INSTITUTE OF PHYSICS PUBLISHING INVERSE PROBLEMS

Inverse Problems **19** (2003) 91–127 PII: S0266-5611(03)38970-1

Calibration of the local volatility in a trinomial tree using Tikhonov regularization

S Crépey

Artabel SA, 69 rue de Paris, F-91400 Orsay, France

E-mail: stephane.crepey@artabel.net

Received 2 July 2002, in final form 14 November 2002
Published 9 December 2002
Online at stacks.iop.org/IP/19/91

Abstract

Following an approach introduced by Lagnado and Osher (Lagnado R and Osher S 1997 *J. Comput. Finance* **1** 13–25), we study the application of Tikhonov regularization to the financial inverse problem of calibrating a local volatility function from observed vanilla option prices. Moreover, we provide a unified treatment for this problem in two different settings: first, the generalized Black–Scholes model, and second, a trinomial tree discretization. We present serial and parallel implementations of the method in the discrete setting, using a probabilistic interpretation to compute, at significantly reduced cost, the gradient of the cost criterion. We illustrate the stability of this regularized calibration procedure by numerical examples. Finally we extend this methodology to the problem of calibration with American option prices.

1. Introduction

1.1. Presentation of the problem

Following an approach introduced by Lagnado and Osher [1], we study the inverse problem, in finance, of calibrating a local volatility function from observed option prices, using Tikhonov regularization. We consider the problem in two different settings: first, the generalized Black–Scholes model, and second, a trinomial tree discretization. In Crépey [2, 3], we have already considered this problem from a theoretical point of view, in the continuous setting of the generalized Black–Scholes model. Here, we shall consider a suitable discretization and the associated regularized calibration algorithm. Moreover, we also consider the problem of calibration with American option prices.

Let us recall the framework in Crépey [2, 3]. In market finance, a European *call* (*put*) option with maturity date T and strike K, on an underlying asset S, means a right to buy (sell), at price K, a unit of S at time T. We consider a theoretical financial market, with two traded

assets: cash, with constant interest rate r, and a risky stock, with diffusion price process driven by a standard Brownian motion W:

$$\mathrm{d}S_t = S_t(\rho(t, S_t)\,\mathrm{d}t + \sigma(t, S_t)\,\mathrm{d}W_t), \qquad t > t_0; \qquad S_{t_0} = S_0.$$

Moreover, the stock is assumed to yield a continuously compounded dividend at constant rate q. Suppose finally the market to be liquid, non-arbitrable and perfect. By arbitrage arguments and Itô's lemma, European calls/puts on S then have a theoretical fair price within the model that we will denote by $\Pi_{T,K}(t_0, S_0; a)$, where $a \equiv \sigma^2/2$ and

$$\Pi_{T,K}(t_0, S_0; a) = \mathrm{e}^{-r(T-t_0)}\mathrm{E}_P^{t_0,S_0}(S_T - K)^{+/-}. \tag{1}$$

Here P denotes the so-called *risk-neutral* probability, under which

$$\mathrm{d}S_t = S_t((r-q)\,\mathrm{d}t + \sigma(t, S_t)\,\mathrm{d}W_t), \qquad t > t_0; \qquad S_{t_0} = S_0. \tag{2}$$

Alternatively to the probabilistic representation (1), the prices Π can be given as the solution to a differential equation. One can use either the *Black–Scholes* backward parabolic equation [4], in the variables (t_0, S_0), which is

$$\begin{cases} -\partial_t\Pi - (r-q)S\partial_S\Pi - a(t, S)S^2\partial^2_{S^2}\Pi + r\Pi = 0, & t < T \\ \Pi|_T \equiv (S-K)^{+/-}, \end{cases} \tag{3}$$

or the *Dupire* forward parabolic equation [5], in the variables (T, K), given by

$$\begin{cases} \partial_T\Pi - (q-r)K\partial_K\Pi - a(T, K)K^2\partial^2_{K^2}\Pi + q\Pi = 0, & T > t_0 \\ \Pi|_{t_0} \equiv (S_0-K)^{+/-}. \end{cases} \tag{4}$$

Notice that (1) can be viewed as the Feynman–Kac representation for the solution of (3). As for (4), it is a Fokker–Planck equation integrated twice with respect to the space variable K, using the formal identity

$$\partial^2_{K^2}(S_0 - K)^{+/-} \equiv \delta_{S_0}(K),$$

where δ_{S_0} denotes the Dirac mass at S_0.

In (2)–(4), the yields r and q are assumed to be known constants. However, they could in fact be any deterministic known functions of time. The local volatility function σ, or $a \equiv \sigma^2/2$, is an unknown function of time and stock. The calibration problem is the inverse problem that amounts to inferring the local volatility function a from market-quoted prices of liquid, so-called *vanilla*, options, typically European calls and puts with various strikes and maturities. The local volatility function thus inferred is then used to price *exotic* (non-vanilla) options, and value hedge ratios or derivative exposure, consistently with the market. This problem, known as *fitting the smile* by market practitioners, is hence the reconstruction of a local volatility function, supposed to be prevailing as the underlier's risk-neutral dynamics. It is indeed important for applications that the reconstruction of such a prevailing dynamics be as fair as possible. But this calibration problem is under-determined (since the set of observed prices is finite) and ill-posed, so that *ad hoc* stabilizing procedures must be used. A variant of the problem, also considered in this paper, consists of the calibration of a local volatility function with *American* option prices. We will refer to the former as *the European*, and the latter as *the American*, calibration problem.

1.2. Survey of the literature

This calibration problem has received intensive study over the last ten years, see for instance [1–3, 5–21] and references therein. The American calibration problem is addressed in [22–24]. All these references illustrate both the mathematical interest in the problem and its relevance from the financial point of view. The Markovian feature of the generalized Black–Scholes

model (2) is an attractive framework for market practitioners. The drawback is that this model is more relevant in specific market conditions, namely those where the volatility is strongly correlated to the market level. In practice, this model is more specifically dedicated to equity index option markets, especially for shorter-dated options. These are mainly American option markets, which makes the American calibration problem all the more important. For more details on these econometric and empirical aspects, we refer the reader to the monograph by Rebonato [12]. In more general market conditions, one should move to non-Markovian models, such as stochastic volatility models [25], jump-diffusion models [26, 27] or models of dynamics of implied volatility [28, 29].

