

American Options. Pricing and Volatility Calibration

Yves Achdou*

*UFR Mathématiques, Université Paris 7, Case 7012, 75251 PARIS Cedex 05, France and
Laboratoire Jacques-Louis Lions, Université Paris 6.
Email: achdou@math.jussieu.fr*

Olivier Pironneau

*Laboratoire Jacques-Louis Lions, Université Pierre et Marie Curie, Boîte courrier 187, 75252
Paris Cedex 05. France.
Email: pironneau@ann.jussieu.fr*

In this paper, we discuss calibration of the local volatility with American put options. The calibration method uses a least-square formulation and a descent algorithm. Pricing is done by solving parabolic variational inequalities, for which solution procedures by active set methods are discussed.

The strategy consists in computing the optimality conditions and the descent direction needed by the optimization loop. This approach has been implemented both at the continuous and discrete levels. It requires a careful analysis of the underlying variational inequalities and of their discrete counterparts. In the numerical example presented here (American options on the FTSE index), the squared volatility is parameterized by a bicubic spline.

Keywords: American options, variational inequalities, pricing, calibration of volatility, inverse problems, optimality conditions, finite element methods

*Control Systems: Theory, Numerics and Applications
30 March - 1 April 2005
Rome*

*Speaker.

1. Introduction

The Black-Scholes model involves a risky asset and a risk-free asset whose price at time t is $S_t^0 = e^{rt}$, where r is the interest rate; it assumes that the price of the risky asset is a solution to the following stochastic differential equation,

$$dS_t = S_t(\mu dt + \sigma_t dB_t), \quad (1.1)$$

where W_t is a standard Brownian motion on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Here σ_t is a positive number, called the *volatility*. In what follows, it will be convenient to work with the squared volatility $\eta_t = \sigma_t^2$.

A *European option on the underlying risky asset*, is a contract which permits to its owner a benefit $P_\circ(S_T)$ at time T . The function P_\circ is called the payoff function and the date T is the maturity. With the Black-Scholes assumptions, it is possible to prove that the option's price at time t is given by

$$P_t = P_e(S_t, t) \equiv \mathbb{E}^*(e^{-r(T-t)} P_\circ(S_T) | F_t), \quad (1.2)$$

where the expectation \mathbb{E}^* is taken with respect to the so-called risk-neutral probability \mathbb{P}^* (equivalent to \mathbb{P} and under which $dS_t = S_t(r dt + \sqrt{\eta_t} dW_t)$, W_t being a standard Brownian motion under \mathbb{P}^* and F_t being the natural filtration of W_t). It can be seen that the pricing function P_e solves the parabolic partial differential equation:

$$\frac{\partial P_e}{\partial t} + \frac{\eta S^2}{2} \frac{\partial^2 P_e}{\partial S^2} + rS \frac{\partial P_e}{\partial S} - rP_e = 0. \quad (1.3)$$

In contrast with European options, American options can be exercised any time before maturity: *An American vanilla call (resp. put) option is a contract giving its owner the right to buy (resp. sell) a share of a specific common stock at a fixed price K before a certain date T .* More generally, for a payoff function P_\circ , the American option with payoff P_\circ and maturity T can be exercised at any $t < T$, yielding the payoff $P_\circ(S_t)$.

Using the notion of strategy with consumption, the Black-Scholes model leads to the following formula for pricing an American option with payoff P_\circ : under the risk neutral probability,

$$P_t = P(S_t, t) \equiv \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}^* \left(e^{-r(\tau-t)} P_\circ(S_\tau) \middle| F_t \right), \quad (1.4)$$

where $\mathcal{T}_{t,T}$ denotes the set of stopping times in $[t, T]$ (see [16] for the proof of this formula). It can be seen that for an American vanilla call on a non dividend paying stock, the formula (1.4) coincides with (1.2), so American and European vanilla calls have the same price. This means that an American vanilla call should not be exercised before maturity.

It can be shown that the price P of the American put of strike K and maturity T is given as a solution to

$$\begin{aligned} \frac{\partial P}{\partial t} + \frac{\eta S^2}{2} \frac{\partial^2 P}{\partial S^2} + rS \frac{\partial P}{\partial S} - rP &\leq 0, & P(S, t) &\geq P_\circ(S), & t \in [0, T], & S > 0, \\ \left(\frac{\partial P}{\partial t} + \frac{\eta S^2}{2} \frac{\partial^2 P}{\partial S^2} + rS \frac{\partial P}{\partial S} - rP \right) (P - P_\circ) &= 0 & t \in [0, T], & S > 0, \\ P(S, t = T) &= P_\circ(S), & S > 0, \end{aligned} \quad (1.5)$$

where

$$P_{\circ}(S) = (K - S)_{+}.$$

The volatility is the difficult parameter of the Black-Scholes model. It is convenient to take it to be constant but then the computed options' prices do not match the market prices. Conversely, taking a family of options available on the market and inverting for each of them the Black-Scholes formula does not yield a constant volatility: for each option, one obtains a different *implied* volatility, and the implied volatility is often a convex function of the strike K , which is known in finance as the *smile effect*.

There are essentially three ways to improve on the Black-Scholes model with a constant volatility:

- *Use a local volatility*, i.e. assume that the volatility is a function of time and of the stock price. Then one has to *calibrate the volatility* from the market data, i.e. to find a volatility function which permits to recover the prices of the options available on the market.
- *assume that the volatility is itself a stochastic process*, see for example [10, 8].
- *generalize the Black-Scholes model by assuming that the spot price is for example a Lévy process*, see [7] and references therein.

In this paper, we deal with the first approach: the calibration problem consists in finding $\eta(S, t)$ from the observations of

- the spot price S_{\circ} today,
- the prices $(\bar{P}_i)_{i \in I}$ of a family of options with different maturities and different strikes $(T_i, K_i)_{i \in I}$.

A way to solve the calibration problem is to

$$\text{find } \eta \in \mathcal{H} \text{ minimizing } J(\eta) + J_R(\eta), \quad J(\eta) = \sum_{i \in I} |P(S_{\circ}, 0, K_i, T_i) - \bar{P}_i|^2, \quad (1.6)$$

where \mathcal{H} is a suitable closed subset of a possibly infinite dimensional function space, J_R is a suitable Tychonoff regularization functional, and where $P(S_{\circ}, 0, K_i, T_i)$ is the price of the option with strike K_i and maturity T_i computed with the local volatility η .

Note that if \mathcal{H} is not carefully chosen and if $J_R = 0$, then the problem is unstable too. Note also that computing $J(\eta)$ requires solving $\#(I)$ different problems of the type (1.3) for European options or (1.5) for American options, so this approach is expensive.

There has been a number of valuable studies on the calibration of volatility with European options of which it is difficult to make a complete account here. In Lagnado and Osher [14, 15] a least square method is used, the volatility is discretized by splines using matlab and the computation of the gradient of J with respect to η is done either by numerical differences or by Adol-C. In Achdou et al [3], calibration with European option is made easier by using Dupire's equation. An alternative to least squares is the pioneering method by Avellaneda et al [6] based on the maximization of an entropy function via dynamic programming. Up to our knowledge, calibration with American options has not been very much discussed yet in the mathematical finance literature.

The paper is divided in five parts: we first give theoretical results on the variational inequality for pricing American options and on the free boundary which delimitates the region of exercise. The second part of the paper is devoted to the finite element method for pricing American options. We show in particular that under some assumptions, there is a free boundary in the discrete problem too, and we discuss two algorithms for computing the price of the American option, both based on active sets strategies. The third part is devoted to volatility calibration with American options: we consider problem (1.6) with

$$J_R(\eta) = \int_0^T \int_0^{\bar{S}} a(S\partial_S\eta)^2 + b(\partial_t\eta)^2 + c(S\partial_{tS}^2\eta)^2 + d(S^2\partial_{SS}^2\eta)^2 + e(\eta - \eta_g)^2.$$

The minimization problem is constrained, and optimality conditions are found for (1.6); it is also proved that differentiability with respect to η holds provided a strict complementarity condition is satisfied. We give an example of calibration with American options on the FTSE index, where the squared volatility is discretized with bicubic splines. The results contained in the four first parts are proved in [1, 5, 4]. In the last part of the paper, we investigate the numerical method used for pricing in the case when an American option on a portfolio containing two assets. At a given time, the free boundary is a curve. Two key ingredients for option pricing are discussed: mesh adaptivity and the algorithm for solving the discrete variational inequality.

2. The variational inequality and the free boundary

All the proofs of the results below can be found in [1].

Calling t the time to maturity, the problem becomes

$$\begin{aligned} \frac{\partial P}{\partial t} - \frac{\eta(S,t)S^2}{2} \frac{\partial^2 P}{\partial S^2} - rS \frac{\partial P}{\partial S} + rP &\geq 0, \quad P \geq P_o, \\ \left(\frac{\partial P}{\partial t} - \frac{\eta(S,t)S^2}{2} \frac{\partial^2 P}{\partial S^2} - rS \frac{\partial P}{\partial S} + rP \right) (P - P_o) &= 0, \end{aligned} \quad (2.1)$$

with Cauchy data

$$P|_{t=0} = P_o. \quad (2.2)$$

We focus on the case of a vanilla put, i.e. the payoff function is $P_o(S) = (K - S)_+$.

