounds of the true price. Ideally, this low-biased estimator should be complemented by a high-biased estimator, but we leave this point for future work. The bias of the estimator $h(\theta_0^*)$ is unknown, because θ_i^* is not $\mathcal{F}_{t_i}$-measurable, as it incorporates values of the N_1 Brownian paths after t_i. A similar problem occurs in the pricing of American options: the exercise strategy generated by the Longstaff-Schwartz algorithm is not adapted to the Brownian filtration, and one should simulate new independent paths to get a low-biased estimate of the price; this lower bound is complemented by an upper bound *à la* Rogers (Rogers (2002)). We find that it is uncomfortable to deal with estimators with unknown bias, whereas a low-biased estimator gives a certain, robust, though partial, information.

Remark 3.1. Solving (Q_i) requires choosing an initial value θ_i^0. For $i \le n-2$, we choose $\theta_i^0 = \theta_{i+1}^*$.

Remark 3.2. An elegant way to avoid resimulations consists in multiplying H_T by the relevant likelihood ratio. Indeed, if we denote $\mathbb{P}^{\theta_i}$ the probability under which (7) holds and $p_{\hat{\xi}}(s,t,x,y)$ the log-normal density, we have

$$\mathbb{E}^{\theta_i}[H_T] = \mathbb{E}^{\hat{\xi}}\left[H_T \prod_{t_i \le t_k < T} \frac{p_{\xi_{t_k}}}{p_{\hat{\xi}}}(t_k, t_{k+1}, X_{t_k}, X_{t_{k+1}})\right]$$

where the t_k's are the discretization times in the simulation of process X, and include the t_i's. A similar technique is used in Broadie and Glasserman (1997) to price American options in high dimension. Unfortunately, the likelihood ratio has great variance, unless D is 'small', so that empirical averages are poor estimates of the right hand side expectation above.

Remark 3.3. In Remark 2.4, we pointed out that in the case where the price of an option depends on path-dependent variables whose values can change continuously, the Hamiltonian f may depend not only on the gammas, and thus may differ from (4). In the parametric approach, one directly parameterizes the optimal covariance matrix, regardless of the form of the Hamiltonian. As a consequence, this method works whether the path-dependent variables change continuously or only at discrete dates.

3.3. **Choice of the parameterization.** The parameterization of the maximization set, i.e., the choice of relevant functions λ_{t_i}'s, is a crucial step in our procedure. A wrong choice would lead to a rough estimate of the optimal volatilities and correlations, hence to a rough lower bound price. To build a good parameterization, one can proceed as follows:

- Choose some relevant path-dependent variables $A = (A^1, \ldots, A^q)$.
- For a grid of dates, asset values and path-dependent values (t, X, A), compute Monte-Carlo gammas $\Gamma(t, X, A)$ in the Black-Scholes model with some covariance matrix $\hat{\xi}$.
- For each point (t, X, A) in the grid, build the solution $(\sigma^{*\alpha}(t, X, A), \rho^{*\alpha\beta}(t, X, A))_{1 \leq \alpha < \beta \leq d}$ to the problem

$$f(X, \Gamma(t, X, A)) = \frac{1}{2} \max_{(\sigma^\alpha, \rho^{\alpha\beta})_{1 \leq \alpha < \beta \leq d}} \sum_{\alpha,\beta=1}^{d} \rho^{\alpha\beta} \sigma^\alpha \sigma^\beta X^\alpha X^\beta \Gamma^{\alpha\beta}(t, X, A)$$

- For each date t in the time grid, graph the optimal solutions $(X, A) \mapsto \sigma^{*\alpha}(t, X, A)$ and $(X, A) \mapsto \rho^{*\alpha\beta}(t, X, A)$ and guess a parameterization for them.

Because this preliminary process involves many computations of Monte-Carlo prices, it is very time-consuming *but* it is done once for all for each product. An alternative tool using our second Monte-Carlo approach is sketched in Remark 5.2.

3.4. **The single-asset case.** To illustrate the general algorithm stated above, let us focus on the single-asset case. In practice, a simulation-based valuation method is needed even in this case, when the value of an option depends on three or more path-dependent variables. We know that the price of an option in the UVM is the solution of the Black-Scholes-Barenblatt PDE, and that the solution of this PDE is of bang-bang type: at each date, the optimal volatility is either $\underline{\sigma}$ or $\overline{\sigma}$. This means that we can restrict the maximization set to functions λ_{t_i} taking values in $\{\underline{\sigma}, \overline{\sigma}\}$. Stated otherwise, it is enough to specify functions $\gamma_{t_i}(\cdot; \theta)$ such that $\lambda_{t_i}(X, A; \theta_i) = \overline{\sigma}$ if $\gamma_{t_i}(X, A; \theta_i) \geq 0$, $= \underline{\sigma}$ otherwise. The sign of $\gamma_{t_i}(\cdot; \theta)$ hence represents the sign of the gamma at date $t \in [t_i, t_{i+1})$.

This particular form for λ_{t_i} speeds up evaluations of $h(\theta_i)$ in Step 2, see (6), because only some of the N_1 paths must be resimulated. Imagine we resimulated paths and computed the value $h(\theta_i^m)$ for some value θ_i^m. The next step in our solver for (Q_i) requires that we do the same for some next value θ_i^{m+1}. It often happens that for many simulated paths at steps m and $m + 1$, we have

$$\forall t_k \in [t_i, t_{i+1}), \quad \lambda_{t_i}(X_{t_k}^{m+1}, A_{t_k}^{m+1}; \theta_i^{m+1}) = \lambda_{t_i}(X_{t_k}^m, A_{t_k}^m; \theta_i^m)$$

so that these paths need not be resimulated. This happens when the two consecutive values $\gamma_{t_i}(X_{t_k}^{m+1}, A_{t_k}^{m+1}; \theta_i^{m+1})$ and $\gamma_{t_i}(X_{t_k}^m, A_{t_k}^m; \theta_i^m)$ have same sign for all the discretization dates t_k in the (small) interval $[t_i, t_{i+1})$. For instance, assume that γ_{t_i} is continuous in θ_i and that the sequence $(\theta_i^m)_{m \in \mathbb{N}}$ converges, as one expects, to some maximum θ_i^*. Then the sign of $\gamma_{t_i}(X_{t_k}^m, A_{t_k}^m; \theta_i^m)$ will stay constant after some iterations on m (provided that the limit $\gamma_{t_i}(X_{t_k}^*, A_{t_k}^*; \theta_i^*)$ be non-zero for all $t_k \in [t_i, t_{i+1})$).

If γ_{t_i} is continuous in θ_i, then h is piecewise constant as a function of θ_i. As a consequence, optimization routines based on the computation of gradients and hessians are of no use. We use the so-called downhill simplex method, see Flannery *et al.* (2007).

3.5. **The two-asset case.** In the two-asset case with $D = [\underline{\sigma}^1, \overline{\sigma}^1] \times [\underline{\sigma}^2, \overline{\sigma}^2] \times [\underline{\rho}, \overline{\rho}]$, our parametric approach requires that we provide:

- a function $\kappa_{t_i}(\cdot;\theta_i)$ such that $\rho_{t_i}(X, A; \theta_i) = \overline{\rho}$ if $\kappa_{t_i}(X, A; \theta_i) \geq 0$, $= \underline{\rho}$ otherwise, i.e., κ_{t_i} represents the sign of the cross gamma,
- a function $\sigma^1_{t_i}(\cdot;\theta_i)$ taking values in $[\underline{\sigma}^1, \overline{\sigma}^1]$,
- a function $\sigma^2_{t_i}(\cdot;\theta_i)$ taking values in $[\underline{\sigma}^2, \overline{\sigma}^2]$.

Although we know that the solution is not of bang-bang type, we can pick two functions $\sigma^1_{t_i}(\cdot;\theta_i)$ and $\sigma^2_{t_i}(\cdot;\theta_i)$ taking values resp. in $\{\underline{\sigma}^1, \overline{\sigma}^1\}$ and $\{\underline{\sigma}^2, \overline{\sigma}^2\}$. This speeds up the solution of (Q_i) as explained in the single-asset case subsection.

Numerical examples with one or two assets will be given in Section 6. We now present our second approach for valuing options by simulation in the UVM. This second approach uses a link between nonlinear PDEs and BSDEs that we briefly introduce in the next section (details can be found in Ma and Yong (2000) and Pham (2009)).

4. A brief introduction to backward stochastic differential equations

We consider a d-dimensional Itô process defined by the (forward) SDE

$$dX_t = \sigma(t, X_t)dW_t$$

with the initial condition $X_0 = x \in \mathbb{R}^d$. Under classical Lipschitz and growth conditions on the volatility σ, this SDE admits a unique strong solution. First-order BSDEs (in short 1-BSDEs) differ from (forward) SDEs in that we impose the terminal value.

4.1. First-order BSDEs.

4.1.1. *Definition, existence and uniqueness.*

Definition 4.1 (1-BSDE, Ma and Yong (2000))**.** A solution of a (Markovian) 1-BSDE is a couple $(Y, Z) \in \mathbb{R} \times \mathbb{R}^d$ of $(\mathcal{F}_t)$-*adapted* Itô processes satisfying

$$dY_t = -f(t, X_t, Y_t, Z_t)dt + Z_t.dW_t$$

with the terminal condition $Y_T = g(X_T)$. The above equation means that

$$Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s.dW_s \tag{8}$$

The deterministic function f is called the *driver*. We emphasize that the diffusion term is part of the solution.