Let us come back to our calibration problem. The first attempts to solve it viewed the problem as one of differentiation from discrete data, interpolating the finite set of observed call prices π for $(T, K) \in obs$ to a full set $\tilde{\Pi}$ of call prices for every (T, K), and recovering a from Dupire's equation (4). This approach was introduced in Dupire [5] and developed in Rubinstein [8], Derman and Kani [6], Derman *et al* [7] or Andersen and Brotherton-Ratcliffe [9], among others. It was also applied to the American calibration problem in Chriss [22] or Barle and Cakici [23]. But it does not address the fact that the calibration problem is ill-posed. So the local volatility surface thus calibrated was quite unstable and dependent upon the interpolation procedure used for π. Alternative parametric approaches came up against the lack of a natural parametrization for the local volatility function. Dumas *et al* [10] used an elementary parametrization, while Coleman *et al* [11] worked with splines. Finally, Bouchouev *et al* [13, 14] studied the stationary case $\sigma(S)$, or the case where the term and strike structures of the local volatility function are decoupled, taking $\sigma(t, S) \equiv \rho(t)\sigma(S)$.

Various regularization methods were also applied to the problem. All of them use a *prior* a_0, which reflects *a priori* information about a. We can classify these methods into two categories.

The first one uses *entropic regularization*, after reformulation of the problem as one of constrained stochastic control, with cost criterion

$$\rho(a, a_0)^2 = E_P \int_{T=t_0}^{\overline{T}} (a - a_0)^2 \ ,$$

and constraints $\Pi|_{obs}(a) = \pi$ or $d(\Pi|_{obs}(a), \pi) \leqslant \delta$, where $d(\Pi|_{obs}(a), \pi)$ denotes a distance between the model prices $\Pi(a)$ and the observed prices π. The probabilistic representation (1) for option prices then allows us to give to this constrained stochastic control problem a dual formulation, using Lagrange multipliers. This approach was introduced with equality constraints by Avellaneda *et al* [15], while Samperi [16] deepened some theoretical aspects.

The second one, which is also the one considered in this paper, uses *Tikhonov variational regularization* by the H^1-norm of $a - a_0$. For seminal and survey references regarding Tikhonov regularization, we refer the reader to Tikhonov [30] and Engl *et al* [31, ch 10]. This approach tackles the calibration problem as a minimization problem, of a cost criterion

$$J_\alpha(a) \equiv d(\Pi|_{obs}(a), \pi)^2 + \alpha\rho(a, a_0)^2.$$

Here $d(\Pi|_{obs}(a), \pi)$ denotes a distance between the model prices $\Pi(a)$ and the observed prices π, α is the so-called *regularization parameter* and ρ is a penalty designed to keep a close to the prior a_0, namely $\rho(a, a_0)^2 \equiv \|a - a_0\|^2_{H^1}$, where

$$\|u\|^2_{H^1} \equiv \iint u^2 + \|\nabla u\|^2,$$

the H^1-squared norm of u. This formulation of the calibration problem was introduced by Lagnado and Osher [1]. Jackson *et al* [17] coupled this approach with a spline parametrization for the local volatility function. Bodurtha and Jermakyan [18] developed a linearized variant.

Berestycki *et al* [20] used as a state variable the implied Black–Scholes volatility corresponding to the model prices $\Pi(a)$, rather than these prices themselves. Achdou and Pironneau [21] used as a state variable the prices Π in the variables (T, K) rather than (t_0, S_0), with H^2-regularization and a finite element discretization. Huang and Pang [24] studied the American calibration problem. In my PhD thesis [19, Part IV] (in French), I have proposed an implicit finite differences implementation which can be used as an alternative to the one in Lagnado and Osher [1]. The idea was to use an *approximate* gradient of the cost criterion computed at the expense of pricing the options and solving *one* Dupire equation, instead of pricing the options and solving one Black–Scholes equation with source term *by option and mesh node* to obtain the *exact* gradient in the original presentation by Lagnado and Osher. The present paper is about an innovative trinomial tree implementation of the method, in which the *exact* gradient can be computed at the expense of pricing the options and solving *one* Fokker–Planck equation. Therefore the accuracy of the original method is preserved but the computational time is drastically reduced. Typically the calibration time can be reduced from about 1 h to about 1 min or less on a standard serial Pentium PC. The use of parallelism allows one to gain a further factor. Moreover we extend this methodology to the American calibration problem.

1.3. Overview

Let us now give an overview of the paper. Section 2 states results of the general theory, surveyed for instance in Engl *et al* [31, ch 10], of the Tikhonov regularization method for ill-posed nonlinear inverse problems. In the generalized Black–Scholes model (2), results regarding stability at fixed $\alpha > 0$, as well as convergence and convergence rates when $\alpha \to 0$, have been shown for the Tikhonov regularized calibration method in Crépey [2, 3]. We refer the reader to the latter for every detail, while section 3 recalls the main results useful for a good understanding of the present paper. In section 4, we obtain the analogous results for a Markov chain approximation of (2), including a probabilistic interpretation for the gradient of the cost criterion. This results in section 5 in an efficient implementation of the method, as accurate as, and faster than, Lagnado and Osher's algorithm [1]. Moreover we propose an MPI-parallel implementation. Numerical experiments on real data sets illustrate the effectiveness and robustness of the resulting regularized calibration procedure (section 6). In section 7 we benchmark our approach with other methods. Finally in section 8 we consider the American calibration problem.

The numerical experiments presented in this paper were performed with the `CalibrationEngine` (of P Cohort, S Crépey, S Farcy and C Martini, Artabel SA). This embeds, in a unified software, both serial and MPI-parallel implementations of various approaches to the calibration problem, in particular a splined-based method of differentiation from discrete data using Dupire's equation (4), a trinomial tree implementation of the variant of Avellaneda *et al* [15] constrained stochastic control approach with inequality constraints, and the trinomial tree implementation of the Tikhonov regularization method which constitutes the core of this paper.

2. Tikhonov regularization of nonlinear inverse problems

Let us first recall some known results regarding Tikhonov regularization of ill-posed nonlinear inverse problems. The reference for this section is the work by Engl *et al* [32].

2.1. Definitions and general assumptions

In [32] the authors consider a nonlinear operator between Hilbert spaces

$$\mathcal{H} \supseteq \mathcal{A} \ni a \overset{\Pi}{\longmapsto} \Pi(a) \in \mathcal{H}'.$$

The operator Π is supposed to be *weakly sequentially closed*. This means that for every $\mathcal{A}$-valued sequence a_n such that a_n converges weakly to a in $\mathcal{H}$ and $\Pi(a_n)$ converges weakly to π in $\mathcal{H}'$, then $a \in \mathcal{A}$ and $\pi = \Pi(a)$. We will only be concerned with the special case where $\mathcal{A}$ is a closed convex subset of $\mathcal{H}$, $\mathcal{H}' = \mathbb{R}^M$ for some fixed $M \in \mathbb{N}^\star$, and Π satisfies the following stronger assumption.