To write the variational formulation of (2.1) (2.2), we need to use the Sobolev space

$$V = \{v \in L^2(\mathbb{R}_+) : S \frac{dv}{dS} \in L^2(\mathbb{R}_+)\}, \quad (2.3)$$

and we call \mathcal{X} the subset of V :

$$\mathcal{X} = \{v \in V, v \geq P_o \text{ in } \mathbb{R}_+\}. \quad (2.4)$$

Since the function of V are continuous, the inequality in (2.4) has a pointwise meaning. The set \mathcal{X} is a closed and convex subset of V , because convergence in V implies pointwise convergence. We

introduce the bilinear form a_t :

$$\begin{aligned} a_t(v, w) &= \int_{\mathbb{R}_+} \frac{S^2 \eta(S, t)}{2} \frac{\partial v}{\partial S} \frac{\partial w}{\partial S} dS \\ &+ \int_{\mathbb{R}_+} \left(-r(t) + \eta(S, t) + \frac{S}{2} \frac{\partial \eta}{\partial S}(S, t) \right) S \frac{\partial v}{\partial S} w dS \\ &+ r \int_{\mathbb{R}_+} vw dS \end{aligned} \quad (2.5)$$

We make the assumptions: there exist two positive constants, $\underline{\eta}$ and $\bar{\eta}$ such that for all $t \in [0, T]$ and all $S \in \mathbb{R}_+$,

$$0 < \underline{\eta} \leq \eta(S, t) \leq \bar{\eta}. \quad (2.6)$$

There exists a positive constant C_η such that for all $t \in [0, T]$ and all $S \in \mathbb{R}_+$,

$$\left| S \frac{\partial \eta}{\partial S}(S, t) \right| \leq C_\eta. \quad (2.7)$$

These imply that the bilinear form a_t is continuous on V uniformly in t , and Gårding's inequality : for a non negative constant λ depending only on $\bar{\eta}$, $\underline{\eta}$ and C_η ,

$$a_t(v, v) \geq \frac{\underline{\eta}}{4} |v|_V^2 - \lambda \|v\|_{L^2(\mathbb{R}_+)}^2, \quad \forall v \in V. \quad (2.8)$$

The weak form of (2.1) is to

find $P \in C^0([0, T]; L^2(\mathbb{R}_+)) \cap L^2(0, T; \mathcal{X})$ such that $\frac{\partial P}{\partial t} \in L^2(0, T; V')$, satisfying $P|_{t=0} = P_\circ$, and

$$\forall v \in \mathcal{X}, \quad \left(\frac{\partial P}{\partial t}(t), v - P(t) \right) + a_t(P(t), v - P(t)) \geq 0. \quad (2.9)$$

Theorem 1. *With η satisfying assumptions (2.6) and (2.7), the problem (2.9) has a unique solution P which belongs to $C^0([0, T] \times [0, +\infty))$ with $P(0, t) = K, \forall t \in [0, T]$, and is such that $S \frac{\partial P}{\partial S}, \frac{\partial P}{\partial S} \in L^2(0, T; V)$, $S \frac{\partial P}{\partial S} \in C^0([0, T]; L^2(\mathbb{R}_+))$ and $\frac{\partial P}{\partial t} \in L^2(0, T; L^2(\mathbb{R}_+))$.*

The function P is also greater than or equal to P_e , the price of the vanilla European put.

The quantities $\|P\|_{L^2(0, T; V)}$, $\|P\|_{L^\infty(0, T; L^2(\mathbb{R}_+))}$, $\|S \frac{\partial P}{\partial S}\|_{L^2(0, T; V)}$, $\|\frac{\partial P}{\partial S}\|_{L^2(0, T; V)}$,

$\|S \frac{\partial P}{\partial S}\|_{L^\infty(0, T; L^2(\mathbb{R}_+))}$, $\|\frac{\partial P}{\partial t}\|_{L^2(0, T; L^2(\mathbb{R}_+))}$, are bounded by constants depending only on K , $\bar{\eta}$, $\underline{\eta}$ and C_η .

We have that

$$-1 \leq \frac{\partial P}{\partial S} \leq 0, \quad \forall t \in (0, T], \text{ a.a. } S > 0. \quad (2.10)$$

There exists a function $\gamma : (0, T] \rightarrow [0, K)$, such that $\forall t \in (0, T)$, $\{S \text{ s.t. } P(S, t) = P_\circ(S)\} = [0, \gamma(t)]$. The function γ is upper semi-continuous, right continuous in $[0, T)$, and, for each $t \in (0, T)$, γ has a left-limit at t .

Calling μ the function $\mu = \frac{\partial P}{\partial t} + A_t P$, where A_t is the linear operator: $V \rightarrow V'$; for all $v, w \in V$, $A_t v = -\frac{\eta(S, t) S^2}{2} \frac{\partial^2 v}{\partial S^2} - r S \frac{\partial v}{\partial S} + r v$, we have

$$\mu = r K 1_{\{P=P_\circ\}}. \quad (2.11)$$

In other words, a.e., one of the two conditions $P = P_\circ$ and $\mu = 0$ is not satisfied: there is strict complementarity in (2.1).

Finally, there exists $\gamma_0 > 0$ depending only on $\bar{\eta}$ and K such that

$$\gamma(t) \geq \gamma_0, \quad \forall t \in [0, T]. \quad (2.12)$$

3. Pricing American options with a finite element method

3.1 A finite element method

All the proof of the results below are given in [5].

We localize the problem on $(0, \bar{S})$, so V becomes

$$V = \{v \in L^2((0, \bar{S})); S \frac{\partial v}{\partial S} \in L^2((0, \bar{S})); v(\bar{S}) = 0\}$$

(where \bar{S} is large enough so that $P_\circ(\bar{S}) = 0$), and $\mathcal{X} = \{v \in V, v \geq P_\circ\}$. The variational inequality is (2.9) with new meanings for V , \mathcal{X} , and a_t .

Moreover, if $\gamma_0 \in (0, K)$ as in (2.12) is known, one can focus on the smaller interval $[\underline{S}, \bar{S}]$, with $0 \leq \underline{S} < \gamma_0$ and obtain the equivalent weak formulation:

$$\text{find } P \in L^2((0, T, \mathcal{X}) \cap C^0([0, T]; L^2(\Omega))), \text{ with } \frac{\partial P}{\partial t} \in L^2(0, T; V')$$

such that $P(t=0) = P_\circ$ and (2.9) for all $v \in \mathcal{X}$, with the new definition of the closed set \mathcal{X} :

$$\mathcal{X} = \{v \in V, v \geq P_\circ \text{ in } (0, \bar{S}], P = P_\circ \text{ in } (0, \underline{S}]\}. \quad (3.1)$$

We introduce a partition of the interval $[0, T]$ into subintervals $[t_{n-1}, t_n]$, $1 \leq n \leq N$, with $\Delta t_i = t_i - t_{i-1}$, $\Delta t = \max_i \Delta t_i$ and a partition of the interval $[0, \bar{S}]$ into subintervals $\omega_i = [S_{i-1}, S_i]$, $1 \leq i \leq N_h + 1$, such that $0 = S_0 < S_1 < \dots < S_{N_h} < S_{N_h+1} = \bar{S}$. The size of the interval ω_i is called h_i and we set $h = \max_{i=1, \dots, N_h+1} h_i$. The mesh \mathcal{T}_h of $[0, \bar{S}]$ is the set $\{\omega_1, \dots, \omega_{N_h+1}\}$. In what follows, we will assume that both the strike K and the real number \underline{S} coincide with nodes of \mathcal{T}_h : there exist $\alpha < \kappa$, $0 \leq \alpha < \kappa < N_h + 1$ such that $S_\kappa = K$ and $S_{\alpha-1} = \underline{S}$. We define the discrete space V_h by

$$V_h = \{v_h \in V, \forall \omega \in \mathcal{T}_h, v_h|_\omega \in \mathcal{P}_1(\omega)\}, \quad (3.2)$$

where $\mathcal{P}_1(\omega)$ is the space of linear functions on ω .

Since K is a node of \mathcal{T}_h , $P_\circ \in V_h$, and since \underline{S} is also a node of \mathcal{T}_h , we can define the closed subset \mathcal{X}_h of V_h by

$$\begin{aligned} \mathcal{X}_h &= \{v \in V_h, v \geq P_\circ \text{ in } [0, \bar{S}], \quad v = P_\circ \text{ in } [0, \underline{S}]\} \\ &= \{v \in V_h, v(S_i) \geq P_\circ(S_i), \quad i = 0, \dots, N_h + 1, v(S_i) = P_\circ(S_i), \quad i < \alpha\}. \end{aligned} \quad (3.3)$$

The discrete problem arising from an implicit Euler scheme is:

find $(P^n)_{0 \leq n \leq N} \in \mathcal{X}_h$ satisfying

$$P^0 = P_\circ, \quad (3.4)$$

and for all n , $1 \leq n \leq N$,

$$\forall v \in \mathcal{X}_h, \quad (P^n - P^{n-1}, v - P^n) + \Delta t_n a_{t_n}(P^n, v - P^n) \geq 0. \quad (3.5)$$

Consider λ such that Gårding's inequality (2.8) holds, and take $\Delta t < \frac{1}{\lambda}$, there exists a unique P^n satisfying (3.5).