Existence and uniqueness of solutions of 1-BSDEs are stated in the following theorem:

Theorem 4.2 (Pardoux and Peng (1990))**.** *If $f : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ and $g : \mathbb{R}^d \to \mathbb{R}$ are two deterministic functions satisfying the Lipschitz condition*

$$|g(u) - g(v)| + |f(x) - f(y)| \leq K\left(|u - v| + |x - y|\right)$$

for some constant K independent of $u, v \in \mathbb{R}^d$ and $x, y \in [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$, then there is a unique adapted solution (Y, Z) satisfying

$$\sup_{0 \le t \le T} \mathbb{E}\left[Y_t^2\right] + \mathbb{E}\left[\int_0^T |Z_t|^2 dt\right] < \infty$$

4.1.2. *BSDEs provide an extension of Feynman-Kac's formula for semi-linear PDEs.* Let us consider the semi-linear PDE

$$\partial_t u + \mathcal{L}u + f(t, x, u(t,x), \sigma(t,x)\nabla_x u(t,x)) = 0, \qquad (t,x) \in [0,T) \times \mathbb{R}^d \tag{9}$$

with the terminal condition $u(T,x) = g(x)$ and $\mathcal{L}$ the Itô generator of X. Such an equation appears when we consider a stochastic control problem with no control on the diffusion coefficient. It turns out that $(Y_t = u(t, X_t), Z_t = \sigma(t, X_t)\nabla_x u(t, X_t))$ is the unique solution of a BSDE. To be precise, a straightforward application of Itô's lemma gives

Proposition 4.3 (Generalization of Feynman-Kac's formula)**.** *Let u be a function of $\mathcal{C}^{1,2}$ satisfying (9) and suppose that there exists a constant C such that, for each $(t,x) \in [0,T] \times \mathbb{R}^d$,*

$$|u(t,x)| + |\sigma(t,x)\nabla_x u(t,x)| \le C(1 + |x|)$$

Then $(Y_t = u(t, X_t), Z_t = \sigma(t, X_t)\nabla_x u(t, X_t))$ is the unique solution to 1-BSDE (8).

Conversely, under additional regularity conditions, Pardoux and Peng (1992) show that the solution of BSDE (8) is a viscosity solution of (9). We state the precise theorem in dimension one:

Theorem 4.4 (Pardoux and Peng (1992))**.** *We suppose $d = 1$ and that f and g are uniformly continuous with respect to x. Then the function $u(t,x) \equiv Y_t^x$ is a viscosity solution to PDE (9). Furthermore, if we suppose that for each $R > 0$ there exists a continuous function $m_R : \mathbb{R}^+ \to \mathbb{R}$ such $m_R(0) = 0$ and*

$$|f(t,x,y,z) - f(t,x',y,z)| \le m_R(|x - x'|(1 + |z|))$$

for all $t \in [0,T]$ and $|x|, |x'|, |z| \le R$, then u is the unique *viscosity solution of PDE (9).*

Theorem 4.4 leads to a Monte-Carlo-like numerical solution to PDE (9) via BSDE (8). This requires an efficient discretization scheme for BSDEs. Two approaches have been considered in the literature. The first one uses an approximation of the driving Brownian motion by a random walk on a finite tree (Ma *et al.* (2002)). This is efficient only in low dimension. The second approach, that we explicit below, consists in writing an Euler-like scheme corrected by taking conditional expectations. We present below a scheme introduced in Bouchard and Touzi (2004) that we will extend to 2-BSDEs later.

4.1.3. *Numerical simulation.* Let us set $\Delta t_i = t_i - t_{i-1}$, $\Delta W_{t_i} = W_{t_i} - W_{t_{i-1}}$ and $\Delta = \max_i \Delta t_i$. We have

$$Y_{t_i} - Y_{t_{i-1}} = -\int_{t_{i-1}}^{t_i} f(s, X_s, Y_s, Z_s) ds + \int_{t_{i-1}}^{t_i} Z_s . dW_s$$

We consider the naive Euler-scheme

$$Y_{t_i}^{\Delta} - Y_{t_{i-1}}^{\Delta} = -f(t_{i-1}, X_{t_{i-1}}^{\Delta}, Y_{t_{i-1}}^{\Delta}, Z_{t_{i-1}}^{\Delta})\Delta t_i + Z_{t_{i-1}}^{\Delta} . \Delta W_{t_i} \tag{10}$$

together with the terminal condition $Y^{\Delta}_{t_n} = g(X^{\Delta}_{t_n})$. Of course, given $Y^{\Delta}_{t_i}, Z^{\Delta}_{t_i}$, there are no $\mathcal{F}_{t_{i-1}}$-measurable random variables $Y^{\Delta}_{t_{i-1}}, Z^{\Delta}_{t_{i-1}}$ satisfying the equation above. Therefore we consider the (*implicit*) scheme obtained by taking conditional expectations $\mathbb{E}_{i-1} = \mathbb{E}[\cdot|\mathcal{F}_{t_{i-1}}]$:

$$\begin{aligned} Y^{\Delta}_{t_n} &= g(X^{\Delta}_{t_n}) \\ Y^{\Delta}_{t_{i-1}} &= \mathbb{E}_{i-1}[Y^{\Delta}_{t_i}] + f(t_{i-1}, X^{\Delta}_{t_{i-1}}, Y^{\Delta}_{t_{i-1}}, Z^{\Delta}_{t_{i-1}})\Delta t_i \\ Z^{\Delta}_{t_{i-1}} &= \frac{1}{\Delta t_i}\mathbb{E}_{i-1}[Y^{\Delta}_{t_i}\Delta W_{t_i}] \end{aligned} \tag{11}$$

The last equation is obtained by multiplying both sides of (10) by ΔW_{t_i} and by taking the conditional expectation $\mathbb{E}_{i-1}[\cdot]$. Note that for Bachelier's model, the Malliavin weight for the delta is $\omega = \frac{\Delta W_T}{\Delta T}$. At the end, we need to use regression approximations or Malliavin's integration by part formula to compute the conditional expectations $\mathbb{E}_{i-1}[Y^{\Delta}_{t_i}\Delta W_{t_i}]$ and $\mathbb{E}_{i-1}[Y^{\Delta}_{t_i}]$. This scheme is *implicit* as $Y^{\Delta}_{t_{i-1}}$ appears in both sides of the second equation of (11). As a consequence, we need to use a Picard fixed point method. In order to overpass this difficulty, we can consider an *explicit* scheme where

$$Y^{\Delta}_{t_{i-1}} = \mathbb{E}_{i-1}[Y^{\Delta}_{t_i} + f(t_{i-1}, X^{\Delta}_{t_{i-1}}, Y^{\Delta}_{t_i}, Z^{\Delta}_{t_{i-1}})\Delta t_i]$$

or

$$Y^{\Delta}_{t_{i-1}} = \mathbb{E}_{i-1}[Y^{\Delta}_{t_i}] + f(t_{i-1}, X^{\Delta}_{t_{i-1}}, \mathbb{E}_{i-1}[Y^{\Delta}_{t_i}], Z^{\Delta}_{t_{i-1}})\Delta t_i$$

The implicit scheme converges as stated below.

Theorem 4.5 (Convergence, Bouchard and Touzi (2004))**.** *Let us consider the implicit scheme (11). Let us define $Y^{\Delta}_t = Y^{\Delta}_{t_i}$ and $Z^{\Delta}_t = Z^{\Delta}_{t_i}$ for $t \in [t_i, t_{i+1})$. Then*

$$\limsup_{\Delta \to 0} \Delta^{-1}\left(\sup_{0\le t\le T}\mathbb{E}[|Y^{\Delta}_t - Y_t|^2] + \mathbb{E}\left[\int_0^T |Z^{\Delta}_t - Z_t|^2 dt\right]\right) < \infty$$

The Black-Scholes-Barenblatt equation (3) is not connected to a 1-BSDE as in 1-BSDEs second-order derivatives only arise linearly through Itô's formula from the quadratic variation of the state process. This leads us to introduce second-order BSDEs.

4.2. **Second-order BSDEs.**

4.2.1. *Definition.*

Definition 4.6 (Second-order BSDE, Cheridito *et al.* (2007))**.** Let $(Y_t, Z_t, \Gamma_t, A_t)_{t\in[0,T]}$ be a quadruple of $(\mathcal{F}_t)$-adapted processes taking values in $\mathbb{R}$, $\mathbb{R}^d$, S^d and $\mathbb{R}^d$ respectively.[4] We call (Y, Z, Γ, A) a solution to a (Markovian) 2-BSDE if

$$\begin{aligned} dX_t &= \sigma(t, X_t)dW_t \\ dY_t &= -f(t, X_t, Y_t, Z_t, \Gamma_t)dt + Z_t \diamond \sigma(t, X_t)dW_t \\ dZ_t &= A_t dt + \Gamma_t\sigma(t, X_t)dW_t \\ Y_T &= g(X_T) \end{aligned} \tag{12}$$

[4] S^d stands for the space of d-dimensional symmetric matrices.

where $\diamond$ is the Stratonovich integral. The use of the Stratonovich product is only for convenience and can be replaced with an Itô integral

$$dY_t = \left(-f(t, X_t, Y_t, Z_t, \Gamma_t) + \frac{1}{2}\mathrm{tr}\left(\Gamma_t \sigma\sigma'(t, X_t)\right)\right) dt + Z_t\sigma(t, X_t)dW_t$$

4.2.2. *Second-order BSDEs provide an extension of Feynman-Kac's formula for fully non-linear PDEs.* Let us consider the fully non-linear parabolic PDE

$$\begin{aligned} \partial_t u(t,x) + f(t,x,u,\nabla_x u, \nabla_x^2 u) &= 0, \qquad (t,x) \in [0,T) \times \mathbb{R}^d \\ u(T,x) &= g(x) \end{aligned} \tag{13}$$

As in Proposition 4.3, a straightforward application of Itô's formula gives

Proposition 4.7 (Cheridito *et al.* (2007))**.** *Let u be a smooth function (smooth enough to apply Itô's formula) satisfying PDE (13). Then $(Y_t = u(t, X_t), Z_t = \nabla_x u(t, X_t), \Gamma_t = \nabla_x^2 u(t, X_t), A_t = (\partial_t + \mathcal{L})\nabla_x u(t, X_t))$ is a solution to 2-BSDE (12), where $\mathcal{L}$ is the Itô generator of X.*

Hence, in finance, Y is the price, Z the delta and Γ the gamma. Since in (13) f depends on the spatial second partial derivative of u, a stochastic representation of u is a 2-BSDE and cannot be a 1-BSDE.