Assumption 2.1 (Compacity). $\Pi(a_n) \to \Pi(a)$ in $\mathbb{R}^M$ if $a_n \to a$ weakly in $\mathcal{H}$ as $n \to \infty$.

We will also assume that Π is twice Gateaux differentiable as follows.

Assumption 2.2 (Differentiability). There exist linear and bilinear forms $d\Pi(a)$ on $\mathcal{H}$ and $d^2\Pi(a)$ on $\mathcal{H}^2$ such that

$$\Pi(a + \varepsilon h) = \Pi(a) + \varepsilon d\Pi(a) \cdot h + \frac{\varepsilon^2}{2} d^2\Pi(a) \cdot (h, h) + \mathrm{o}(\varepsilon^2); \qquad a, a + h \in \mathcal{A}$$

with

$$\begin{aligned} &\|d\Pi(a) \cdot h\| \leqslant C\|h\|; && a \in \mathcal{A}, h \in \mathcal{H} \\ &\|d^2\Pi(a) \cdot (h, h')\| \leqslant C\|h\|\|h'\|; && a \in \mathcal{A}, h, h' \in \mathcal{H}, \end{aligned}$$

where C is a constant independent of $a \in \mathcal{A}$.

Remark 2.3. Usually the assumption is stated as *Fréchet* differentiability of the direct operator, but only Gateaux differentiability will be required in the following. In any case the operators that we will consider in sections 3 and 4 are actually Fréchet differentiable, even if not stated as such.

In practice, the set of observations, or input data for the inverse problem, is only known up to some noise δ. For instance, market prices π^δ are defined as bid–ask spreads. Moreover, any numerical procedure to tackle an inverse problem entails some computational burden η. Tikhonov regularization allows us to overcome such data noise and computational burden.

Definition 2.4 (α-solution). *By an α-solution of the inverse problem for Π with prior $a_0 \in \mathcal{H}$, noisy data $\pi^\delta \in \mathbb{R}^M$ and computational burden $\eta \geqslant 0$, we mean any $a_\alpha^{\delta,\eta} \in \mathcal{A}$ such that*

$$J_\alpha^\delta(a_\alpha^{\delta,\eta}) \leqslant J_\alpha^\delta(a) + \eta, \qquad a \in \mathcal{A}$$

where

$$2J_\alpha^\delta(a) \equiv \|\Pi(a) - \pi^\delta\|^2 + \alpha\|a - a_0\|^2. \tag{5}$$

2.2. Properties of Tikhonov regularized solutions

First, the so-called *stability*, in the following theorem, means the sequential compacity of the α-solutions, at any fixed $\alpha > 0$.

Theorem 2.5 (Stability; [32, theorem 2.1]). *Assume $\pi^{\delta_n} \to \pi^\delta$ and $\eta_n \to \eta \equiv 0$ when $n \to +\infty$. Then any sequence of α-solutions $a_\alpha^{\delta_n,\eta_n}$ admits a subsequence which converges towards an α-solution $a_\alpha^{\delta,\eta\equiv 0}$.*

Remark 2.6. Engl *et al* [32, theorem 2.1] assume $\eta_n \equiv 0$. However, their proof extends immediately to the case of a sequence η_n that tends to 0 as $n \to \infty$.

Assuming further that the data lie in the range of the model leads to asymptotic properties of α-solutions as $\alpha \to 0$.

Assumption 2.7 (Range property). $\pi \in \Pi(\mathcal{A})$.

Definition 2.8 (a_0-MNS). *Under assumption 2.7, we shall call an a_0-minimal norm solution (a_0-MNS) of the inverse problem for Π with data π, any solution a that minimizes $\|a - a_0\|$ over the set of all solutions.*

Theorem 2.9 (Convergence; [32, theorem 2.3]). *Under assumption 2.7, suppose that*

$$(n \in \mathbb{N}) \qquad \|\pi - \pi^{\delta_n}\| \leqslant \delta_n,$$
$$(n \to +\infty) \qquad \alpha_n, \delta_n^2/\alpha_n, \eta_n/\alpha_n \longrightarrow 0.$$

Then any sequence $a_{\alpha_n}^{\delta_n,\eta_n}$ admits a subsequence which converges towards an a_0-MNS a. Moreover, $a_{\alpha_n}^{\delta_n,\eta_n} \to a$, if a is the unique a_0-MNS of the inverse problem for Π with data π.

Remark 2.10. In the special case where the direct operator Π is linear, Tikhonov regularization then appears as an approximating scheme for the pseudo-inverse of Π.

Finally, one can get convergence rate estimates, uniform over all data $\pi \in \Pi(\mathcal{A})$ sufficiently close and smooth with respect to the prior a_0. Let $d\Pi(a)^\star$ denote the adjoint of the operator $h \mapsto d\Pi(a) \cdot h$.

Theorem 2.11 (Convergence rates; [32, theorem 2.4 and remark 2.5]). *Under assumption 2.7, suppose that*

$$\|\pi - \pi^\delta\| \leqslant \delta, \qquad \alpha \sim \delta, \qquad \eta = \mathrm{O}(\delta^2).$$

Then $\|a_\alpha^{\delta,\eta} - a\| = \mathrm{O}(\sqrt{\delta})$, for any a_0-MNS a of the inverse problem for Π with data π such that

$$a - a_0 = d\Pi(a)^* \lambda, \tag{6}$$

for some sufficiently small $\lambda \in \mathbb{R}^M$.

Proof. In our notation, the condition on λ in [32, theorem 2.4 and remark 2.5] can be rewritten as

$$2\left\langle \lambda, \int_0^1 (1-t) d^2\Pi(a + t(a_\alpha^{\delta,\eta} - a))(a_\alpha^{\delta,\eta} - a, a_\alpha^{\delta,\eta} - a)\, \mathrm{d}t \right\rangle \leqslant \rho \|a_\alpha^{\delta,\eta} - a\|^2,$$

for some $\rho < 1$. But, provided that λ is sufficiently small in $\mathbb{R}^M$, this indeed follows from our differentiability assumption 2.2 by the Cauchy–Schwarz inequality. □

Remark 2.12. This theorem is the only one that requires the differentiability assumption 2.2.

Remark 2.13. An interesting feature of Tikhonov regularization is that the data set π does not have to belong to the range of the direct operator for applicability of the method—even if assumption 2.7 is the simplest assumption for the previous results regarding convergence and convergence rates (in fact, the minimum assumption for such results is the existence of a least squares solution to the inverse problem, see [33, proposition 3.2]).

To alleviate the notation, we shall assume in the following sections 3–5 that there are only calls in the calibration data. The next two sections give two concrete settings in which assumptions 2.1 and 2.2 will be seen to hold true, so that all the results in this section are applicable.