Let $(w^i)_{i=0, \dots, N_h}$ be the nodal basis of V_h , and let \mathbf{M} and \mathbf{A}^m in $\mathbb{R}^{(N_h+1) \times (N_h+1)}$ be the mass and stiffness matrices defined by

$$\mathbf{M}_{i,j} = (w^i, w^j), \mathbf{A}_{i,j}^m = a_{t_m}(w^j, w^i), \quad 0 \leq i, j \leq N_h.$$

Calling

$$U^n = (P^n(S_0), \dots, P^n(S_{N_h}))^T \text{ and } U^0 = (P_o(S_0), \dots, P_o(S_{N_h}))^T,$$

(3.5) is equivalent to

$$\begin{aligned} (\mathbf{M}(U^n - U^{n-1}) + \Delta t_n \mathbf{A}^n U^n)_i &\geq 0, & \text{for } i \geq \alpha, \\ U_i^n &= U_i^0 & \text{for } i < \alpha, \\ U^n &\geq U^0, \\ (U^n - U^0)^T (\mathbf{M}(U^n - U^{n-1}) + \Delta t_n \mathbf{A}^n U^n) &= 0. \end{aligned} \quad (3.6)$$

We call \mathbf{M}_α , respectively \mathbf{A}_α^n , the block of \mathbf{M} , respectively \mathbf{A}^n , corresponding to $\alpha \leq i, j \leq N_h$.

3.2 The discrete exercise boundary

One may ask if there is a well defined exercise boundary $t \rightarrow \gamma_h(t)$ also in the discrete problem. A positive answer has been given by Jaillet et al [13] in the case of a constant volatility, an implicit Euler scheme and a uniform mesh in the logarithmic variable. The main argument of the proof lies in the fact that the solution to the discrete problem is non decreasing with respect to the variable t . With a local volatility, this may not hold (see the numerical example below). The result of Jaillet et al has been completed for a local volatility in [5], in the special case when the mesh is uniform in the variable S : here too, the discrete problem has a free boundary. The proof does not rely any longer on the monotonic character of the discrete solution with respect to t but on the discrete analogue of the bounds (2.10), i.e. $-1 \leq \frac{\partial P}{\partial S} \leq 0$. This is proved by studying a suitable penalized problem (which is the discrete version of a semilinear parabolic equation with a non decreasing and convex non linearity) and by using a discrete maximum principle on the partial derivative with respect to S (for this reason, a uniform mesh is needed). We can summarize this by

Theorem 2. *Let η verify (2.6) and (2.7), and choose $\Delta t < \frac{1}{2\lambda}$, with λ given in (2.8). Assume that the grid \mathcal{T}_h is uniform and that $\underline{S} > 0$. Assume also that the parameters h and $\frac{h^2}{\Delta t}$ are small enough so that the matrices \mathbf{A}_α^n and $\mathbf{M}_\alpha + \Delta t_n \mathbf{A}_\alpha^n$ are tridiagonal irreducible M-matrices for all n , $1 \leq n \leq N$. There exist N real numbers γ_h^n , $1 \leq n \leq N$, such that*

$$\begin{aligned} \underline{S} &\leq \gamma_h^n < K, \\ \gamma_h^n &\text{ is a node of } \mathcal{T}_h, \\ \forall i, 0 \leq i \leq N_h, \quad P^n(S_i) = P_o(x_i) &\Leftrightarrow S_i \leq \gamma_h^n. \end{aligned} \quad (3.7)$$

We believe that this may be extended to somewhat more general meshes.

3.3 A front-tracking algorithm

Here, we propose an algorithm for computing the solution of (3.5) assuming that the free boundary is the graph of a function. In our experience, this algorithm, based on tracking the free boundary, is more robust (and slightly more expensive) than the Brennan and Schwartz algorithm (see [13]). Since the free boundary is the graph of a function, the idea is to look for γ_h^n by

- Start from $\gamma_h^n = \gamma_h^{n-1}$,
- solve the discrete problem corresponding to

$$\begin{aligned} \frac{P^n - P^{n-1}}{\Delta t_n} - \frac{\eta(S, t_n) S^2}{2} \frac{\partial^2 P^n}{\partial S^2} - rS \frac{\partial P^n}{\partial S} + rP^n = 0 \text{ for } \gamma_h^n < S < \bar{S}, \\ P^n = P_o \text{ for } 0 \leq S \leq \gamma_h^n, \end{aligned}$$

and $P^n(\bar{S}) = 0$,

- if P^n satisfies (3.5), stop else shift the point γ_h^n to the next node on the mesh left/right according to which constraint is violated by P^n .

With the notations introduced above, the algorithm for computing P_h^n is as follows:

Algorithm

choose k such that $\gamma_h^{n-1} = S_k$; set found=false;

while(not found)

.. solve

$$\begin{aligned} (\mathbf{M}(U^n - U^{n-1}) + \Delta t_n \mathbf{A}^n U^n)_i = 0, \quad \text{for } i \geq k, \\ U_i^n = U_i^0 \quad \text{for } i < k. \end{aligned} \tag{3.8}$$

.. **if** $((U^n - U^0)_{k+1} < 0)$

.. found=false; $k = k + 1$;

.. **else**

.. compute $a = (\mathbf{M}(U^n - U^{n-1}) + \Delta t_n \mathbf{A}^n U^n)_{k-1}$;

.. **if** $(a < 0)$

.. found=false; $k = k - 1$;

.. **else** found=true

In our tests, we have computed the average (over the time steps) number of iterations to obtain the position of the free boundary: it was found that (with a rather fine time-mesh), this number is smaller than 2.

3.4 A regularized active set strategy

The algorithm above is not easy to generalize in higher dimensions. For an algorithm based on active sets and generalizable in any dimension, we have to regularize first the problem. Following [12], we first go back to the semi-discrete problem: find $P^n \in \mathcal{X}$ such that

$$\forall v \in \mathcal{X}, \quad (P^n - P^{n-1}, v - P^n) + \Delta t_n a_{t_n}(P^n, v - P^n) \geq 0.$$

For any positive constant c , this is equivalent to finding $P^n \in V$ and a Lagrange multiplier $\mu \in V'$ such that

$$\begin{aligned} \forall v \in V, \quad & \left(\frac{P^n - P^{n-1}}{\Delta t_n}, v \right) + a_{t_n}(P^n, v) - \langle \mu, v \rangle = 0, \\ & \mu = \max(0, \mu - c(P^n - P^0)). \end{aligned} \quad (3.9)$$

When using an iterative method for solving (3.9), i.e. when constructing a sequence $(P^{n,m}, \mu^m)$ for approximating (P^n, μ) , the Lagrange multiplier μ^m may not be a function if the derivative of the $P^{n,m}$ jumps, whereas μ is generally a function. Therefore, a dual method (i.e. an iterative method for computing μ) may be difficult to use. As a remedy, K.Ito and K.Kunisch [12] consider a one parameter family of regularized problems based on smoothing the equation for μ by

$$\mu = \alpha \max(0, \mu - c(P^n - P^0)), \quad (3.10)$$

for $0 < \alpha < 1$, which is equivalent to

$$\mu = \max(0, -\chi(P^n - P^0)), \quad (3.11)$$

for $\chi = c\alpha/(1 - \alpha) \in (0, +\infty)$. We may consider a generalized version of (3.11):

$$\mu = \max(0, \bar{\mu} - \chi(P^n - P^0)), \quad (3.12)$$

where $\bar{\mu}$ is a fixed function. This turns to be useful when the complementarity condition is not strict.

It is now possible to study the fully regularized problem

$$\begin{aligned} \forall v \in V, \quad & \left(\frac{P^n - P^{n-1}}{\Delta t_n}, v \right) + a_{t_n}(P^n, v) - \langle \mu, v \rangle = 0, \\ & \mu = \max(0, \bar{\mu} - \chi(P^n - P^0)), \end{aligned} \quad (3.13)$$

and prove that it has a unique solution, with μ a square integrable function. A primal-dual active set algorithm for solving (3.13) is

Primal-dual active set algorithm

1. Choose $P^{n,0}$, set $k = 0$

2. Loop

(a) Set

$$\mathcal{A}^{-,k+1} = \{S : \bar{\mu}^k(S) - \chi(P^{n,k}(S) - P^0(S)) > 0\}$$

$$\text{and } \mathcal{A}^{+,k+1} = (0, \bar{S}) \setminus \mathcal{A}^{-,k+1}.$$

(b) Solve for $P^{n,k+1} \in V: \forall v \in V,$

$$\left(\frac{P^{n,k+1} - P^{n-1}}{\Delta t_n}, v \right) + a_{t_n}(P^{n,k+1}, v) - (\bar{\mu} - \chi(P^{n,k+1} - P^0), 1_{\mathcal{A}^{-,k+1}} v) = 0. \quad (3.14)$$

(c) Set

$$\mu^{k+1} = \begin{cases} 0 & \text{on } \mathcal{A}^{+,k+1}, \\ \bar{\mu} - \chi(P^{n,k+1} - P^0) & \text{on } \mathcal{A}^{-,k+1} \end{cases} \quad (3.15)$$

(d) Set $k = k + 1$.