Only a uniqueness result exists for 2-BSDEs. The precise result is based on two assumptions:

A(f): $f : [0,T) \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times S^d \to \mathbb{R}$ is continuous, Lipschitz in y uniformly in (t,x,z,Γ) and for some $C, p > 0$

$$|f(t,x,y,z,\Gamma)| \leq C\left(1 + |y| + |x|^p + |z|^p + |\Gamma|^p\right)$$

A(comp): if $\omega : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ (resp. u) is a lower semi-continuous (resp. u.s.c) viscosity supersolution (resp. subsolution) of (13) with $\omega(t,x) \geq -C(1+|x|^p)$ and $u(t,x) \leq C(1+|x|^p)$ then $\omega(T,\cdot) \geq u(T,\cdot)$ implies that $\omega \geq u$ on $[0,T] \times \mathbb{R}^d$.

Theorem 4.8 (Uniqueness, Cheridito *et al.* (2007))**.** *Under* **A(f)** *and* **A(comp)**, *for every g having polynomial growth, there is at most one solution to (12).*

4.2.3. *Numerical simulation.* Proceeding by analogy with 1-BSDEs, Cheridito *et al.* (2007) obtain the following discretization scheme for 2-BSDEs:

Scheme Cheridito *et al.*:

$$Y^\Delta_{t_n} = g(X^\Delta_{t_n}) \qquad Z^\Delta_{t_n} = \nabla g(X^\Delta_{t_n})$$

$$Y^\Delta_{t_{i-1}} = \mathbb{E}_{i-1}[Y^\Delta_{t_i}] + \left(f(t_{i-1}, X^\Delta_{t_{i-1}}, Y^\Delta_{t_{i-1}}, Z^\Delta_{t_{i-1}}, \Gamma^\Delta_{t_{i-1}}) - \frac{1}{2}\mathrm{tr}[\sigma(t_{i-1}, X^\Delta_{t_{i-1}})\sigma(t_{i-1}, X^\Delta_{t_{i-1}})'\Gamma^\Delta_{t_{i-1}}]\right)\Delta t_i$$

$$Z^\Delta_{t_{i-1}} = \frac{1}{\Delta t_i}\sigma(t_{i-1}, X^\Delta_{t_{i-1}})'^{-1}\mathbb{E}_{i-1}[Y^\Delta_{t_i}\Delta W_{t_i}]$$

$$\Gamma^\Delta_{t_{i-1}} = \frac{1}{\Delta t_i}\mathbb{E}_{i-1}[Z^\Delta_{t_i}\Delta W'_{t_i}]\sigma(t_{i-1}, X^\Delta_{t_{i-1}})^{-1}$$

The quantity P&L $\times\, \Delta t_i$ where

$$\text{P\&L} \equiv f(t_{i-1}, X^\Delta_{t_{i-1}}, Y^\Delta_{t_{i-1}}, Z^\Delta_{t_{i-1}}, \Gamma^\Delta_{t_{i-1}}) - \frac{1}{2}\mathrm{tr}[\sigma(t_{i-1}, X^\Delta_{t_{i-1}})\sigma(t_{i-1}, X^\Delta_{t_{i-1}})'\Gamma^\Delta_{t_{i-1}}] \tag{14}$$

is a Theta-Gamma P&L correction term between t_{i-1} and t_i. Because the driver f generally depends on the variable Y, to solve for $Y^{\Delta}_{t_{i-1}}$, one may use a Picard scheme. This is unnecessary when we deal with the UVM, because in this case f depends only on (X, Γ), see Section 5 below. This scheme requires a final condition for Z: $Z^{\Delta}_{t_n} = \nabla g(X^{\Delta}_{t_n})$. As the payoffs g usually considered in finance are not smooth, this scheme may perform poorly. For instance, for a digital option, Z_{t_n} is a Dirac mass. We suggest a new scheme in the next section.

Using an induction argument, one can show that the random variables $Y^{\Delta}_{t_i}$, $Z^{\Delta}_{t_i}$ and $\Gamma^{\Delta}_{t_i}$ are deterministic functions of $X^{\Delta}_{t_i}$. It then follows that the conditional expectations above can be replaced by

$$\begin{aligned}\mathbb{E}_{i-1}[Y^{\Delta}_{t_i}\Delta W_{t_i}] &= \mathbb{E}[Y^{\Delta}_{t_i}\Delta W_{t_i}|X^{\Delta}_{t_{i-1}}]\\ \mathbb{E}_{i-1}[Z^{\Delta}_{t_i}\Delta W'_{t_i}] &= \mathbb{E}[Z^{\Delta}_{t_i}\Delta W'_{t_i}|X^{\Delta}_{t_{i-1}}]\end{aligned}$$

Note that PDE (13) does not depend on the function $\sigma(\cdot,\cdot)$ which can be chosen arbitrarily. A proof of convergence of **Scheme Cheridito *et al.*** has been obtained recently in Fahim *et al.* (2009) in the case where the diffusion coefficient dominates the partial gradient of the Theta-Gamma P&L with respect to its Hessian component.

5. Solving the UVM with BSDEs

The Black-Scholes-Barenblatt (BSB) PDE (3) is a particular case of (13) with the driver f depending only on x and $\nabla^2_x u$:

$$f(X,\Gamma) = \frac{1}{2}\max_{(\sigma^{\alpha},\rho^{\alpha\beta})_{1\le\alpha<\beta\le d}\in D}\sum_{\alpha,\beta=1}^{d}\rho^{\alpha\beta}\sigma^{\alpha}\sigma^{\beta}X^{\alpha}X^{\beta}\Gamma^{\alpha\beta}$$

The 2-BSDE associated to the BSB equation is

$$\begin{aligned}dX^{\alpha}_t &= \hat{\sigma}^{\alpha}X^{\alpha}_t dW^{\alpha}_t, \qquad dW^{\alpha}_t dW^{\beta}_t = \hat{\rho}^{\alpha\beta}dt, \qquad 1\le\alpha<\beta<d\\ dY_t &= -f(X_t,\Gamma_t)dt + \sum_{\alpha=1}^{d} Z^{\alpha}_t \diamond \hat{\sigma}^{\alpha}X^{\alpha}_t dW^{\alpha}_t\\ dZ^{\alpha}_t &= A^{\alpha}_t dt + \sum_{\beta=1}^{d}\Gamma^{\alpha\beta}_t\hat{\sigma}^{\beta}X^{\beta}_t dW^{\beta}_t\\ Y_T &= g(X_T)\end{aligned} \tag{15}$$

We are free to choose the diffusion $\sigma(\cdot,\cdot)$: we pick a log-normal dynamics for X with some constant volatility $\hat{\sigma}^{\alpha}$ and some constant correlation $\hat{\rho}^{\alpha\beta}$.

From Theorem 2.1, under regularity and growth assumptions on the payoff g, the solution u to PDE (3) is smooth. By Proposition 4.7, $Y_t = u(t, X_t)$ is a solution to (15). Since D is compact, $\mathbf{A(f)}$ and $\mathbf{A(comp)}$ hold so that Theorem 4.8 applies and Y is the unique solution to (15).

5.1. **A new numerical scheme.** We suggest a new numerical scheme for the Black-Scholes-Barenblatt 2-BSDE (15). After dropping the Δ superscript, this scheme reads

New scheme for UVM:

$$
\begin{aligned}
&X^\alpha_{t_i} = X^\alpha_0 e^{-(\hat\sigma^\alpha)^2 \frac{t_i}{2} + \hat\sigma^\alpha W^\alpha_{t_i}}, \qquad \mathbb{E}[\Delta W^\alpha_{t_i} \Delta W^\beta_{t_i}] = \hat\rho^{\alpha\beta} \Delta t_i \\
&Y_{t_n} = g(X_{t_n}) \\
&Y_{t_{i-1}} = \mathbb{E}[Y_{t_i} | X_{t_{i-1}}] + \left(f(X_{t_{i-1}}, \Gamma_{t_{i-1}}) - \frac{1}{2} \sum_{\alpha,\beta=1}^{d} \hat\rho^{\alpha\beta} \hat\sigma^\alpha \hat\sigma^\beta X^\alpha_{t_{i-1}} X^\beta_{t_{i-1}} \Gamma^{\alpha\beta}_{t_{i-1}} \right) \Delta t_i \qquad (16) \\
&(\Delta t_i)^2 \hat\sigma^\alpha \hat\sigma^\beta X^\alpha_{t_{i-1}} X^\beta_{t_{i-1}} \Gamma^{\alpha\beta}_{t_{i-1}} = \mathbb{E}\left[Y_{t_i} \left(U^\alpha_{t_i} U^\beta_{t_i} - \Delta t_i \hat\rho^{-1}_{\alpha\beta} - \Delta t_i \hat\sigma^\alpha U^\alpha_{t_i} \delta_{\alpha\beta} \right) \middle| X_{t_{i-1}} \right]
\end{aligned}
$$

with $U^\alpha_{t_i} \equiv \sum_{\beta=1}^d \hat\rho^{-1}_{\alpha\beta} \Delta W^\beta_{t_i}$. Note that our scheme differs from **Scheme Cheridito *et al.*** in that we have changed the discretization for the Gamma Γ by using explicitly the Malliavin weight for a log-normal diffusion with volatility $\hat\sigma$ and correlation $\hat\rho$.