3. Continuous setting

The first of these settings is the generalized Black–Scholes model, already considered in Crépey [2, 3]. The results presented in this section are not new, except theorem 3.3, which does not appear explicitly in Crépey [2, 3] but can be derived in the same manner as theorem 3.2. We state these results once again here to emphasize their analogy with the results obtained in the next section. These will be obtained in a new setting, which can be seen as a discretization of the generalized Black–Scholes model with trinomial trees, or explicit finite differences.

3.1. The setting

Given a plane strip $Q =]t_0, \overline{T}[\times\mathbb{R}$, constant bounds $0 < \underline{a} \leqslant \overline{a}$ and a real measurable prior a_0 on Q such that $\underline{a} \leqslant a_0 \leqslant \overline{a}$, let us denote by

$$\mathcal{A} \equiv \{a \in a_0 + H^1(Q);\ \underline{a} \leqslant a \leqslant \overline{a}\},$$

where $H^1(Q)$ is the usual Sobolev space on Q such that

$$\|u\|^2_{H^1(Q)} \equiv \iint_Q u^2 + \|\nabla u\|^2.$$

Let there also be given a finite subset $\mathcal{F} \subset \overline{Q}$ with $|\mathcal{F}| = M \in \mathbb{N}^\star$. We define Π as the *pricing functional*

$$\mathcal{A} \ni a \overset{\Pi}{\longmapsto} \Pi_{\mathcal{F}}(t_0, y_0 = 0; a) \in \mathbb{R}^M,$$

where

$$\Pi_{T,k}(t_0, y_0; a) = \mathrm{e}^{-r(T-t_0)}\mathrm{E}_P^{t_0,S_0}(S_T - K)^+$$

denotes, for $(T, k) \in \mathcal{F}$, the price of the European call with maturity T and strike $K = \exp(k)$, at the current date t_0 and underlying asset value S_0, in the generalized Black–Scholes model (2) expressed with the logarithmic variable

$$y \equiv \ln\left(\frac{S}{S_0}\right) - (r - q - \underline{a})(t - t_0).$$

Now, it follows from the results in Crépey [2, 3] that the pricing functional Π satisfies the regularity assumptions 2.1 and 2.2. We refer the reader to Crépey [2, 3] for the technical details and other results. The proofs use $W^1_{2,p}$ estimates for Black–Scholes and Dupire equations, as well as viscosity considerations. The main difficulty is to show the estimates in assumption 2.2 for a bound C independent of $a \in \mathcal{A}$. Notice that Crépey [2, 3] uses the logarithmic variable $y = \ln(S)$, instead of y as above. The present choice of variables gives rise to better stability conditions after discretization of the problem (see section 4.1). With this choice, all the results in Crépey [2, 3] are easily seen to be applicable, except those, irrelevant here, concerning the analysis of the stability of the method with respect to perturbations of (t_0, S_0).

3.2. The gradient

Furthermore,

Theorem 3.1 (See [2, 3]). *For $(T, k) \in \mathcal{F}$, the derivative of $\Pi_{T,k}(t_0, y_0; a)$ in the direction $h \in H^1(Q)$, admits the following Feynman–Kac representation:*

$$d\Pi_{T,k}(t_0, y_0; a) \cdot h = \int_{t=t_0}^{T} \int_{y=-\infty}^{\infty} h(t, y)\Gamma_{T,k}(t, y; a)\gamma_{t_0,y_0}(t, y; a)\,\mathrm{d}y\,\mathrm{d}t, \quad (7)$$

where $\Gamma_{T,k}(t, y; a) \equiv (\partial^2_{y^2} - \partial_y)\Pi_{T,k}(t, y; a)$ *and* $\gamma_{t_0,y_0}(t, y; a)$ *denotes, for almost every* $t > t_0$*, the transition probability density between* t_0 *and* t *discounted at rate* r*, that is,* $\mathrm{e}^{-r(t-t_0)} \times$ *the density, of the process*

$$y_t \equiv \ln\left(\frac{S_t}{S_0}\right) - (r - q - \underline{a})(t - t_0).$$

Finally, let Δ denote the Laplacian operator on $H^2(Q)$.

Theorem 3.2 (See [2, 3]). *Assuming moreover that* a *is uniformly continuous with respect to its space variable* y*, then condition (6) means that* $\Lambda \equiv a - a_0$ *is the unique strong solution in* $H^2(Q)$ *of the following nonlocal problem:*

$$\begin{cases} \Lambda - \Delta\Lambda = \displaystyle\sum_{(T,k)\in\mathcal{F};\, t\leqslant T} \lambda_{T,k}\Gamma_{T,k}(t, y; a)\gamma_{t_0,y_0}(t, y; a), & Q \text{ a.e.} \\ \partial_n\Lambda = 0, \qquad \partial Q \text{ a.e.} \end{cases} \tag{8}$$

where the normal derivative $\partial_n\Lambda \in L_2(\partial Q)$ *is well defined, for* $\Lambda \in H^2(Q)$.

In the same way, let ∇ denote the Gateaux derivative in $H^1(Q)$.

Theorem 3.3 (See [2, 3, 19]). *If* a *is uniformly continuous with respect to its space variable* y*, then* $u \equiv \nabla J^\delta_\alpha(a) - \alpha(a - a_0)$ *belongs to* $H^2(Q)$*, and* u *is the unique strong solution in* $H^2(Q)$ *of the following problem:*

$$\begin{cases} u - \Delta u = \displaystyle\sum_{(T,k)\in\mathcal{F};\, t\leqslant T} (\Pi_{T,k}(t_0, y_0; a) - \pi^\delta_{T,k})\Gamma_{T,k}(t, y; a)\gamma_{t_0,y_0}(t, y; a), & Q \text{ a.e.} \\ \partial_n u = 0, \qquad \partial Q \text{ a.e.} \end{cases} \tag{9}$$

where the normal derivative $\partial_n u \in L_2(\partial Q)$ *is well defined, for* $u \in H^2(Q)$.

To simplify the notation, we will assume in the next two sections that all maturities with observed prices fall at steps of a constant time subdivision of $[t_0, \overline{T}]$, where $\overline{T} \equiv \max_{(T,k)\in\mathcal{F}} T$.

4. The problem within a trinomial tree

It turns out that there is a natural discretization of the setting of the problem in section 3.1, which keeps all the required properties. The idea is to use a Markov chain algorithm to specify the same problem in a fully discrete setting. In order to handle the key point that the local volatility may vary within a range, we adopt a trinomial tree method where the mesh is fixed once and for all, and the local volatility varies from node to node.