Calling A_n the operator from V to V' : $\langle A_n v, w \rangle = \left(\frac{v}{\Delta t_n}, w \right) + a_{t_n}(v, w)$ and $F : V \times L^2(\mathbb{R}_+) \rightarrow V' \times L^2(\mathbb{R}_+)$

$$F(v, \mu) = \begin{pmatrix} A_n v + \mu - \frac{P^{n-1}}{\Delta t_n} \\ \mu - \max(0, \bar{\mu} - \chi(v - P^0)) \end{pmatrix},$$

it is proved in [12] that $G(v, \mu) : V \times L^2(\mathbb{R}_+) \rightarrow V' \times L^2(\mathbb{R}_+)$ defined by

$$G(v, \mu)h = \begin{pmatrix} A_n h_1 + h_2 \\ h_2 - \chi 1_{\{\bar{\mu} - \chi(v - P^0) > 0\}} h_1 \end{pmatrix}$$

is a generalized derivative of F in the sense that

$$\lim_{\|h\| \rightarrow 0} \frac{\|F(v + h_1, \mu + h_2) - F(v, \mu) - G(v + h_1, \mu + h_2)h\|}{\|h\|} = 0;$$

Note that

$$G(P^{n,k}, \mu^k)h = \begin{pmatrix} A_n h_1 + h_2 \\ h_2 - \chi 1_{\mathcal{A}^{-,k+1}} h_1 \end{pmatrix}.$$

Thus the primal-dual active set algorithm above can be seen as a semi-smooth Newton method applied to F , i.e.

$$(P^{n,k+1}, \mu^{k+1}) = (P^{n,k}, \mu^k) + G^{-1}(P^{n,k}, \mu^k)F(P^{n,k}, \mu^k). \quad (3.16)$$

Indeed, calling $(\delta P^n, \delta \mu) = (P^{n,k+1} - P^{n,k}, \mu^{k+1} - \mu^k)$, it is straightforward to see that in the primal-dual active set algorithm, we have

$$\begin{aligned} A_n \delta P^n + \delta \mu &= -A_n P^{n,k} - \mu^k + \frac{P^{n-1}}{\Delta t_n}, \\ \delta \mu &= -\mu^k \text{ on } \mathcal{A}^{+,k+1}, \\ \delta \mu - \chi \delta P^n &= -\mu^k + \bar{\mu} - \chi(P^{n,k} - P^0) \text{ on } \mathcal{A}^{-,k+1}, \end{aligned}$$

which is precisely (3.16).

In [12], Ito and Kunish, by using the results proved in [11], establish that the primal-dual active set algorithm converges from any initial guess, and that if the initial guess is sufficiently close to the solution of (3.13), then the convergence is superlinear.

To compute numerically the solution of (3.9), it is possible to compute successively the solutions $(P^n(\chi_\ell), \mu(\chi_\ell))$ of (3.13) for a sequence of parameters (χ_ℓ) converging to $+\infty$: to compute $(P^n(\chi_{\ell+1}), \mu(\chi_{\ell+1}))$, one uses the primal-dual active set algorithm with initial guess $(P^n(\chi_\ell), \mu(\chi_\ell))$.

Notice that it is possible to use the same algorithm for the fully discrete problem. Convergence results hold in the discrete case if there is a discrete maximum principle. The algorithm amounts to solving a sequence of systems of linear equations, and the matrix of the system varies at each iteration.

3.5 Mesh adaption

One of the key features of the finite element method is that it permits to compute reliable and often very efficient error indicators for the error between the exact and discrete solutions. We do not wish to develop much on this topic here, because this needs long and technical arguments. We rather refer to [4]. The strategy relies on the fact that the error due to the time discretization can be estimated separately: there are indicators for the error due to time discretization, and at each time step, indicators for the error due to the discretization with respect to the price variable. The error indicators are local. Therefore, they tell us where the time grid, and the mesh in the S variable should be refined.

To illustrate this, we consider an American put, with strike $K = 100$. The interest rate is 0.04 as above, but the volatility is local and we choose:

$$\sigma(S, t) = 0.1 + 0.1 * 1_{100(t-0.5)^2 + \frac{(S-90)^2}{100} < 2}(S, t),$$

so the volatility is piecewise constant and takes the value 0.2 in an ellipse and 0.1 outside. With such a choice, the exercise boundary is expected to change slope as it enters and comes out the region where $\sigma = 0.2$. On Figure 1, we plot the volatility surface as a function of S and t . The exercise boundary is displayed on Figure 3: we see that the free boundary does change slope when the volatility jumps. We see also that refinement is crucial in order to catch properly the exercise boundary. Note that the function γ is not monotone. On Figure 2, two meshes are displayed: we see that the refinement follows the free boundary. On Figure 4, the error indicators with respect to S are plotted: here again, we see that the error indicators are large near the free boundary, where the function P is singular.

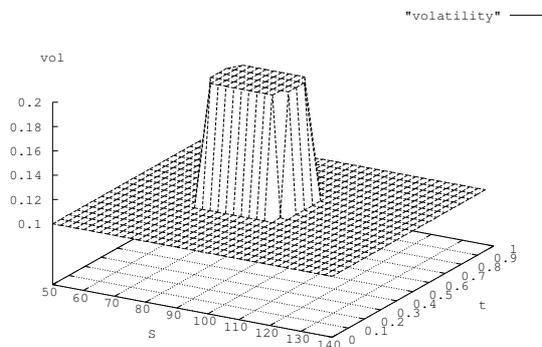


Figure 1: The local volatility surface

4. Calibration with American options

The calibration problem consists in finding η from the observations of

- the spot price S_0 today,

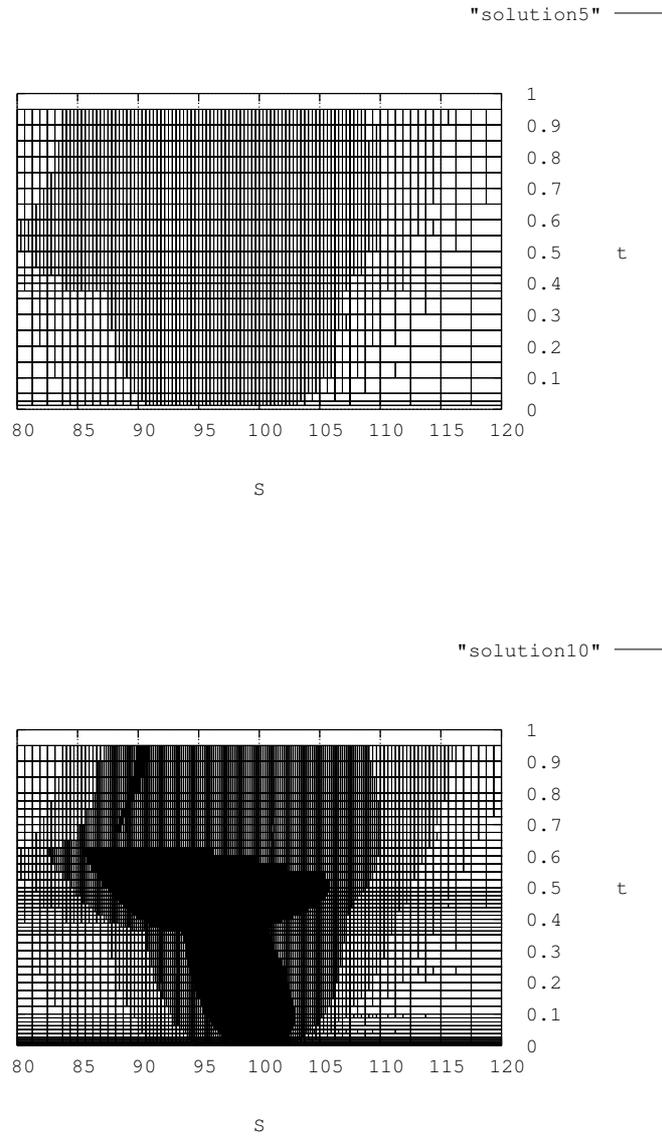


Figure 2: Two successive mesh refinements: the mesh is refined along the exercise boundary, see next figure

- the prices $(\bar{P}_i)_{i \in I}$ of a family of American puts with different maturities and different strikes $(T_i, K_i)_{i \in I}$.

We call $T = \max_{i \in I} T_i$. We assume that for any $i \in I$, the maturity T_i coincides with some node of the time grid, i.e. there exists $N_i \leq N$ such that $t_{N_i} = T_i$. We also assume that for any $i \in I$, the strike K_i is a node of the S -grid, i.e. there exists $\kappa_i < N_h$ such that $K_i = S_{\kappa_i}$.

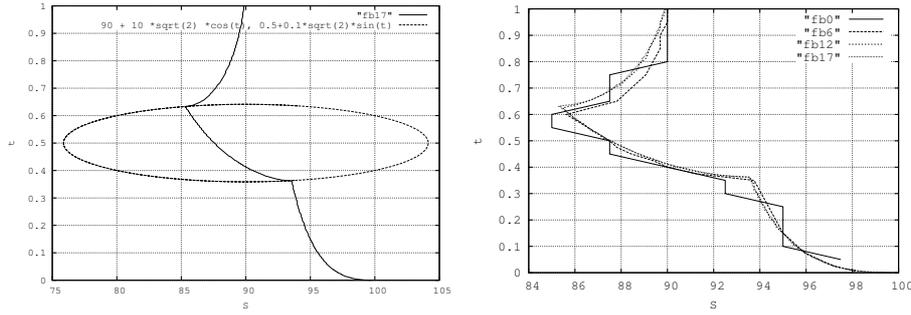


Figure 3: Left: the exercise boundary for the final mesh and the ellipse where the volatility jumps: there are two singularities corresponding to the jumps of volatility. Right: the exercise boundaries for different mesh refinements