Considering the case where $i = n$ helps understand what **New scheme** does: since we simulated log-normal X's, the price $Y_{t_{n-1}}$ is the sum of the Black-Scholes price $\mathbb{E}[Y_{t_n} | X_{t_{n-1}}]$ and the Theta-Gamma correction P&L

$$
\left(f(X_{t_{n-1}}, \Gamma_{t_{n-1}}) - \frac{1}{2} \sum_{\alpha,\beta=1}^{d} \hat\rho^{\alpha\beta} \hat\sigma^\alpha \hat\sigma^\beta X^\alpha_{t_{n-1}} X^\beta_{t_{n-1}} \Gamma^{\alpha\beta}_{t_{n-1}} \right) \Delta t_n
$$

This last term requires that we estimate the Gamma at time t_{n-1}. We use the Black-Scholes Gamma, as given by the last equation in (16), which uses the Malliavin weight for the lognormal diffusion.

This new scheme performs better in our numerical experiments. In their recent preprint, Fahim *et al.* (2009) study two schemes: **Scheme Cheridito *et al.*** introduced above, and **Scheme Fahim *et al.*** where the Gamma process is directly estimated from the price process:

Scheme Fahim *et al.*:

$$
\begin{aligned}
&Y^\Delta_{t_n} = g(X^\Delta_{t_n}) \\
&Y^\Delta_{t_{i-1}} = \mathbb{E}_{i-1}[Y^\Delta_{t_i}] + \left(f(t_{i-1}, X^\Delta_{t_{i-1}}, Y^\Delta_{t_{i-1}}, Z^\Delta_{t_{i-1}}, \Gamma^\Delta_{t_{i-1}}) - \frac{1}{2} \mathrm{tr}[\sigma(t_{i-1}, X^\Delta_{t_{i-1}}) \sigma(t_{i-1}, X^\Delta_{t_{i-1}})' \Gamma^\Delta_{t_{i-1}}] \right) \Delta t_i \\
&\Gamma_{t_{i-1}} = \mathbb{E}_{i-1}[Y_{t_i} (\sigma(t_{i-1}, X^\Delta_{t_{i-1}})')^{-1} \frac{\Delta W_{t_i} (\Delta W_{t_i})' - \Delta t_i \mathrm{Id}}{\Delta t_i^2} \sigma(t_{i-1}, X^\Delta_{t_{i-1}})^{-1}]
\end{aligned}
$$

The term $(\sigma')^{-1} \frac{\Delta W_{t_i} (\Delta W_{t_i})^\dagger - \Delta t_i \mathrm{Id}}{\Delta t_i^2} \sigma^{-1}$ is the Gamma Malliavin weight for the Bachelier model. From the definition of process (X_t), it seems more appropriate to use **New scheme** which involves the Gamma Malliavin's weight for a log-normal process. In the particular case where P&L $= 0$, see (14), the non-linear PDE reduces to a Black-Scholes PDE and **New scheme** is exact, contrary to **Scheme Cheridito *et al.*** and **Scheme Fahim *et al.***. Also, for convex or concave European payoffs, **New scheme** is exact, contrary to the two others.

Fahim *et al* state that **Scheme Cheridito *et al.*** performs better than **Scheme Fahim *et al.*** in their numerical experiments. Note that they only tested their schemes on smooth payoffs ($g(x) = x^\eta$). Since g is smooth in all their examples, they set $Z_T = \nabla g(X_T)$. This is not applicable for discontinuous payoffs such as a digital option: $g(x) = 1_{x \geq K}$.

New scheme requires computing $\frac{d(d+1)}{2} + 1$ conditional expectations at each discretization date. For this purpose, as for the valuation of American options, one can use

- Parametric regression as in the Tsitsiklis-Van Roy (2001) or the Longstaff-Schwartz (2001) methods (see Gobet *et al.* (2005) in the case of 1-BSDEs):

$$\mathbb{E}[Y_{i+1}|X_i = x] \approx \sum_{k=1}^{N} c_k \phi_k(x)$$

- Non-parametric regression:

$$\mathbb{E}[Y_{i+1}|X_i = x] \approx \frac{\mathbb{E}[Y_{i+1}\delta_N(X_i - x)]}{\mathbb{E}[\delta_N(X_i - x)]}$$

  with $\delta_N(\cdot)$ a kernel approximating a Dirac mass at zero.
- Likelihood ratio weight:

$$\mathbb{E}[Y_{i+1}|X_i = x] = \mathbb{E}\left[Y_{i+1}\frac{p(t_i, t_{i+1}, x, X_{i+1})}{p(0, t_{i+1}, X_0, X_{i+1})}\right]$$

  with p the joint log-normal distribution.
- Malliavin's weight (see Bouchard and Touzi (2004) in the case of 1-BSDEs) .

Remark 5.1 (Transaction cost). For transaction costs, the third equation in **New scheme** is replaced by

$$Y_{t_{i-1}} = \mathbb{E}[Y_{t_i}|X_{t_{i-1}}] + \sqrt{\frac{2}{\pi\delta t}} \sum_{\alpha=1}^{d} k_\alpha X^\alpha_{t_{i-1}} \sqrt{\sum_{\beta,\gamma=1}^{d} \rho^{\beta\gamma}\sigma^\beta\sigma^\gamma X^\beta_{t_{i-1}} X^\gamma_{t_{i-1}} \Gamma^{\alpha\beta}_{t_{i-1}} \Gamma^{\alpha\gamma}_{t_{i-1}}} \Delta t_i$$

where X follows an Itô diffusion process.

In our numerical experiments for the second approach, we have used non-parametric regressions (with Gaussian kernels) in one dimension and parametric regressions in two dimensions with suitable basis functions. Likelihood ratio weights have too much variance. Before presenting our results, we test the algorithm in the case of an at-the-money call option with payoff $(X_T - X_0)^+$. This enables us to state the details of our algorithm.

5.2. **First example: at-the-money call option.** We take $T = 1$, $X_0 = 100$, $\underline{\sigma} = 0.1$, $\overline{\sigma} = 0.2$ and $\hat{\sigma} = 0.15$. The true price, $\mathcal{C} = 7.97$, is the Black-Scholes price with the upper volatility as stated in Remark 2.2. The results are reported in Table 1. In this table, $N_1 = 2^{M_1}$ is the number of paths used to compute the conditional expectations, and $\Delta = t_i - t_{i-1} = 1/n$ is the constant discretization time step.

Δ	M_1	12	13	14	15	16	17
1/2	Price	8.04	8.06	8.00	8.00	8.00	8.00
1/4	Price	8.53	8.29	8.00	7.87	7.89	7.86
1/8	Price	9.28	8.72	8.02	7.79	7.78	7.68

TABLE 1. At-the-money call option valued using the BSDE approach with volatility $\hat{\sigma} = 0.15$. The true price is $\mathcal{C} = 7.97$.

As Δ becomes small, we need more and more simulations to obtain an accurate price. E. Gobet *et al.* (2005), for 1-BSDEs, and A. Fahim *et al.* (2009), for 2-BSDEs, also noticed that the numerical scheme diverges when Δ goes to zero, M_1 being fixed. For instance, for $\Delta = 1/8$, the price has not converged yet with $M_1 = 17$. Furthermore, as observed on this simple example, the algorithm produces an unpredictable bias. In order to build a low-biased estimate, we simulate N_2 replicas of

$$dX_t^\alpha = \sigma_t^{*\alpha} X_t^\alpha dW_t^\alpha, \qquad dW_t^\alpha dW_t^\beta = \rho_t^{*\alpha\beta} dt, \qquad 1 \le \alpha < \beta \le d$$

in an independent second Monte-Carlo procedure, where $\sigma_t^{*\alpha}$ and $\rho_t^{*\alpha\beta}$ are the solutions to

$$\max_{(\sigma^\alpha, \rho^{\alpha\beta})_{1\le\alpha<\beta\le d} \in D} \sum_{\alpha,\beta=1}^{d} \rho^{\alpha\beta}\sigma^\alpha\sigma^\beta X_t^\alpha X_t^\beta \Gamma_t^{\alpha\beta} \tag{17}$$

and $\Gamma_t^{\alpha\beta} = \varphi(t, X_t)$, with φ the result of the regression step. Because the covariance matrix is suboptimal, we obtain a low-biased estimator. This is a commonly used technique for the pricing of American options in Monte-Carlo, see Glasserman (2003). We run the second Monte-Carlo with $N_2 = 2^{15}$ paths and a time step of $\Delta_2 = 1/400$ for the forward discretization of X. Results are reported in Table 2.