4.1. The setting

Define $\underline{\sigma}$ and $\overline{\sigma}$ such that $\underline{a}, \overline{a} \equiv \underline{\sigma}^2/2, \overline{\sigma}^2/2$. We choose a Markov chain with time step $\tau = (\overline{T} - t_0)/N$ and space step $\varepsilon \equiv \beta\overline{\sigma}\sqrt{\tau}$, where β is some fixed parameter, the so-called *stretch factor*. Starting from $y_0 = 0$, we look for a scheme with, at each node y_n, a local transition probability $p = P(y_{n+1} = y_n - \varepsilon | y_n)$, $p' = P(y_{n+1} = y_n + \varepsilon | y_n)$, so that $P(y_{n+1} = y_n | y_n) = 1 - p - p'$.

Then, given a local node volatility $a \equiv a(t_{n+1}, y_n) \in [\underline{a}, \overline{a}]$, it is easy to show that the choice

$$p = \left(\frac{a}{\varepsilon^2} + \frac{(a - \underline{a})}{2\varepsilon}\right)\tau, \qquad p' = \left(\frac{a}{\varepsilon^2} - \frac{(a - \underline{a})}{2\varepsilon}\right)\tau$$

will yield nonnegative weights as long as the following stability conditions hold true:

$$\varepsilon \leqslant \frac{2\overline{a}}{(\overline{a}-\underline{a})}, \qquad 1 \leqslant \beta; \tag{10}$$

and also that the first and second moments of the Markov chain, after the change of variables

$$y \mapsto S = S_0 \exp(y + (r - q - \underline{a})(t - t_0)),$$

will match those of the continuous diffusion (2) with an $o(\tau)$ accuracy as $\tau \to 0$.

Now we set exactly the same problem as in section 3.1, except that the space $\mathcal{A}$ is the set of all possible $a \equiv \sigma^2/2$ at the nodes of the sub-tree $\mathcal{I}$ starting at (t_1, y_0) with $\underline{a} \leqslant a \leqslant \overline{a}$, and Π means the price, or discounted expectation of the payoff function, of a call in the tree with local volatility. Obviously, $\mathcal{A}$ can be identified in a natural way with the product set $[\underline{a}, \overline{a}]^{N^2}$. Moreover, denoting by α_t and α_y auxiliary regularization parameters to be defined later, and by $\mathcal{B}$ the set of bottom nodes of $\mathcal{I}$, we endow $\mathcal{A}$ with the inner product

$$\langle u, v\rangle_{h^1} = \alpha_t D_t(u, v) + \alpha_y D_y(u, v), \tag{11}$$

where

$$D_t(u, v) \equiv \tau^{-1}\varepsilon \sum_{(t_n, y_m)\in\mathcal{I}} (u(t_{n-1}, y_m) - u(t_n, y_m))(v(t_{n-1}, y_m) - v(t_n, y_m)) \tag{12}$$

$$\begin{aligned} D_y(u, v) \equiv \tau\varepsilon^{-1}\Bigg(&\sum_{(t_n, y_m)\in\mathcal{I}} (u(t_n, y_{m+1}) - u(t_n, y_m))(v(t_n, y_{m+1}) - v(t_n, y_m)) \\ &+ \sum_{(t_n, y_m)\in\mathcal{B}} (u(t_n, y_m) - u(t_n, y_{m-1}))(v(t_n, y_m) - v(t_n, y_{m-1}))\Bigg), \end{aligned} \tag{13}$$

with homogeneous boundary conditions, Neumann at the origin of the tree and Dirichlet elsewhere.

In this setting, the pricing functional $a \mapsto \Pi(a)$ is multilinear, hence continuous. More precisely,

$$\Pi_{T,k}(t_0, y_0; a) = \sum_{\mathcal{J}\subseteq\mathcal{I}} C_{\mathcal{J}} \prod_{(t,y)\in\mathcal{J}} a(t, y), \tag{14}$$

where $C_{\mathcal{J}}$ is a linear combination, with rational functions of ε and τ as coefficients, of the values of the call payoff at the terminal nodes of the tree. These are the

$$(S_0 \mathrm{e}^{(y_m + (r-q-\underline{a})(T-t_0))} - \mathrm{e}^k)^+,$$

where $|y_m| \leqslant N\varepsilon$ and $N = (\overline{T} - t_0)/\tau$. The first and second partial derivatives of $\Pi_{T,k}(t_0, y_0; a)$ with respect to the values of the local volatility function a at the nodes of $\mathcal{I}$ inherit the structure (14), whence the differentiability assumption 2.2 for

$$\Pi \equiv \Pi_{T,k}(t_0, y_0; a), \qquad (T, k) \in \mathcal{F}$$

where the bound C in this assumption is independent of $a \in \mathcal{A}$, as wished. Finally, Π satisfies the regularity assumptions 2.1 and 2.2.

4.2. The gradient

Furthermore,

Theorem 4.1. *For $(T, k) \in \mathcal{F}$, the partial derivative of $\Pi_{T,k}(t_0, y_0; a)$ with respect to the value of the local volatility function a at node $(t_n, y_m) \in \mathcal{I}$ with $t_n \leqslant T$ admits the following Feynman–Kac representation:*

$$d\Pi_{T,k}(t_0, y_0; a) \cdot \delta(t_n, y_m) = \Gamma_{T,k}(t_n, y_m; a)\gamma_{t_0,y_0}(t_{n-1}, y_m; a) \exp(-r\tau)\tau, \tag{15}$$

where

$$\Gamma_{T,k}(t_n, y_m; a) \equiv \left(\frac{1}{\varepsilon^2} + \frac{1}{2\varepsilon}\right)\Pi_{T,k}(t_n, y_{m-1}; a) - \frac{2}{\varepsilon^2}\Pi_{T,k}(t_n, y_m; a) + \left(\frac{1}{\varepsilon^2} - \frac{1}{2\varepsilon}\right)\Pi_{T,k}(t_n, y_{m+1}; a),$$

and γ_{t_0,y_0} *denotes the map of Arrow–Debreu prices in the tree at the current phase* (t_0, y_0).

The Arrow–Debreu prices are the transition probabilities discounted at rate r. They can thus be computed forward in the tree using a Fokker–Planck discrete equation that expresses the composition of discounted probabilities.