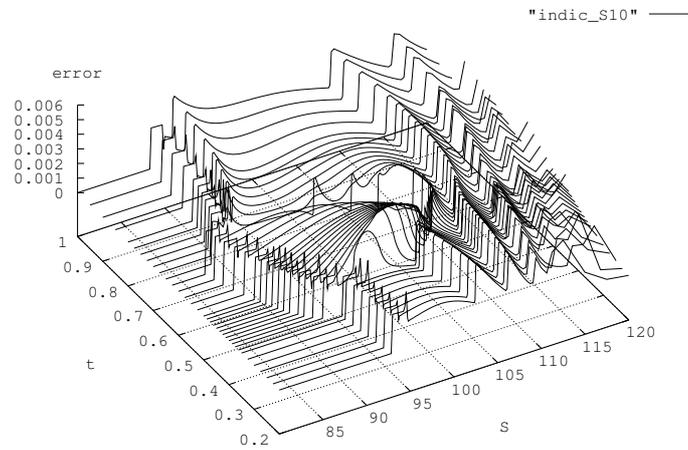


Figure 4: Indicators for the error due to the discretization with respect to S

We consider the least square problem:

$$\text{find } \eta \in \mathcal{H} \text{ minimizing } J(\eta) + J_R(\eta), \quad J(\eta) = \sum_{i \in I} |P_i^{N_i}(S_o) - \bar{P}_i|^2, \quad (4.1)$$

where \mathcal{H} is a suitable closed subset of a possibly infinite dimensional function space, J_R is a suitable Tychonoff regularization functional, and

$$\text{find } (P_i^n)_{0 \leq n \leq N_i}, P_i^n \in \mathcal{K}_{h,i} \text{ satisfying}$$

$$P_i^0 = P_{o,i}, \quad (4.2)$$

and for all $n, 1 \leq n \leq N_i,$

$$\forall v \in \mathcal{K}_{h,i}, \quad (P_i^n - P_i^{n-1}, v - P_i^n) + \Delta t_n a_{T_i - t_n}(P_i^n, v - P_i^n) \geq 0, \quad (4.3)$$

where $P_{\circ,i} = (K_i - S)_+$. We call $\mu_{i,j}^n$ the real number

$$\mu_{i,j}^n = (P_i^n - P_i^{n-1}, w^j - P_i^n) + \Delta t_n a_{T_i - t_n}(P_i^n, w^j - P_i^n). \quad (4.4)$$

4.1 Optimality conditions

In [1], the inverse problem corresponding to the continuous counterpart of (4.3) is studied and optimality conditions are given for suitable choices of \mathcal{H} and J_R . Here, we aim at finding optimality conditions for the fully discrete problem (4.1).

In order to find optimality conditions for the present least-square problem, we replace the state equations (4.2) (4.3) by the above mentioned penalized problem, whose penalty parameter, called ε , will tend to 0. Doing so, we obtain a new least square problem, for which necessary optimality conditions are easily found. Then, we pass to the limit as ε goes to zero: we obtain the following result:

Theorem 3. *Let η^* be a minimizer of (4.1) which can be found as a limit of a sequence η_ε^* of minimizers for the penalized problems, and let $(P_i^{*,n})_{i \in I}$ be the solutions to (4.2) (4.3) with $\eta = \eta^*$. There exist $y_i^{*,n} \in \tilde{V}_h$, and $\alpha_{i,j}^n \in \mathbb{R}$, $1 \leq n \leq N_i$, $\rho \leq j \leq N_h$, $i \in I$, such that $\forall v \in \tilde{V}_h$,*

$$\begin{aligned} (y_i^{*,N_i}, v) + \Delta t_{N_i} \left(a_0^*(v, y_i^{*,N_i}) + \sum_{j=\rho}^{N_h} \alpha_{i,j}^{N_i} v(S_j) \right) &= 2(P_i^{*,N_i}(S_\circ) - \bar{P}_i) v(S_\circ), \\ (y_i^{*,n} - y_i^{*,n+1}, v) + \Delta t_n \left(a_{T_i - t_n}^*(v, y_i^{*,n}) + \sum_{j=\rho}^{N_h} \alpha_{i,j}^n v(S_j) \right) &= 0, \quad 1 \leq n < N, \end{aligned} \quad (4.5)$$

with, for all j, n , $\rho \leq j \leq N_h$, $1 \leq n \leq N_i$,

$$\alpha_{i,j}^n (P_i^{*,n}(S_j) - P_\circ(S_j)) = 0, \quad \mu_{i,j}^{*,n} y_i^{*,n}(S_j) = 0, \quad \alpha_{i,j}^n y_i^{*,n}(S_j) \geq 0,$$

such that for any $\eta \in \mathcal{H}$, noting by $\delta\eta = \eta - \eta^*$,

$$\begin{aligned} 0 \leq & \langle DJ_R(\eta^*), \delta\eta \rangle \\ & - \frac{1}{2} \sum_{i \in I} \sum_{n=1}^{N_i} \Delta t_n \sum_{j=\rho}^{N_h} S_j^2 \delta\eta(S_j, T_i - t_n) y_i^{*,n}(S_j) \left(\frac{P_i^{*,n}(S_j) - P_i^{*,n}(S_{j-1})}{h_j} \right. \\ & \left. + \frac{P_i^{*,n}(S_j) - P_i^{*,n}(S_{j+1})}{h_{j+1}} \right). \end{aligned}$$

4.2 Differentiability

Proposition 1. *Take $\Delta t \leq \frac{1}{2\lambda}$ with λ as in (2.8). Assume that for all $\eta \in \mathcal{H}$ verifying (2.6) (2.7), the parameters h and $\frac{h^2}{\min_n \Delta t_n}$ are small enough so that the matrices \mathbf{A}_l^n and $\mathbf{M}_l + \Delta t_n \mathbf{A}_l^n$ are tridiagonal irreducible M -matrices for all n , $1 \leq n \leq N$ and l , $\alpha \leq l < N_h$. Let $\eta \in \mathcal{H}$ be such that the strict complementarity conditions*

$$P_i^n(S_j) > P_{\circ,i}(S_j) \Leftrightarrow \mu_{i,j}^n = 0, \quad (4.6)$$

for all $i \in I$ and for all $j, \rho \leq j \leq N_h$, where P_i^n is the solution to (2.3) (2.4), and $\mu_{i,j}^n = (P_i^n - P_i^{n-1}, w^j) + \Delta t_n a_{T-t_n}(P_i^n, w^j)$. The functional J is differentiable at η , and for any admissible variation χ of η ,

$$\begin{aligned} & \langle DJ(\eta), \chi \rangle = \\ & -\frac{1}{2} \sum_{i \in I} \sum_{n=1}^N \Delta t_n \sum_{j=\rho}^{N_h} S_j^2 \chi(S_j, T-t_n) y_i^n(S_j) \left(\frac{P_i^n(S_j) - P_i^n(S_{j-1})}{h_j} \right. \\ & \left. + \frac{P_i^n(S_j) - P_i^n(S_{j+1})}{h_{j+1}} \right). \end{aligned} \quad (4.7)$$

where $y_i^n = y^n(\eta)_i \in \tilde{V}_h$, $\alpha_{i,j}^n \in \mathbb{R}$, $\rho \leq j \leq N_h$, are the solution to: $\forall v \in \tilde{V}_h$,

$$\begin{aligned} & (y_i^{N_i}, v) + \Delta t_{N_i} \left(a_0(v, y_i^{N_i}) + \sum_{j=\rho}^{N_h} \alpha_{i,j}^{N_i} v(S_j) \right) = 2(P_i^{N_i}(S_o) - \bar{P}_i)v(S_o), \\ & (y_i^n - y_i^{n+1}, v) + \Delta t_n \left(a_{T-t_n}(v, y_i^n) + \sum_{j=\rho}^{N_h} \alpha_{i,j}^n v(S_j) \right) = 0, \quad 1 \leq n < N, \end{aligned}$$

with

$$\alpha_{i,j}^n (P_i^n(S_j) - P_o(S_j)) = 0, \quad \mu_{i,j}^n y_i^n(S_j) = 0, \quad \alpha_{i,j}^n y_i^n(S_j) \geq 0.$$

4.3 Algorithm

We describe the simplest possible projected descent method in the space Y , where the descent direction is computed thanks to the considerations above. The degrees of freedom of a function $\chi \in Y$ are the values of χ at some nodes of a grid and we call them $(\Lambda_\ell^*(\chi))_{1 \leq \ell \leq L}$, (Λ_ℓ^* is the linear form on Y which maps χ to its value at a given node). We endow Y with the basis $(\Lambda_\ell(\chi))_{1 \leq \ell \leq L}$ defined by $\Lambda_\ell^*(\Lambda_k) = \delta_{\ell k}$, and we define the inner product $(\sum_{\ell=1}^L a_\ell \Lambda_\ell, \sum_{\ell=1}^L b_\ell \Lambda_\ell)_Y = \sum_{\ell=1}^L a_\ell b_\ell$.

Algorithm

- Choose $\eta \in \mathcal{H}$, $\varepsilon > 0$ and $\rho > 0$, set $e = +\infty$.
- **While** $e > \varepsilon$ **do**
 1. Compute $(P_i)_{i \in I}$ by (4.2) (4.3), by using for example one of the algorithms proposed in §3.3 and $J(\eta) + J_R(\eta)$, $J(\eta) = \sum_{i \in I} |P_i^{N_i}(S_o) - \bar{P}_i|^2$,
 2. For all $i \in I$, compute $(y_i^n)_{1 \leq n \leq N_i}$, $y_i^n \in \tilde{V}_h$ satisfying (4.5).
 3. compute $\zeta \in Y$ such that for all $\chi \in Y$,

$$\begin{aligned} & (\zeta, \chi)_Y \\ & = -\frac{1}{2} \sum_{i \in I} \sum_{n=1}^{N_i} \Delta t_n \sum_{j=\rho}^{N_h} S_j^2 \chi(S_j, T-t_n) y_i^n(S_j) \left(\frac{u_i^n(S_j) - u_i^n(S_{j-1})}{h_j} \right. \\ & \left. + \frac{u_i^n(S_j) - u_i^n(S_{j+1})}{h_{j+1}} \right). \end{aligned} \quad (4.8)$$

4. set $\tilde{\eta} = \pi_{\mathcal{H}}(\eta - \rho(\text{grad} J_R(\eta) + \zeta))$, $e = \|\tilde{\eta} - \eta\|$, $\eta = \tilde{\eta}$, where $\pi_{\mathcal{H}}$ is the projection on \mathcal{H} .