Δ	M_1	12	13	14	15	16	17
1/2	Price	7.93	7.94	7.95	7.95	**7.96**	**7.96**
1/4	Price	7.88	7.90	7.92	7.93	7.95	7.94
1/8	Price	7.53	7.93	7.58	7.60	**7.96**	**7.96**

TABLE 2. At-the money call option valued using the BSDE approach with volatility $\hat{\sigma} = 0.15$ and an independent second MC. The true price is $\mathcal{C} = 7.97$.

Since the estimator is low-biased, the price is larger than *each* of the prices in Table 2. In this table, we have highlighted in bold the higher price estimates. Table 2 states that the true price in the UVM is *certainly* more than or equal to 7.96. From Table 1, i.e., when the bias is unknown, one cannot make such a claim and it is hard to guess where the true price is. This is the reason why we prefer to work with estimators with known bias.

5.3. **The algorithm.** The final meta-algorithm for pricing can be summarized in the following steps:

- Simulate N_1 replicas of X with a log-normal diffusion.
- Apply the backward algorithm **New scheme** using a regression approximation. In a high-dimensional problem, the parametric regression is the most appropriate.
- Simulate N_2 independent replicas of X using the gamma functions computed at the previous step.

Note that the payoffs g that we will use in our numerical experiments below do not satisfy the regularity condition $g \in \mathcal{C}^3$ under which we stated the existence and uniqueness of 2-BSDE (15). However, even for these non-smooth payoffs, our discretization scheme seems to converge numerically.

Remark 5.2 (Another tool to choose a relevant parameterization: Combination of the BSDE and parametric approaches). Here we present an alternative to Subsection 3.3. We may use the BSDE algorithm to choose an efficient parameterization in the parametric approach. For each date t in the time grid, we graph the optimal $\sigma_t^{*\alpha}$ and $\rho_t^{*\alpha\beta}$ solutions to (17) as functions of X_t (and, possibly, of path-dependent variables) and guess a parameterization for them. Note that contrary to the method suggested in Subsection 3.3, this process involves only one Monte-Carlo computation.

5.4. **About the generation of the first N_1 paths.** In both the parametric and the BSDE algorithms, the volatility function $\sigma(\cdot,\cdot)$ is arbitrary. We can choose different volatility functions $\sigma(\cdot,\cdot)$'s to generate the first N_1 replicas of X. Different $\sigma(\cdot,\cdot)$'s lead to different sets $(X^{(p)}, A^{(p)})_{1\le p\le N_1}$ which serve in the optimization or regression procedures, hence to different optimal covariance matrix estimates. Of course, the exact matrix does not depend on $\sigma(\cdot,\cdot)$. Here, we have chosen a log-normal diffusion with a mid-volatility $\hat{\sigma}$ but other choices are possible. For instance:

- Before proceeding to Step 3, we may repeat Steps 1 and 2, replacing $(\hat{\rho}^{\alpha\beta}\hat{\sigma}^{\alpha}\hat{\sigma}^{\beta})$ by the optimal covariance matrix estimate as computed at Step 2. This should improve the estimate and result in a better lower bound for the price in Step 3. In the BSDE approach, this should reduce the contribution of the P&L Gamma-Theta term (14) to $Y_{t_{i-1}}$ in (16) and therefore the impact of the error in the computation of $\Gamma_{t_{i-1}}$ through regression.
- Allowing the $(X^{(p)}, A^{(p)})_{1\le p\le N_1}$ points to completely cover the diffusion support of the UVM should result in more accurate optimization or regression results. This can be achieved by choosing different volatility function $\sigma(\cdot,\cdot)$ for different paths. For instance, in the single-asset case, for payoffs depending on realized variance, we may choose

$$\sigma^{(p)} = \begin{cases} \underline{\sigma} & \text{if } p < N_1/3 \\ (\underline{\sigma}+\overline{\sigma})/2 & \text{if } N_1/3 \le p < 2N_1/3 \\ \overline{\sigma} & \text{if } 2N_1/3 \le p \end{cases}$$

6. Numerical experiments

Here we present numerical results for both approaches.

In our experiments, we take $T = 1$, and, for each asset α, $X_0^\alpha = 100$, $\underline{\sigma}^\alpha = 0.1$, $\overline{\sigma}^\alpha = 0.2$ and we use the constant mid-volatility $\hat{\sigma}^\alpha = 0.15$ to generate the first N_1 replicas of X. We also pick $t_i = i/n$, so that $\Delta = 1/n$. In the pricing stage, the $N_2 = 2^{15}$ replicas of X use a time step $\Delta_2 = 1/400$. **First parameterization method:** In the gamma calibration stage, we pick $N_1 = 2^{M_1}$ with $M_1 = 12$, and the N_1 replicas of X use a time step $t_{k+1} - t_k = 1/100$. **Second BSDE method:** We allow M_1 to vary from 12 to 17.

6.1. **Options with one underlying.** Let us first consider three payoffs depending on a single asset: a call spread, a digital option and a call Sharpe.

Call spread. (See Table 3) Let us test our two algorithms in the case of a call spread option with payoff $(X_T - K_1)^+ - (X_T - K_2)^+$. We pick $K_1 = 90$ and $K_2 = 110$. The true price (PDE) is $\mathcal{C}_{\text{PDE}} = 11.20$ and the Black-Scholes price with the mid-volatility is $C_{\text{BS}} = 9.52$. **First parameterization method:** Because of the shape of the call spread payout, we pick the following parameterization of the volatility: $\theta_i \in \mathbb{R}$ and $\lambda_{t_i}(X;\theta_i) = \overline{\sigma}$ if $\theta_i - \ln(X/X_0) \ge 0$, $= \underline{\sigma}$ otherwise, i.e., $\gamma_{t_i}(X;\theta_i) = \theta_i - \ln(X/X_0)$. For instance, for $\Delta = 1/4$, the numerical optimal gamma frontier is given by $(\theta_0^*, \theta_1^*, \theta_2^*, \theta_3^*) = (0.02, 0.02, 0.02, 0.02)$. **Second BSDE method:** We use non-parametric regressions.

Δ	Param	M_1	12	13	14	15	16	17
1/2	**11.19**	Price	11.08	11.07	11.06	11.06	11.06	11.06
1/4	**11.19**	Price	11.01	**11.12**	11.06	11.07	11.11	11.11
1/8	11.18	Price	10.74	10.55	10.73	11.01	11.04	11.11

TABLE 3. Call spread valued using the parametric approach (column 2) and the BSDE approach (columns 4 to 9). The true price (PDE) is $\mathcal{C}_{\text{PDE}} = 11.20$.

Since our parameterization of the optimal gamma frontier is exact in this case, the parametric approach gives very accurate results. Besides, as $\theta_{t_i}^*$ varies very slowly with t_i, this method proves to be efficient even with $\Delta = 1/2$, i.e., even when the gamma frontier is updated twice a year. The BSDE approach captures the right magnitude of the price but is not able to produce a lower bound greater than 11.12.

In Table 4, we illustrate the fact that although the diffusion coefficient $\hat{\sigma}$ can be chosen arbitrarily in both algorithms, an optimal choice can lead to better numerical results (see Subsection 5.4). We have computed the price of the above call spread option as a function of the volatility $\hat{\sigma}$. We pick $\Delta = 1/4$ and $M_1 = 16$. We also report the Black-Scholes price with volatility $\hat{\sigma}$. **First parameterization method:** The one-dimensional downhill simplex method, with initial guess $\theta_3^0 = 0$ and a first simplex side of 0.02, gives the same optimal $(\theta_0^*, \theta_1^*, \theta_2^*, \theta_3^*) = (0.02, 0.02, 0.02, 0.02)$ for all $\hat{\sigma}$ from 2% to 30%. For $\hat{\sigma} = 50\%$, it finds $(\theta_0^*, \theta_1^*, \theta_2^*, \theta_3^*) = (0, 0, 0, 0)$. **Second BSDE method:** It turns out that, among all the constant volatilities tested, the optimal constant volatility corresponds to the mid-volatility $\hat{\sigma} = 15\%$.

Algo/$\hat{\sigma}$	2%	5%	10%	15%	20%	30%	50%
Param	**11.19**	**11.19**	**11.19**	**11.19**	**11.19**	**11.19**	11.14
BSDE	9.36	10.72	11.01	**11.12**	11.07	11.00	10.81
Black-Scholes	10.00	9.97	9.76	9.52	9.30	8.87	8.06

TABLE 4. Call spread valued using the parametric and BSDE approaches with different volatilities $\hat{\sigma}$. The true price (PDE) is $\mathcal{C}_{\text{PDE}} = 11.20$. $\Delta = 1/4$ and $M_1 = 16$. We also report the Black-Scholes price with volatility $\hat{\sigma}$.

Digital option. (See Table 5) Let us now test the algorithms with the digital option delivering $100 \times 1_{X_T \geq K}$. We pick $K = 100$. The true price (PDE) is $\mathcal{C}_{\text{PDE}} = 63.33$ and the Black-Scholes price with the mid-volatility is $C_{\text{BS}} = 46.54$. **First parameterization method:** Because the payouts are similar, we pick the same parameterization as for the call spread. **Second BSDE method:** We use non-parametric regressions.

Again the parameterization of the optimal gamma frontier is exact in this case, so the parametric method gives very accurate results. The BSDE method performs well, except when Δ is too small. This example shows that discontinuous payoffs can be accurately priced by our algorithms as well.