Proof. By direct calculus,

$$\Pi_{T,k}(t_{n-1}, y_l; a + \delta(t_n, y_m)) - \Pi_{T,k}(t_{n-1}, y_l; a) = \mathbf{1}_{\{l=m\}}\Gamma_{T,k}(t_n, y_m; a)\exp(-r\tau)\tau.$$

So, by linearity, $d\Pi_{T,k}(t_0, y_0; a) \cdot \delta(t_n, y_m)$ is equal to the value at (t_0, y_0) of the diffusion backward in the tree of a single nonzero term

$$\Gamma_{T,k}(t_n, y_m; a)\exp(-r\tau)\tau$$

at (t_{n-1}, y_m). Now, (15) is nothing but the Feynman–Kac representation for $d\Pi_{T,k}(t_0, y_0; a) \cdot \delta(t_n, y_m)$. □

Finally, let us define the following discretized Laplacian operator on $\mathbb{R}^{N^2}$:

$$\Delta_\tau^\varepsilon u \cdot (t_n, y_m) \equiv \alpha_t(u(t_{n-1}, y_m) - 2u(t_n, y_m) + u(t_{n+1}, y_m))/\tau^2 + \alpha_y(u(t_n, y_{m-1}) - 2u(t_n, y_m) + u(t_n, y_{m+1}))/\varepsilon^2, \qquad (t_n, y_m) \in \mathcal{I}$$

with homogeneous boundary conditions, Neumann (Dirichlet) at the origin and at an artificial $t_{N+1} \equiv \overline{T} + \tau$ time step of the tree (elsewhere).

Theorem 4.2. *Condition (6) means that* $\Lambda \equiv a - a_0$ *is the unique solution in* $\mathbb{R}^{N^2}$ *of the following nonlocal problem:*

$$-\Delta_\tau^\varepsilon \Lambda \cdot (t_n, y_m) = \sum_{(T,k)\in\mathcal{F};\, t_n \leqslant T} \lambda_{T,k}\Gamma_{T,k}(t_n, y_m; a)\gamma_{t_0,y_0}(t_{n-1}, y_m; a) \times \exp(-r\tau)/\varepsilon, \qquad (t_n, y_m) \in \mathcal{I}. \qquad (16)$$

Proof. Operate an elementary discrete Green identity on the following adjunction relations:

$$\langle \delta(t_n, y_m), \Lambda\rangle_{h^1} = \sum_{(T,k)\in\mathcal{F}} \lambda_{T,k} d\Pi_{T,k}(t_0, y_0; a) \cdot \delta(t_n, y_m), \qquad (t_n, y_m) \in \mathcal{I}$$

and scale by $\tau\varepsilon$. □

In the same way, let ∇ denote the gradient with respect to the h^1-inner product (11) in $\mathbb{R}^{N^2}$.

Theorem 4.3. $u \equiv \nabla J_\alpha^\delta(a) - \alpha(a - a_0)$ *is the unique solution in* $\mathbb{R}^{N^2}$ *of the following problem:*

$$-\Delta_\tau^\varepsilon u \cdot (t_n, y_m) = \sum_{(T,k)\in\mathcal{F};\, t_n \leqslant T} (\Pi_{T,k}(t_0, y_0) - \pi_{T,k}^\delta)\Gamma_{T,k}(t_n, y_m; a)\gamma_{t_0,y_0}(t_{n-1}, y_m; a) \times \exp(-r\tau)/\varepsilon, \qquad (t_n, y_m) \in \mathcal{I}. \qquad (17)$$

Remark 4.4. Observe the analogy between the identities (15)–(17) and (7)–(9), respectively.

5. A trinomial tree variant of the Lagnado–Osher algorithm

5.1. Outline

Hence, a natural way to tackle numerically the Tikhonov regularized calibration problem in the tree consists of η-minimizing J_α^δ on $a \in [\underline{a}, \overline{a}]^{N^2}$, where $a \equiv \sigma^2/2$ and N is the number of time steps in the tree, using the gradient with respect to the h^1-inner product (11) in $\mathbb{R}^{N^2}$. This gradient can be computed by solving (17), by cyclic reduction using for example `Genbun` in `Slatec` at `Netlib`, or by fast Fourier transform. We will instead use the *Euclidean* gradient, corresponding, as is seen immediately from above, to the right-hand side of (17) minus $\alpha\Delta_\tau^\varepsilon(a - a_0)$, all of it factored by $\tau\varepsilon$. We thus spare the computational cost of solving (17). But our main point here is that sophisticated black-box bound constrained gradient descent minimization routines, such as, for instance, `Lbfgs_b` in `Toms` at `Netlib`, require the *Euclidean* gradient as their argument.

With respect to Lagnado and Osher's algorithm [1], the main interest of this tree implementation follows from the probabilistic representation (15). This allows us to compute the gradient of the cost criterion by pricing the options and solving *one* Fokker–Planck equation in the tree, instead of pricing the options and solving one Black–Scholes equation with source term *by option and mesh node* in the Lagnado–Osher original presentation. Indeed we refer the reader to [1, p 17, equation (2.12) and step 4 in the algorithm]. Step 4 in the algorithm requires them to solve one Black–Scholes equation with source term (2.12) for each i, j, m, n, where (i, j) and (m, n) parametrize the options and mesh nodes, whence 'a large number of direct solutions to the PDE (2.12)' are required. In contrast, our tree implementation amounts to computing the adjoint

$$\gamma_{t_0,y_0}(\cdot; a)(\Pi_{T,k}(t_0, y_0\cdot; a) - \pi_{T,k}^\delta)$$

for the state $\Pi_{T,k}(.; a), (T, k) \in \mathcal{F}$, exactly in reverse mode. The expression for the gradient can be coded by hand; no automatic differentiation tool is needed. The accuracy of the Lagnado–Osher algorithm is preserved, but the computational time is drastically reduced. Typically the calibration time can be reduced from about one hour to about one minute or less on a standard serial Pentium PC. Moreover a parallel implementation allows one to gain a further factor. To do so, one shares between the available processors, for each maturity with observed prices, the computations relative to the options with various strikes. This can be done by using, for instance, the `MPI` library.

Another interest of this tree implementation is that explicit finite difference computations in the tree are less costly than implicit methods for computing option prices. Of course, explicit schemes are subject to the stability condition (10). But, to handle this condition, one only needs to take a space step $\varepsilon \leqslant \frac{2\overline{a}}{(\overline{a}-\underline{a})}$, and then a time step τ such that $\overline{\sigma}\sqrt{\tau} \leqslant \varepsilon$. One could also object that explicit finite differences converge slower than implicit ones, so that the gain in using an explicit scheme would be overwhelmed by the need to use a finer discretization. This is discussed in the appendix.