- `end_do`

The complete justification of the algorithm above is still an open question because it is not proved that $-\text{grad}J_R(\eta) - \zeta$ is always a descent direction. However, from Proposition 1, we know that most often, ζ is exactly $\text{grad}J(\eta)$: in this case, the algorithm coincides with a projected gradient method.

In the numerical tests below, we have used variants of this algorithm (an interior point algorithm due to J. Herskovits[9]-it is a quasi-Newton algorithm which can handle general constraints), which have proved very robust. In particular, we never experienced breakups caused by the fact that the direction ζ is not a descent direction.

Parallelism The algorithm above can be parallelized in a very natural way on a distributed memory machine with N_p processors, because the computations of the pairs (P_i, y_i) , $i \in I$ are independent from each other. We split I in $I = \cup_{k=1}^{N_p} I_k$ in order to balance the amount of work among the processors, the processor labelled k being responsible for the sums over $i \in I_k$ in $J(\eta)$ and (4.8). Note that the complexity of the computation of P_i, y_i depends on i , so load balancing is not straightforward. The data for η and ζ are replicated on the N_p processors. The processor labelled k computes its own contribution to $J(\eta)$ and to (4.8), i.e. the sums over $i \in I_k$, in an independent manner, then communications are needed for assembling the sums over $i \in I$ in $J(\eta)$ and in (4.8).

For programming, we have used C++ with the message passing library *mpi*.

4.4 Results with American Puts on the FTSE Index

In this paragraph, we consider American puts on the FTSE index. The data correspond to June 6, 2001. We thank José Da Fonseca for providing us with the data.

The price of the underlying asset is $x_0 = 5890$. The American puts correspond to four different maturities: 0.122, 0.199, 0.295, 0.55 years. We set $T = 0.55$. The interest rate r varies with time, so r is replaced by $r(t)$ in (4.3), and this function is known. For these maturities, the prices of the observed options vs strike are plotted on Figure 5. The aim is to find the volatility surface from these prices. The volatility is discretized by functions that are the sum of

- a piecewise affine function in the S -variable which is constant in the regions $S < 1000$ and $S > 9000$ and affine in the region $1000 < S < 9000$
- a bicubic spline in the region $1000 < S < 9000$, $|t - T/2| < T/2 + 0.1$, whose value and derivatives vanish on the boundary of this rectangle. The control points of the spline are plotted on Figure 6, where the time variable is $T - t$. We see that the control points are not uniformly distributed: the mesh is refined for small times t and at the money.

The grid for u is non uniform with 745 nodes in the S -direction and 210 nodes in the t direction. For simplicity, the grid is chosen in such a way that the points $(T_i, K_i)_{i \in I}$ coincide with some grid nodes.

The (squared) volatility obtained at convergence is displayed on Figure 7: the surface has a smile shape. The relative errors between the observed prices and those computed at convergence are plotted on Figure 8, top. They are rather large for small values of K . However, we have to realize that the available observed prices are themselves given with a round-off error, which is exactly 0.5.

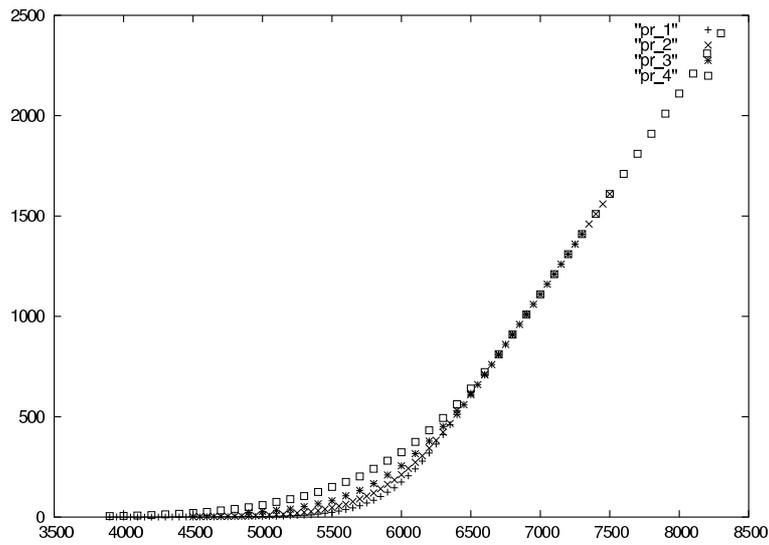


Figure 5: The data for the inverse problem: the prices of a family of American puts on the FTSE index

On Figure 8, bottom, we have plotted the relative round-off error on the observed prices. Doing so, we see that the relative errors on the prices at convergence are of the same order as the round-off error on the observed prices. Therefore, it is very natural that the optimization program cannot improve on this level of error.

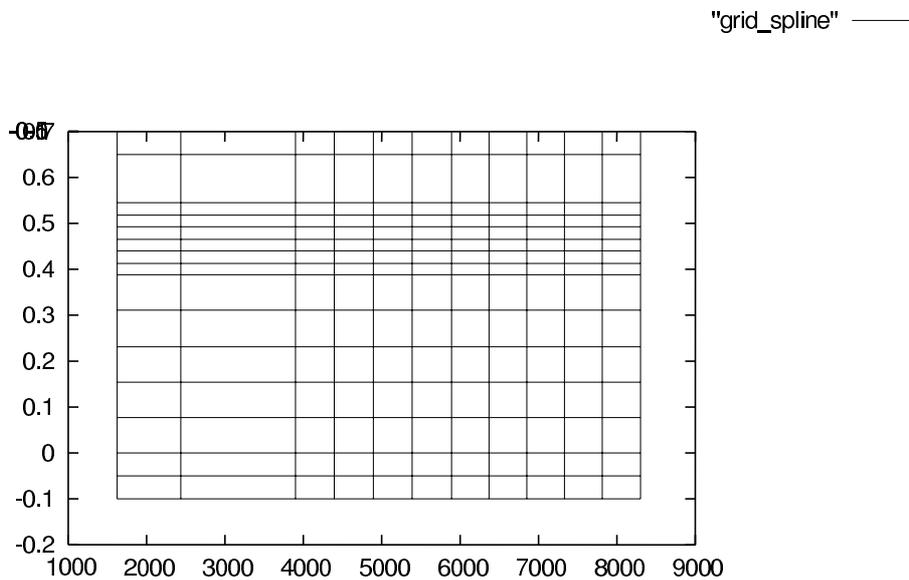


Figure 6: The control points of the bicubic splines

5. Pricing an American option on a basket containing two assets

The aim of this section is to price an American option on a basket containing two assets. We

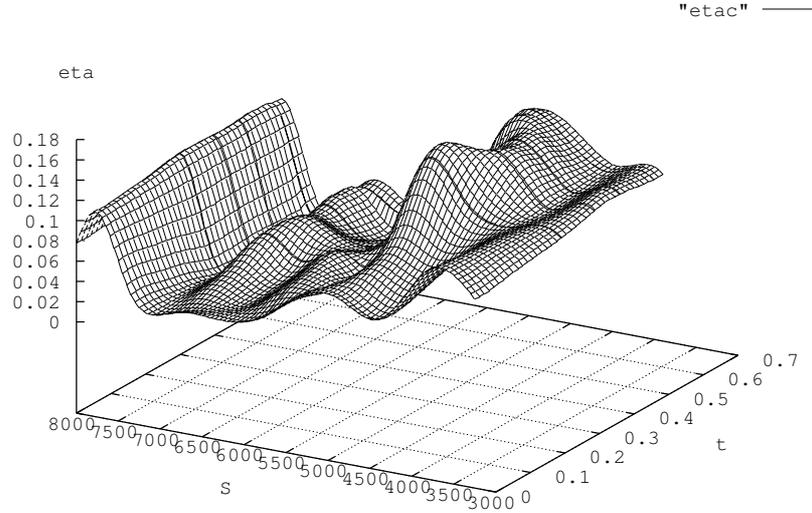


Figure 7: The squared volatility surface obtained by running the calibration program

assume that the prices of the underlying assets obey a system of stochastic differential equations:

$$\begin{aligned} dS_{1t} &= S_{1t}(\mu_1 dt + \frac{\sigma_1}{\sqrt{1+\rho^2}}(dW_{1t} + \rho dW_{2t})), \\ dS_{2t} &= S_{2t}(\mu_2 dt + \frac{\sigma_2}{\sqrt{1+\rho^2}}(\rho dW_{1t} + dW_{2t})), \end{aligned} \quad (5.1)$$

where W_{1t} and W_{2t} are two independent standard Brownian motions. For simplicity, we assume that σ_1 and σ_2 are positive constants, but generalization to functions $\sigma_1(S_1, S_2, t)$ and $\sigma_2(S_1, S_2, t)$ can be considered. The parameter ρ is the correlation factor: $-1 < \rho < 1$. Also for simplicity, we assume that the interest rate r of the risk-free asset is constant.