Call Sharpe. (See Table 6) To finish with the single-asset examples, let us test the algorithms with a call Sharpe option delivering $(X_T - 100)^+/\sqrt{V_T}$ where $V_T = \frac{1}{T}\sum_{l=1}^{12}\left(\ln \frac{X_{t_l}}{X_{t_{l-1}}}\right)^2$ is the

Δ	Param	M_1	12	13	14	15	16	17
1/2	63.13	Price	**62.83**	62.83	62.74	62.75	62.75	62.74
1/4	**63.14**	Price	62.53	62.86	62.77	62.35	62.45	62.43
1/8	62.68	Price	60.06	59.16	60.56	60.59	60.94	60.53

TABLE 5. Digital option valued using the parametric approach (column 2) and the BSDE approach (columns 4 to 9). The true price (PDE) is $\mathcal{C}_{\text{PDE}} = 63.33$ and the Black-Scholes price with the mid-volatility is $C_{\text{BS}} = 46.54$.

realized volatility computed using monthly returns. At time t, the option value depends on X_t and on the two path-dependent variables

$$\begin{aligned} A_t^1 &= \sum_{\{l|t_l \leq t\}} \left(\ln \frac{X_{t_l}}{X_{t_{l-1}}} \right)^2 \\ A_t^2 &= X_{\sup_{\{l|t_l \leq t\}} t_l} \end{aligned}$$

First parameterization method: We take $\theta_i = (\theta_i^1, \theta_i^2) \in \mathbb{R}^2$ and pick $\gamma_{t_i}(X, A; \theta_i) = \theta_i^1 \sqrt{A^1} + \theta_i^2 - \ln(X/X_0)$. **Second BSDE method:** It is notably difficult to find a convenient basis to compute the conditional expectations and we assume as a first approximation that $\mathbb{E}_{i-1}[\cdot] \equiv \mathbb{E}[\cdot|X_{t_{i-1}}, A_{t_{i-1}}^1, A_{t_{i-1}}^2] \simeq \mathbb{E}[\cdot|X_{t_{i-1}}]$. The latter is computed using a one dimensional non-parametric regression.

Δ	Param	M_1	12	13	14	15	16	17	18
1/2	54.98	Price	47.73	47.18	48.82	48.09	48.10	48.01	48.09
1/4	**55.55**	Price	46.93	47.34	48.01	48.92	48.67	49.38	49.44
1/12	54.32	Price	48.03	49.26	49.78	50.87	51.11	51.66	**52.12**

TABLE 6. Call Sharpe valued using the parametric approach (column 2) and the BSDE approach (columns 4 to 9). The true price (PDE) is $\mathcal{C}_{\text{PDE}} = 58.4$. The Black-Scholes price with mid-volatility is $\mathcal{C}_{\text{BS}} = 40.71$.

In this case, the optimal gamma frontier at date t_i, i.e., the optimal function $\gamma_{t_i}(\cdot)$, does not fall within the parametized set $\{\gamma_{t_i}(\cdot;\theta_i)|\theta_i \in \mathbb{R}^2\}$ (see Subsection 3.4). Hence the parametric approach is less accurate than it was in the cases of the call spread and digital options. Nevertheless, it gives lower bounds around 55, that is, it captures the correct magnitude of the exact move, from 40.7 in the Black-Scholes model to 58.4 in the UVM. If we choose $\theta_i \in \mathbb{R}$ and $\gamma_{t_i}(X, A; \theta_i) = \theta_i - \ln(X/X_0)$ with $\Delta = 1/4$, we get $(\theta_0^*, \theta_1^*, \theta_2^*, \theta_3^*) = (0.075, 0.075, 0.025, -0.0125)$ and a lower bound estimate of 53.85, a value close to 55.55. This shows that the realized variance A^1 hardly affects the gamma frontier. Since the conditional expectations in the BSDE approach were computed as if they only depended on X, the BSDE lower bound prices can be compared to 53.85.

Remark 6.1 (Continuously computed variance). The BSDE approach can easily be adapted to the case when the realized variance V_t changes continuously. In Remark 2.4, we have shown that in this case the price of the option in the UVM can be written $u(t, X_t, V_t)$ where u is solution to

$$\begin{aligned} \partial_t u(t,x,v) + f(x, \partial_x^2 u(t,x,v), \partial_v u(t,x,v)) &= 0, \qquad (t,x,v) \in [0,T) \times \mathbb{R}_+^* \times \mathbb{R}_+^* \\ f(x, \partial_x^2 u, \partial_v u) &= \max_{\underline{\sigma} \leq \sigma \leq \overline{\sigma}} \sigma^2 \left(\frac{1}{2} x^2 \partial_x^2 u + \partial_v u \right) \end{aligned}$$

We can associate a two-dimensional 2-BSDE on the (X,V) plane to this fully non-linear PDE:

$$\begin{aligned} dX_t &= \hat{\sigma} X_t dW_t^0 \\ dV_t &= \hat{\sigma}^2 dt + \eta dW_t^1 \\ dY_t &= \left(-f(X_t, \Gamma_t^{XX}, Z_t^V) + \mathcal{L}^{X,V} u(t, X_t, V_t)\right) dt + Z_t^S \hat{\sigma} X_t dW_t^0 + Z_t^V \eta dW_t^1 \end{aligned}$$

with

$$\mathcal{L}^{X,V} u(t, X_t, V_t) = \frac{1}{2}\hat{\sigma}^2 \left(X^2 \Gamma^{XX} + 2Z^V\right) + \frac{1}{2}\eta^2 \Gamma^{VV}$$

We used $\hat{\sigma}^2$ as the (forward) drift for the variance V, but this is arbitrary. We have introduced a diffusion term for V_t. Here η is a constant and W^1 a Brownian motion orthogonal to W^0. Adding this (purely numerical) volatility term allows us to compute $Z_t^V \equiv \partial_V u$. Just as the solution u of the PDE, the 2-BSDE is independent of η, but the numerical scheme depends on it. A too small or too large value for η would lead to a bad regression-based estimation of Z_t^V.

6.2. **Options with two underlyings and no uncertainty on correlation.** We consider (projective) payoffs depending on two correlated assets with a constant certain correlation ρ which can be written as

$$g(X_T^1, X_T^2) = X_T^1 \mathcal{G}\left(\frac{X_T^2}{X_T^1}\right)$$

This simple payoff form together with the certainty on the correlation parameter makes it possible to reduce the two-dimensional Black-Scholes-Barenblatt PDE to a simple one-dimensional Black-Scholes-Barenblatt PDE. Indeed by changing from the risk-neutral measure $\mathbb{P}$ to the measure $\mathbb{P}^1$ associated to the numéraire X^1, we obtain

$$\mathbb{E}^{\mathbb{P}}[g(X_T^1, X_T^2)|\mathcal{F}_t] = X_t^1 \mathbb{E}^{\mathbb{P}^1}[\mathcal{G}(X_T)|\mathcal{F}_t]$$

Here, $X_t \equiv \frac{X_t^2}{X_t^1}$ is a local martingale under $\mathbb{P}^1$, $(\tilde{W}_t^1, \tilde{W}_t^2) = (W_t^1 - \int_0^t \sigma_s^1 ds, W_t^2 - \rho \int_0^t \sigma_s^1 ds)$ is a Brownian motion with correlation ρ under $\mathbb{P}^1$, and

$$\begin{aligned} dX_t &= X_t \left(\sigma_t^2 d\tilde{W}_t^2 - \sigma_t^1 d\tilde{W}_t^1\right) \\ &\overset{\text{Law}}{=} X_t \sigma_t d\tilde{W}_t \end{aligned}$$

with the control $(\sigma_t)^2$ taking values in some interval $I = [\underline{\sigma}_I^2, \overline{\sigma}_I^2]$. When $\rho \leq 0$,

$$I = \left[\sum_{\alpha=1}^{2} (\underline{\sigma}^\alpha)^2 - 2\rho \underline{\sigma}^1 \underline{\sigma}^2, \sum_{\alpha=1}^{2} (\overline{\sigma}^\alpha)^2 - 2\rho \overline{\sigma}^1 \overline{\sigma}^2\right] \tag{18}$$

When $\rho > 0$, one can get $\underline{\sigma}_I$ and $\overline{\sigma}_I$ using the method described for computing the Hamiltonian (5), after replacing $\sigma^\alpha \sigma^\beta X^\alpha X^\beta \Gamma^{\alpha\beta}$ by $\sigma^\alpha \sigma^\beta$. Finally, we have

$$\max_{\sigma_s^1, \sigma_s^2,\, s \in [t,T]} \mathbb{E}^{\mathbb{P}}[g(X_T^1, X_T^2)|\mathcal{F}_t] = X_t^1 \max_{(\sigma_s)^2 \in I,\, s \in [t,T]} \mathbb{E}^{\mathbb{P}^1}[\mathcal{G}(X_T)|\mathcal{F}_t]$$

Outperformer option. As an example of projective payoff, we consider an outperformer option $(X_T^1 - X_T^2)^+$. Because $\mathcal{G}$ is the put payoff, when $\rho \leq 0$, the true value is given by the Black-Scholes price with high volatilities. **First parameterization method:** We choose $\theta_i = (\theta_i^1, \theta_i^2) \in \mathbb{R}^2$, $\sigma_{t_i}^1(X; \theta_i) = \overline{\sigma}^1$ if $\theta_i^1 - \ln(X^1/X_0^1) \geq 0$, $= \underline{\sigma}^1$ otherwise, and $\sigma_{t_i}^2(X; \theta_i) = \overline{\sigma}^2$ if $\theta_i^2 - \ln(X^2/X_0^2) \geq 0$, $= \underline{\sigma}^2$ otherwise. **Second BSDE method:** We try two choices of basis functions. First, we use X^1 and X^2. Second, we also add the symmetric second-order polynomials $X^\alpha X^\beta$, $\alpha, \beta = 1, 2$. In all cases, a constant is included in the regression. First, we test the method assuming zero correlation (see Tables 7 and 8). Then, we assume a correlation $\rho = -0.5$ (see Table 9).