5.2. Choice of the regularization parameter

Results regarding convergence and convergence rates in section 2.2 relate to *a priori* strategies $\alpha \equiv \alpha(\delta, \eta)$. In the special case where the direct operator Π is linear, various *a posteriori* strategies $\alpha(\delta, \eta; \pi^\delta)$, as well as *error-free* strategies $\alpha(\pi^\delta)$, are known. For a survey, we refer the reader to Engl *et al* [31, sections 4.3–4.5]. *A posteriori* strategies include, for instance, the ones based upon *discrepancy principles*. Well known error-free strategies are the *generalized cross-validation* or the *L-curve* methods. All these strategies come out onto

algorithms that are iterative with respect to α. However, in the linear case, much of the computation can be factorized and done once and for all across all the values of α used. Some of these strategies, such as those based upon discrepancy principles [31, section 10.3], and the associated algorithms iterative with respect to α, extend to the nonlinear framework. But the computations cannot be factorized any more. These methods then require the resolution of several large-scale minimization problems, one for each value of α used, which is often too time-consuming.

Regarding our implementation of the calibration problem, we have heuristically chosen an error-free two-stage strategy as follows. We will come back later to the choices of the bounds $\underline{a}$ and $\overline{a}$, and of the prior a_0. Now, in order to take care of normalization, we minimize a functional slightly different from J_α^δ in (5), namely j_α^δ such that

$$2j_\alpha^\delta(a) \equiv \frac{1}{M}\sum_{(T,k)\in\mathcal{F}}\left(\frac{\Pi_{T,k}(t_0, y_0; a) - \pi_{T,k}^\delta}{\omega_{T,k}}\right)^2 + \alpha_t D_t(a - a_0, a - a_0) + \alpha_y D_y(a - a_0, a - a_0), \tag{18}$$

where the penalization terms were defined in (12), (13) and

$$\omega_{T,k} \equiv \max(\Pi_{T,k}(t_0, y_0; \overline{a}) - \pi_{T,k}^\delta, \pi_{T,k}^\delta - \Pi_{T,k}(t_0, y_0; \underline{a})), \qquad (T,k) \in \mathcal{F}.$$

To determine α_t and α_y, we first solve the *unregularized* calibration problem, by minimization of the least square term in (18), without penalization. Denoting by $a^\star$ the solution thus obtained, we then choose α_t and α_y such that

$$\begin{cases} \dfrac{1}{M}\displaystyle\sum_{(T,k)\in\mathcal{F}}\left(\frac{\Pi_{T,k}(t_0, y_0; a^\star) - \pi_{T,k}^\delta}{\omega_{T,k}}\right)^2 = \alpha_t D_t(a^\star - a_0, a^\star - a_0) + \alpha_y D_y(a^\star - a_0, a^\star - a_0) \\ \alpha_t D_t(a^\star - a_0, a^\star - a_0) = \alpha_y D_y(a^\star - a_0, a^\star - a_0). \end{cases}$$

Moreover, so as to reduce the computation cost, we perform this minimization in a tree with half as many time steps as in the original one. Indeed, at this stage of the computation, we are only interested in the value of the minimum, and not in the solution $a^\star$ itself. Therefore this minimization may be less accurate than the full regularized calibration procedure, namely minimizing (18) in the original tree with the values of α_t and α_y thus determined. Incidentally, computing the regularization parameters in the halved tree results in greater values for these parameters than those which one would obtain in the full tree, hence a greater stability for the overall regularized calibration procedure. We will refer to the resolution of the unregularized and regularized minimization problem as *the first stage* and *the second stage* of the overall regularized calibration procedure.

Notice that the least square term and the penalization term are equal in $j_\alpha^\delta(a^\star)$. Rather like in the L-curve method, the idea is to equilibrate the contributions of the quadratic residual and penalty terms, so as to realize a fair compromise between accuracy and stability in the method. Since the regularization parameters α_t and α_y depend on π^δ, but not directly on δ or η, this is an error-free strategy. This choice is mainly a heuristic one. One could have made other choices, such as *a posteriori* strategies $\alpha(\delta, \eta; \pi^\delta)$, in which one can use the bid–ask spread on the market prices as an estimator for δ, or take δ as the distance between the model and the data, computed by solving the unregularized minimization problem.

6. Numerical experiments

Now we report exhaustive numerical experiments performed on real market data, using the above two-stage trinomial tree regularized calibration algorithm. In the following, the

European call/put *implied volatility*, for given data r, q, t_0, y_0, T, k and price π, means the unique constant volatility such that π is the price of the European call/put in the classic, nongeneralized Black–Scholes model corresponding to these data. The *implied volatility mismatch* between the observed price $\pi^{\delta}_{T,K}$ and the generalized Black–Scholes model price $\Pi_{T,K}(t_0, S_0; a)$ is thus the difference between the corresponding implied volatilities. The *implied volatility surface*, for a local volatility function a, means the surface of implied volatilities corresponding to the call model prices for all $T \geqslant t_0$ and $K > 0$, or (T, K) in a suitable discretization. This computation for an implied volatility surface may thus be considered as an indirect interpolation procedure for the input call prices, or their implied volatilities. Moreover, in contrast to an arbitrary interpolation procedure, the resulting implied volatility surface is inherently arbitrage-free.

We use a trinomial tree as in section 4.1, but with variable time step, so that all maturities with observed prices indeed fall at steps of the time subdivision. To handle the stability condition (10) in the time adaptive tree, we first choose a space step $\varepsilon \leqslant \frac{2\underline{a}}{(\overline{a}-\underline{a})}$, and then a time step τ_i such that $\overline{\sigma}\sqrt{\tau_i} \leqslant \varepsilon$ in each slice $]T_{i-1}, T_i]$ of the tree. The bounds $\underline{a}$ and $\overline{a}$ are fixed thanks to a specific pre-minimization procedure based upon the input implied volatilities. The prior a_0 is taken as a constant equal to the weighted average of the input implied volatilities, with weights proportional to the implied Black–Scholes vegas. Use of a nonconstant prior will also be considered in section 6.5.

The data consist of daily sets of vanilla European call and put option prices on the FTSE and DAX indexes. The current time t_0 corresponding to each of these data sets has been translated to $t_0 = 0$. We filtered out the options with prices outside the well known arbitrage bounds, as well as the less liquid options with moneyness K/S_0 smaller than 0.8 or larger than 1.2. Finally, for data sets with both call and put option prices, the dividend rates that we used were computed so that the observed prices match as well as possible the theoretical call–put parity relation, in the overall least-squares sense. This is an indirect way of correcting the gap between the quotation times for the underlier and the option prices, which are often asynchronous in public data.

6.1. A school example

The first example involves a FTSE data set of call prices at 11 February 2000. These prices are displayed in table 1. The corresponding value of the index was $S_0 = 6219.0$. We used $N = 52$ time steps in the tree, a constant dividend rate based upon a dividend of amount $-0.047\,38$ at maturity $0.191\,781$ and a constant interest rate based upon a zero coupon equal to $0.988\,284$ at maturity $0.191\,781$. Figure 1 displays

(a) the squared local volatility surface σ^2 obtained at the end of the second computation stage,
(b) the implied volatility surface reconstructed thanks to it, and
(c) the accuracy of the overall calibration procedure.