Consider an American option on this two assets basket, whose payoff function is $P_0(S_1, S_2)$. Under the risk neutral probability, and replacing the time with the time to maturity, the price of the option is given by $P_t = P(S_{1t}, S_{2t}, t)$, where the function P is the solution to the set of inequalities:

$$\begin{aligned} \frac{\partial P}{\partial t} - \frac{1}{2} \sum_{k,l=1}^2 \Xi_{k,l} S_k S_l \frac{\partial^2 P}{\partial S_k \partial S_l} - \sum_{k=1}^2 r S_k \frac{\partial P}{\partial S_k} + rP &\geq 0 \quad t \in (0, T], S_1, S_2 > 0, \\ P(S_1, S_2, t) &\geq P_0(S_1, S_2) \quad t \in (0, T], S_1, S_2 > 0, \\ P(S_1, S_2, 0) &= P_0(S_1, S_2) \quad S_1, S_2 > 0. \end{aligned} \quad (5.2)$$

and to the complementarity condition

$$\left(\frac{\partial P}{\partial t} - \frac{1}{2} \sum_{k,l=1}^2 \Xi_{k,l} S_k S_l \frac{\partial^2 P}{\partial S_k \partial S_l} - \sum_{k=1}^2 r S_k \frac{\partial P}{\partial S_k} + rP \right) (P - P_0) = 0 \quad \begin{matrix} t \in (0, T], \\ S_1, S_2 > 0. \end{matrix} \quad (5.3)$$

The tensor Ξ ,

$$\Xi = \begin{pmatrix} \sigma_1^2 & \frac{2\rho}{1+\rho^2} \sigma_1 \sigma_2 \\ \frac{2\rho}{1+\rho^2} \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix}$$

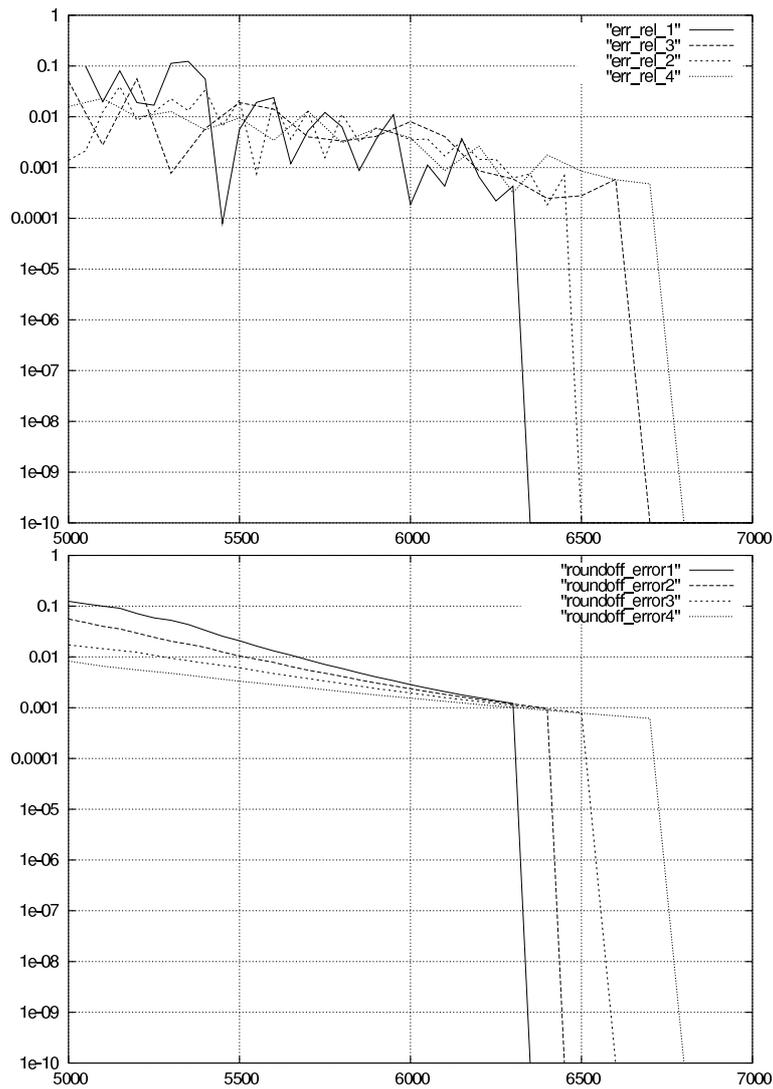


Figure 8: Top : relative errors between the observed prices and those obtained with η found after running the calibration program. A curve corresponds to a given maturity. Bottom: relative round-off error on observed prices. The two errors are of the same order.

is clearly positive definite.

Consider the payoff function

$$P_0(S_1, S_2) = \min((K_1 - S_1)_+, (K_2 - S_2)_+). \tag{5.4}$$

For this choice of P_0 , there is a free boundary, which is now a two-dimensional surface in $\mathbb{R}_+ \times \mathbb{R}_+ \times [0, T]$, which separates the zone where $P(t, S_1, S_2) = P_0(S_1, S_2)$ and where the option should be exercised and the zone where $P(t, S_1, S_2) > P_0(S_1, S_2)$. Finding this exercise boundary is much more difficult than for a single underlying asset. Moreover, the function P exhibits a singularity on the free boundary; therefore, when using a discrete method for pricing the American option, the mesh needs to be refined near the free boundary, which is an unknown of the problem.

This motivates the following strategy:

- we use triangular piecewise linear and continuous finite elements in the variables S_1, S_2 .
- the mesh is refined adaptively to the solution. By and large, starting from a possibly coarse grid, the idea is to adaptively construct a regular mesh (the angles of the triangles are bounded from below by a fixed constant) in the metric generated by an approximated Hessian of the computed solution. Although the theory for this strategy is not as clean as that using residuals, this method gives generally very good results. We have used the same adaptive algorithm as in the freeware FreeFem++, by F. Hecht and O. Pironneau [17], which is available on <http://www.ann.jussieu.fr/hecht/freemem++.html>. This software permits to use 2D finite elements and mesh adaption by means of a user friendly dedicated language.
- the time stepping scheme is an implicit Euler scheme, and mesh adaption is done every 10 time steps.
- At each time step, one has to solve a discrete variational inequality corresponding to an elliptic problem. For that, we use the active set strategy introduced in § 3.4.

In Figure 9, we plot the contours of the pricing function at two different dates: one month and one year to maturity, as well as the adapted mesh. The parameters are $r = 0.05$, $K_1 = K_2 = 100$, $\sigma_1 = \sigma_2 = 0.2$, and $\Xi_{1,2} = 0.15\sigma_1^2$. In Figure 10, we plot the exercise zone and near the region where the exercise boundary crosses the line $S_1 = S_2$. We see that the free boundary has a singularity in this region, and capturing this singularity really requires a well adapted mesh. The method for

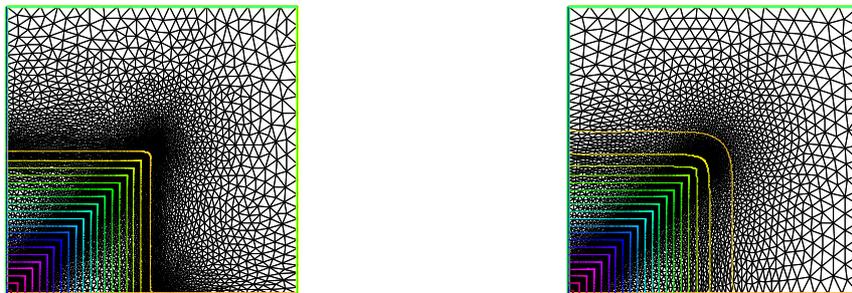


Figure 9: The contours for the price of a put on a basket containing two assets, at $t = 1$ month and $t = 1$ year to maturity. The adapted mesh is plotted too.

computing the discrete solution at each step is the active set algorithm of § 3.4. The parameters are $\chi = 10^7$ and $\bar{\mu} = 0$. In Figure 11, we plot the ℓ^2 norm of the increment of the vector containing the values of the solution at the nodes of the triangulation, between two successive iterates of the semi-smooth Newton method, as a function of the iteration number. We see that the convergence is superlinear, and that ten iterations are enough to reduce the norm of the increment to 10^{-9} . In Figure 12, we plot the increment of nodes in the active set (in absolute value) between two successive iterates of the semi-smooth Newton method, as a function of the iteration number. We see that the increment drops rapidly to some units. In Figure 13, we plot the contours of the pricing