Δ	Param	M_1	12	13	14	15	16	17
1/2	**11.26**	Price	**11.24**	**11.24**	**11.24**	**11.24**	**11.24**	**11.24**
1/4	**11.26**	Price	9.65	10.11	10.04	10.16	10.28	9.69
1/8	**11.26**	Price	9.17	9.45	9.14	9.26	9.40	9.47
1/12	**11.26**	Price	9.17	9.67	9.47	9.38	9.32	9.84

TABLE 7. Outperformer option with 2 uncorrelated assets valued using the parametric approach (column 2) and the BSDE approach (columns 4 to 9). The true price is $\mathcal{C} = 11.25$. Basis functions $= \{1, X^1, X^2\}$.

Δ	Param	M_1	12	13	14	15	16	17
1/2	**11.26**	Price	11.09	11.15	11.15	11.15	11.15	11.15
1/4	**11.26**	Price	10.95	11.07	11.10	11.13	11.16	**11.19**
1/8	**11.26**	Price	10.35	10.71	10.88	10.94	11.07	11.14
1/12	**11.26**	Price	10.39	10.68	10.74	10.91	11.02	11.13

TABLE 8. Outperformer option with 2 uncorrelated assets valued using the parametric approach (column 2) and the BSDE approach (columns 4 to 9). The true price is $\mathcal{C} = 11.25$. Basis functions $= \{1, X^1, X^2, (X^1)^2, (X^2)^2, X^1X^2\}$.

Δ	Param	M_1	12	13	14	15	16	17
1/2	**13.77**	Price	13.58	13.66	**13.68**	13.66	13.66	13.65
1/4	13.74	Price	13.13	13.48	13.50	13.58	13.64	13.68
1/8	13.74	Price	12.45	12.87	13.12	13.20	13.48	13.63
1/12	13.75	Price	12.52	12.82	12.98	13.27	13.44	13.59

TABLE 9. Outperformer option with 2 correlated assets ($\rho = -0.5$) valued using the parametric approach (column 2) and the BSDE approach (columns 4 to 9). The true price is $\mathcal{C} = 13.75$. Basis functions $= \{1, X^1, X^2, (X^1)^2, (X^2)^2, X^1X^2\}$.

The parametric method selects high volatilities everywhere and hence gives excellent results. These results can lie slightly above the true price due to the Monte-Carlo sampling error. In the BSDE approach, the choice of basis functions clearly affects the price estimate. In particular for the trivial basis $= \{1, X_1, X_2\}$, when $\Delta \geq 1/4$, the numerical gammas resulting from regressions often take negative values. This does not happen with $\Delta = 1/2$ nor with the second basis. Furthermore, as observed previously, as Δ becomes smaller, we need more and more simulations to obtain an accurate price. However, as our estimator is low-biased, for a fixed Δ, we should increase M_1 as long as the price increases.

Outperformer spread option. Finally, we move to a more complex projective payoff $(X_T^2 - K_1 X_T^1)^+ - (X_T^2 - K_2 X_T^1)^+$ with $K_1 = 0.9$ and $K_2 = 1.1$. We pick the certain correlation $\rho = -0.5$. As explained above, this option can be valued using our Monte-Carlo algorithms or a numerical solution of the one-dimensional Black-Scholes-Barenblatt PDE (see the beginning of this subsection) with $\underline{\sigma} = 17.32\%$ and $\overline{\sigma} = 34.64\%$. We found $\mathcal{C}_{\mathrm{PDE}} = 11.41$. **First parameterization method:** We choose $\theta_i = (\theta_i^1, \theta_i^2) \in \mathbb{R}^2$, $\sigma^1_{t_i}(X;\theta_i) = \overline{\sigma}^1$ if $\theta_i^1 - \ln(X^2/X^1) \geq 0$, $= \underline{\sigma}^1$ otherwise, and $\sigma^2_{t_i}(X;\theta_i) = \overline{\sigma}^2$ if $\theta_i^2 - \ln(X^2/X^1) \geq 0$, $= \underline{\sigma}^2$ otherwise. **Second BSDE method:** We use two sets of basis functions (see Tables 10 and 11). The second set respects the fact that the exact price can be written as $X^1 u(t, X^2/X^1)$ and clearly improves the price estimator.

Δ	Param	M_1	12	13	14	15	16	17
1/2	**11.37**	Price	11.07	11.07	11.10	**11.11**	11.09	11.10
1/4	11.37	Price	10.32	10.40	10.56	10.70	10.74	10.60
1/8	11.37	Price	9.54	9.51	9.29	8.94	9.05	9.34

TABLE 10. Outperformer spread option with 2 correlated assets ($\rho = -0.5$) valued using the parametric approach (column 2) and the BSDE approach (columns 4 to 9). The true price is $\mathcal{C} = 11.41$. The Black-Scholes price with mid-volatilities is $\mathcal{C}_{\mathrm{BS}} = 9.04$. Basis functions $= \{1, X^1, X^2, (X^1)^2, (X^2)^2, X^1X^2\}$.

Δ	Param	M_1	12	13	14	15	16	17
1/2	**11.37**	Price	11.24	11.25	11.26	11.26	11.27	11.27
1/4	**11.37**	Price	10.85	11.12	11.20	11.27	**11.29**	11.28
1/8	11.37	Price	9.51	9.56	9.60	9.60	10.07	10.31

TABLE 11. Outperformer spread option with 2 correlated assets ($\rho = -0.5$) valued using the parametric approach (column 2) and the BSDE approach (columns 4 to 9). The true price is $\mathcal{C} = 11.41$. The Black-Scholes price with mid-volatilities is $\mathcal{C}_{\mathrm{BS}} = 9.04$. Basis functions $= \{X^1, X^2, \frac{(X^2)^2}{X^1}, \frac{(X^2)^3}{(X^1)^2}\}$. Note that for $\Delta = 1/8$ and $M_1 = 20$ (not reported above), the price has not converged yet (10.83).

The parametric method performs very well, because the optimal volatilities belong to the parametized set we input. Indeed, we know that the true price reads $P(t, x^1, x^2) = x^1 u(t, x^2/x^1)$ with u the value of a call spread on X^2/X^1 in the one-dimensional UVM with

$$I = \left[\sum_{\alpha=1}^{2} (\underline{\sigma}^\alpha)^2 - 2\rho\underline{\sigma}^1\underline{\sigma}^2, \sum_{\alpha=1}^{2} (\overline{\sigma}^\alpha)^2 - 2\rho\overline{\sigma}^1\overline{\sigma}^2\right],$$

because we have chosen $\rho \leq 0$, see (18). Hence the optimal volatilities are $\underline{\sigma}^1, \underline{\sigma}^2$ when $\partial_x^2 u(t, x^2/x^1) < 0$, and $\overline{\sigma}^1, \overline{\sigma}^2$ when $\partial_x^2 u(t, x^2/x^1) \geq 0$. Now, $\mathrm{sign}(\partial_x^2 u(t, x^2/x^1)) = \mathrm{sign}(c_t - x^2/x^1)$ for some constant c_t. This shows that we could even have enforced $\theta_i^1 = \theta_i^2$ in our parameterization of the optimal volatilities, without deteriorating the price.

6.3. **Options with two underlyings and uncertainty on correlation.** When the correlation ρ is uncertain, one cannot reduce the two-dimensional Black-Scholes-Barenblatt PDE to a simple one-dimensional PDE as in Subsection 6.2, even for projective payoffs. To test the accuracy of our two algorithms when correlation is uncertain, we price the outperformer spread option with the same parameters as above, except that the time-t correlation ρ_t is no more constant but is assumed to stay within $[\underline{\rho}, \overline{\rho}]$. We pick $\underline{\rho} = -0.5$, $\overline{\rho} = 0.5$ and $\hat{\rho} = 0$. The (two-dimensional)

PDE price is 12.83. When ρ_t was assumed to be constantly equal to -0.5, the price was 11.41. **First parameterization method:** We choose $\theta_i = (\theta_i^1, \theta_i^2, \theta_i^3) \in \mathbb{R}^3$, $\sigma^1_{t_i}(X;\theta_i) = \overline{\sigma}^1$ if $\theta_i^1 - \ln(X^2/X^1) \geq 0$, $= \underline{\sigma}^1$ otherwise, $\sigma^2_{t_i}(X;\theta_i) = \overline{\sigma}^2$ if $\theta_i^2 - \ln(X^2/X^1) \geq 0$, $= \underline{\sigma}^2$ otherwise, and $\rho_{t_i}(X;\theta_i) = \overline{\rho}$ if $-\theta_i^3 + \ln(X^2/X^1) \geq 0$, $= \underline{\rho}$ otherwise. **Second BSDE method:** We use the basis functions $= \{1, X^1, X^2, \frac{(X^2)^2}{X^1}, \frac{(X^2)^3}{(X^1)^2}\}$ (see Table 12).