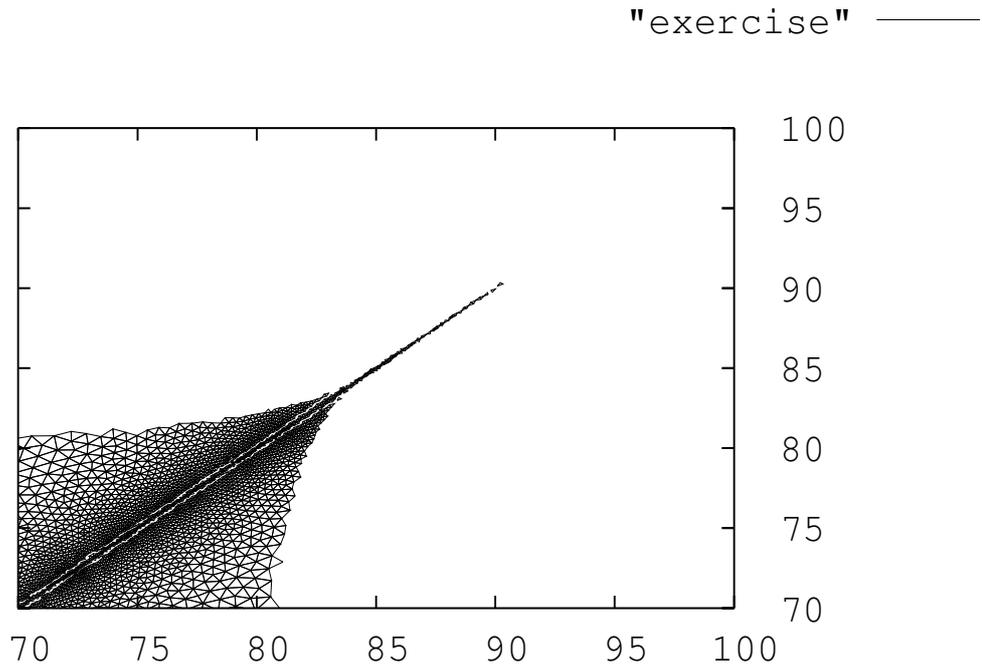


Figure 10: The exercise region one year to maturity (zoom)

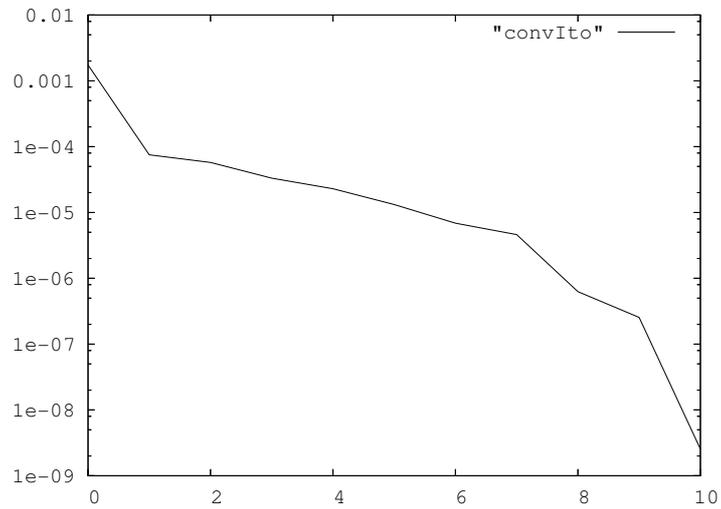


Figure 11: The convergence of the active set strategy

POS(CSTNA2005)020

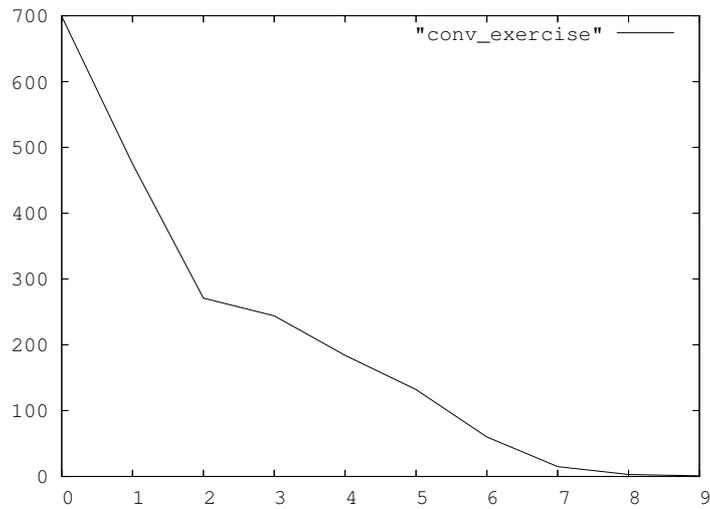


Figure 12: The convergence of the active set strategy

function for a different payoff function, i.e.

$$P_0(S_1, S_2) = \min((K - S_1)_+, (K - S_2)_+) + (K \min((S_1 - K)_+, (S_2 - K)_+))^{1/2}$$

with $K = 100$. The adapted mesh can be seen on the same Figure. In Figure 14, we plot the graph of the pricing function eleven months to maturity.

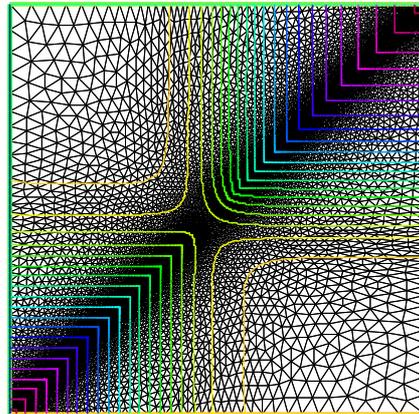


Figure 13: The contours of the option price and the adapted mesh eleven months to maturity

6. Conclusion

Up to our knowledge, calibration with American options has not been much discussed in the literature. In contrast with European options, it does not seem possible to use the Dupire forward

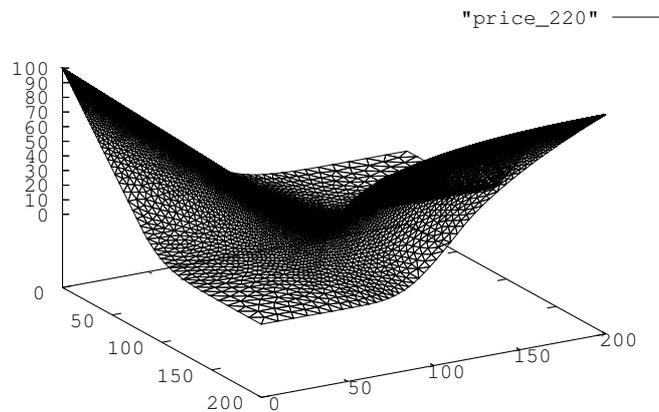


Figure 14: The option price eleven months to maturity

equation, because American option pricing is intrinsically nonlinear.

We have presented a least square method where the descent direction is computed by solving adjoint problems. We have proved that most often, this direction is exactly the gradient of the cost function. The method is computationally intensive, but it can be parallelized in a straightforward manner. We believe that parallelization is necessary for a practical use. The algorithm was used on academical and realistic test cases. We have not taken yet historical data into consideration: they can be used in the Tychonoff functional for stabilizing the calibration and also for speeding up the computations, since the volatility is not expected to vary too much on a daily scale. This remains to be done.

We have also presented a method for pricing efficiently American options on small baskets, involving finite elements, mesh adaption and an active set strategy at each time step.

Finally, let us mention that calibration of Levy processes can also be done with American option. This is the topic of a work in preparation [2].

References

- [1] Y. Achdou. An inverse problem for a parabolic variational inequality arising in the calibration of American options. *SIAM J. Control Optim.*, 43(5):1583–1615, 2005.
- [2] Yves Achdou. An inverse problem for a parabolic variational inequality with an integro-differential operator arising in the calibration of lévy processes with American options. in progress.
- [3] Yves Achdou and Olivier Pironneau. Volatility smile by multilevel least squares. *International Journal of Theoretical and Applied Finance*, 5(6):619–643, 2002.
- [4] Yves Achdou and Olivier Pironneau. *Computational methods for option pricing*. Frontiers in Applied Mathematics. SIAM, 2005.

- [5] Yves Achdou and Olivier Pironneau. A numerical procedure for the calibration of American options. *Applied mathematical finance*, 12(3), 2005.
- [6] M. Avellaneda, M. Friedman, C. Holmes, and D. Samperi. Calibrating volatility surfaces via relative entropy minimization. *Applied Mathematical Finances*, 4:37–64, 1997.
- [7] Rama Cont and Peter Tankov. *Financial modelling with jump processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton, FL, 2004.
- [8] Jean-Pierre Fouque, George Papanicolaou, and K. Ronnie Sircar. *Derivatives in financial markets with stochastic volatility*. Cambridge University Press, Cambridge, 2000.
- [9] J. Herskovits. Feasible direction interior-point technique for nonlinear optimization. *J. Optim. Theory Appl.*, 99(1):121–146, 1998.
- [10] S. Heston. A closed form solution for options with stochastic volatility with application to bond and currency options. *Review with Financial Studies*, pages 327–343, 1993.
- [11] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.*, 13(3):865–888 (electronic) (2003), 2002.
- [12] Kazufumi Ito and Karl Kunisch. Semi-smooth Newton methods for variational inequalities of the first kind. *M2AN Math. Model. Numer. Anal.*, 37(1):41–62, 2003.
- [13] Patrick Jaillet, Damien Lamberton, and Bernard Lapeyre. Variational inequalities and the pricing of American options. *Acta Appl. Math.*, 21(3):263–289, 1990.
- [14] R. Lagnado and S. Osher. Reconciling differences. *RISK*, 10:79–83, 1997.
- [15] R. Lagnado and S. Osher. A technique for calibrating derivative security pricing models: numerical solutions of an inverse problem. *The journal of computational finance*, 1(1):13–25, 1997.
- [16] D. Lamberton and B. Lapeyre. *Introduction au calcul stochastique appliqué à la finance*. Ellipses, 1997.
- [17] O. Pironneau and F. Hecht. *FREEFEM*. www.ann.jussieu.fr.