Δ	Param	M_1	12	13	14	15	16	17
1/2	12.50	Price	12.37	12.40	12.41	12.41	12.39	12.38
1/4	**12.67**	Price	11.58	11.79	12.38	**12.44**	12.44	12.41
1/8	12.67	Price	9.51	10.27	10.08	10.87	11.34	11.43

Table 12. Outperformer spread option valued using the parametric approach (column 2) and the BSDE approach (columns 4 to 9). The correlation is uncertain, with $\underline{\rho} = -0.5$ and $\overline{\rho} = 0.5$. The true price is $\mathcal{C}_{\text{PDE}} = 12.83$. The Black-Scholes price with mid-volatility and mid-correlation is 9.24. Basis functions $= \{1, X^1, X^2, \frac{(X^2)^2}{X^1}, \frac{(X^2)^3}{(X^1)^2}\}$. Note that for $\Delta = 1/8$ and $M_1 = 20$ (not reported above), we get **12.54**.

It is noteworthy that even with our simple parameterization for the optimal volatilities and correlation, the parametric method performs very well.

As a conclusion of our numerical tests, we generally observe that the parametric method performs better than the BSDE method, provided we input a good parameterization. Indeed, if the optimal covariance matrix belongs to the parametized set, then the parametric method will naturally find it. However, if a wrong parameterization is supplied, the parametric method will produce bad results. In order to get more precise results with the BSDE method, we should have used a finer regression procedure, e.g., better regressors and/or more regressors and/or non-parametric regressions in dimension greater than or equal to two. A perfect regression procedure at all dates t_i, on the space spanned by all $\mathcal{F}_{t_i}$-measurable variables, would lead to the exact gamma and price, up to the discretization error. In this paper, we intended to prove the efficiency of the BSDE method, without optimizing its regression step.

7. Conclusion

In this paper, we have provided two efficient, accurate, easily-implementable Monte-Carlo approaches to the pricing of derivatives in the UVM. In the first method, which consists in a series of backward low-dimensional maximizations, the main ingredient is a parameterization of the optimal covariance matrix. The second method uses a recent connection between fully non-linear second-order PDE and 2-BSDE. This second pricing methodology represents, as far as the authors know, the first application of backward stochastic differential equations to pricing models commonly used by practitioners. We hope that this work will initiate research on BSDEs in the community of practitioners.

Besides, these two algorithms can be combined efficiently. We may use the BSDE algorithm to choose an efficient parameterization for the parametric approach.

As a word of caution, we have illustrated in our numerical experiments that results depend greatly on the parameterization/choice of regressors which may require numerical experimentations and good understanding of the financial derivative under consideration. This is a common feature in the pricing of American options. An upper bound *à la* Rogers (2002) should help accessing the accuracy of the lower bound. Finally, as indicated in Remark 2.3, a similar methodology could be used to include transaction costs in general diffusion models. These issues will be the subject of future work.

Acknowledgements. The authors wish to thank the members of the Global Markets Quantitative Research Group at Société Générale, and particularly Lorenzo Bergomi and Benoît Humez, for fruitful discussions and for their comments on preliminary versions of the paper. They are also grateful to two anonymous referees whose comments helped improve the paper.

References

[1] Andersen, L. (1999): *A Simple Approach to the pricing of Bermudan Swaptions in the Multi-Factor Libor Market Model* : Journal of Computational Finance, 3:5–32.
[2] Avellaneda, M., Levy, A., Paras, A. (1995): *Pricing and hedging derivative securities in markets with uncertain volatilities* : Journal of Applied Finance, Vol 1.
[3] Avellaneda, M., Buff, R. (1999): *Combinatorial implications of non-linear uncertain volatility models: the case of barrier options*, Applied Mathematical Finance, Vol 1:1–18.
[4] Avellaneda, M., Kampen, J. (2003) : *On Parabolic Equations with Gauge Function Term and Applications to the Multidimensional Leland Equation*, Applied Mathematical Finance, Volume 10, Issue 3 September 2003, pages 215 - 228.
[5] Bouchard, B., Touzi, N. (2004): *Discrete-time approximation and Monte Carlo simulation of backward stochastic differential equations*, Stochastic Processes and their Applications, 111, 175–206.
[6] Broadie, M., Glasserman, P. (1997): *A stochastic mesh method for pricing high-dimensional American options*, PaineWebber Working Papers in Money, Economics and Finance #PW9804, Columbia Business School, New York.
[7] Buff, R. (2002): *Uncertain Volatility Models - Theory and Application*, Springer Finance Lecture Notes.
[8] Cheridito, P., Soner, M., Touzi, N., Victoir, N. (2007): *Second Order Backward Stochastic Differential Equations and Fully Non-Linear Parabolic PDEs*, Communications on Pure and Applied Mathematics, Volume 60, Issue 7, 1081–1110.
[9] El Karoui, N., Peng, S., Quenez, M.C. (1997): *Backward Stochastic Differential Equations in Finance*, Mathematical Finance, Vol. 47, No. 1, 1–77.
[10] Fahim, A., Touzi, N., Warin, X. (2009): *A Probabilistic Numerical Method for Fully Nonlinear Parabolic PDEs*, preprint.
[11] Flannery, B.P., Press, W.H., Teukolsky, S.A., Vetterling, W.T. (2007): *Numerical Recipes, The Art of Scientific Computing*, Third Edition, Cambridge University Press.
[12] Forsyth, P.A, Vetzal, K.R. (2001): *Implicit Solution of Uncertain Volatility/Transaction Cost Option Pricing with Discretely Observed Barriers*, Appl. Num. Math. 36, 427–445.
[13] Glasserman, P. (2003): *Monte Carlo Methods in Financial Engineering*, Springer.
[14] Gobet, E., Lemor, J.P., Warin, X. (2005): *A regression-based Monte-Carlo method to solve backward stochastic differential equations*, Annals of Applied Probability, Vol.15(3), pp.2172–2002.
[15] Henry-Labordère, A. (1995): *Cours de Recherche Opérationnelle*, Presses Ecole des Ponts.
[16] Leblanc, M., Martini, C. (2000): *Unbounded volatility in the uncertain volatility model*, INRIA Research Report 4065.
[17] Leland, H. E. (1985): *Option pricing and replication with transaction costs*, The Journal of Finance, 40, 283-1301.
[18] Longstaff, F.A., Schwartz, E.S. (2001): *Valuing American Options by Simulation: A Simple Least-Squares Approach*, The Review of Financial Studies, 14(1), 113–147.
[19] Lyons, T.J. (1995): *Uncertain volatility and the risk-free synthesis of derivatives*, Applied Mathematical Finance, 2(2):117–133.
[20] Pardoux, E., Peng, S. (1990): *Adapted Solution of a Backward Stochastic Differential Equation*, Systems Control Lett., 14, 55-61.
[21] Pardoux, E., Peng, S. (1992): *Backward Stochastic Differential Equations and Quasilinear Parabolic Partial Differential Equations*, Lecture Notes in CIS, vol. 176. Springer-Verlag, 200–217.

[22] Ma, J. and Yong, J. (2000): *Forward-Backward Stochastic Differential Equations and Their Applications*, Lecture Notes in Mathematics 1702, Springer.

[23] Ma, J., Protter, P., San Martin, J. and Torres, S. (2002) : *Numerical Method for Backward Stochastic Differential Equations*, Annals of Applied Probability 12, pp. 302316.

[24] Martini, C. (1999): *The Uncertain Volatility Model and American Options*, INRIA Research Report 3697.

[25] Meyer, G. (2006): *The Black-Scholes Barenblatt Equation for Options with Uncertain Volatility and its Application to Static Hadging*, International Journal of Theoretical and Applied Finance, vol 9(5):673–703.

[26] Pham, H. (2005): *On some recent aspects of stochastic control and their applications*, Probability Surveys, vol 2, 506–549.

[27] Pham, H (2009): *Continuous-time stochastic control and applications with financial applications*, Series: Stochastic Modelling and Applied Probability , Vol. 61, Springer.

[28] Pooley, D.M., Forsyth, P.A., Vetzal K.R. (2003): *Numerical convergence properties of option pricing PDEs with uncertain volatility*, IMA Journal of Numerical Analysis, 23(2):241–267.

[29] Pooley, D.M., Forsyth, P.A., Vetzal K.R. (2003): *Two Factor Option Pricing with Uncertain Volatility*, Lecture Notes in Computer Science, vol 2669, Springer.

[30] Rogers, L.C.G. (2002): *Monte-Carlo valuation of American Options*, Mathematical Finance, 12:271–286.

[31] Smith, A.T. (2002): *American options under uncertain volatility*, Applied Mathematical Finance, 9(2):123–141.

[32] Tsitsiklis, J.N., Van Roy, B. (2001): *Regression Methods for Pricing Complex American-Style Options*, IEEE Transactions on Neural Networks, 12(4):694–703.

[33] Von Petersdorff, T. , Schwab, C. (2004): *Numerical Solution of Parabolic Equations in High Dimensions*, Modélisation mathématique et analyse numérique, vol. 38, no1, pp. 93-127 .

[34] Zhang, K., Wang, S. (2009): *A computational scheme for uncertain volatility in option pricing*, to appear in Applied Numerical Mathematics.

Société Générale, Global Markets Quantitative Research Group, 17 cours Valmy, 92987 Paris La Défense cedex, France

E-mail address: julien.guyon@sgcib.com

E-mail address: pierre.henry-labordere@sgcib.com