Observe the typical strike structure of the implied volatility surface, smiling at low time-to-maturity and skewed downwards as the time-to-maturity increases. To show the impact of the regularization term, figure 2 (beware of the scale of the ordinate axis) displays the squared local volatility surface at the end of the first computation stage.

6.2. Stability with respect to calendar time

The second experiment illustrates the stability of the method by showing the local squared (figure 3) and implied (figure 4) volatility surfaces obtained on successive DAX data sets at

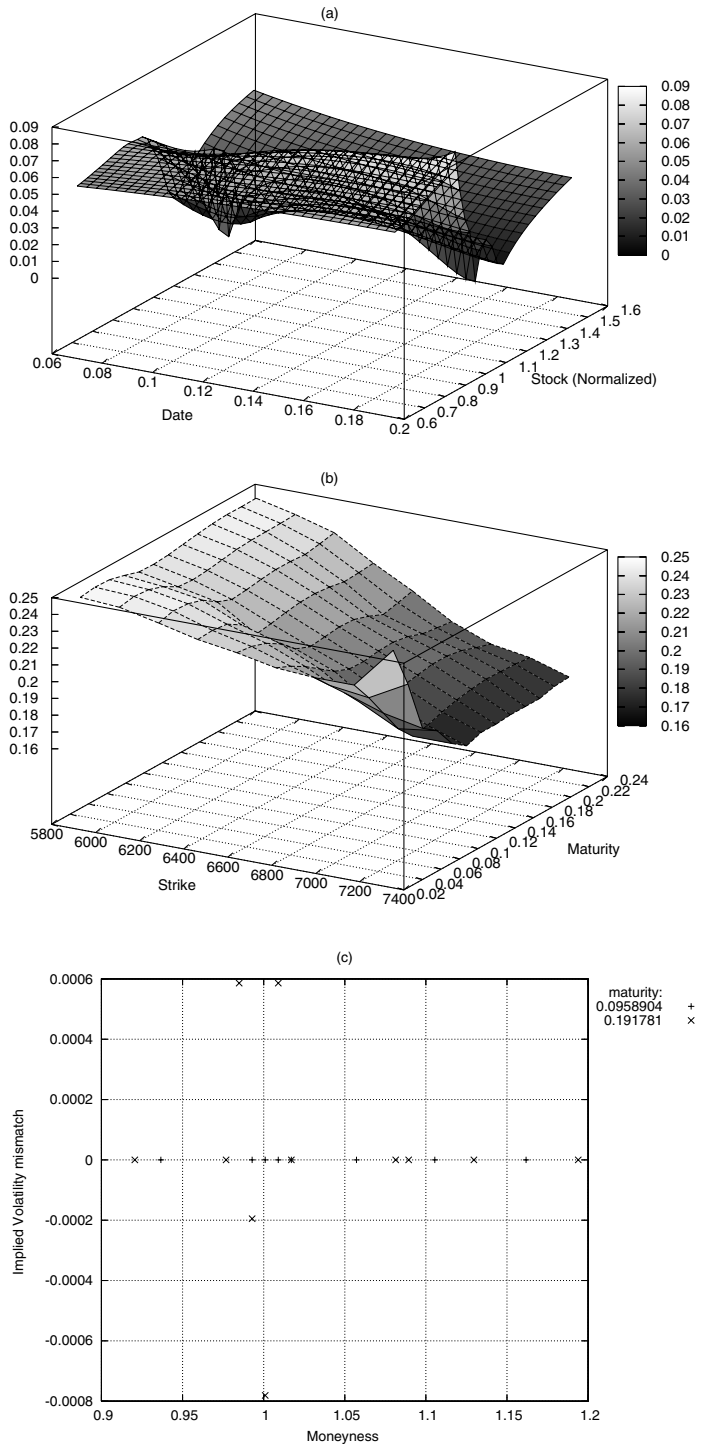

Figure 1. Squared local volatility (a), implied volatility surface (b) and calibration accuracy (c) at the end of the second computation stage, on a FTSE European data set at 11 February 2000.

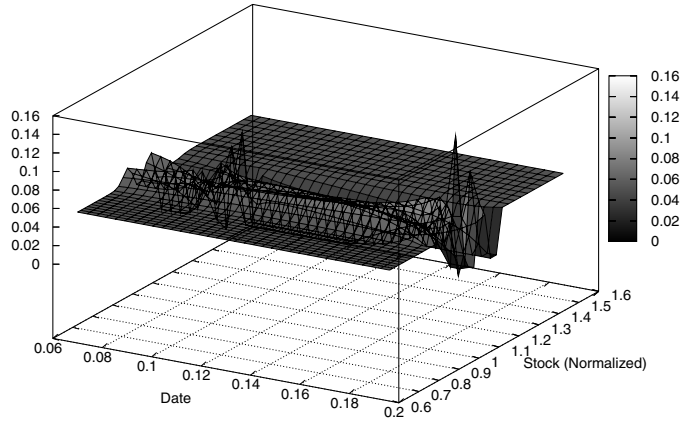

Figure 2. Squared local volatility at the end of the first computation stage, on a FTSE European data set at 11 February 2000.

Table 1. A FTSE data set of European call prices.

Maturity	Strike	Price
0.095 890	5825	469.5
0.095 890	6175	223.5
0.095 890	6225	195.5
0.095 890	6275	169
0.095 890	6325	144.5
0.095 890	6575	56.5
0.095 890	6875	10.5
0.095 890	7225	0.5
0.191 781	5725	631
0.191 781	6075	378
0.191 781	6125	345.5
0.191 781	6175	314.5
0.191 781	6225	284.5
0.191 781	6275	256.5
0.191 781	6325	229.5
0.191 781	6725	72.5
0.191 781	6775	60.5
0.191 781	7025	24
0.191 781	7425	2

May 2 (a), 3 (b) and 4 (c), 2001. We used $N = 31, 26$ and 47 time steps in the trees. Figure 5 displays the accuracy of the calibrations. These are already very satisfying, especially if one considers the small size of the trees we used in these experiments. Notice that the data with shortest time-to-maturity in parts (a) and (b) are not taken into account any longer in part (c), as they become too close to expiry. Recall that in a given generalized Black–Scholes model, hence at fixed local volatility function, the implied volatility surface does vary with calendar time. In contrast, any dynamics of the local volatility surface might reveal some model misspecification. Actually, figure 3 shows that the local volatility functions thus calibrated are fairly regular at fixed date and stable from day to day.

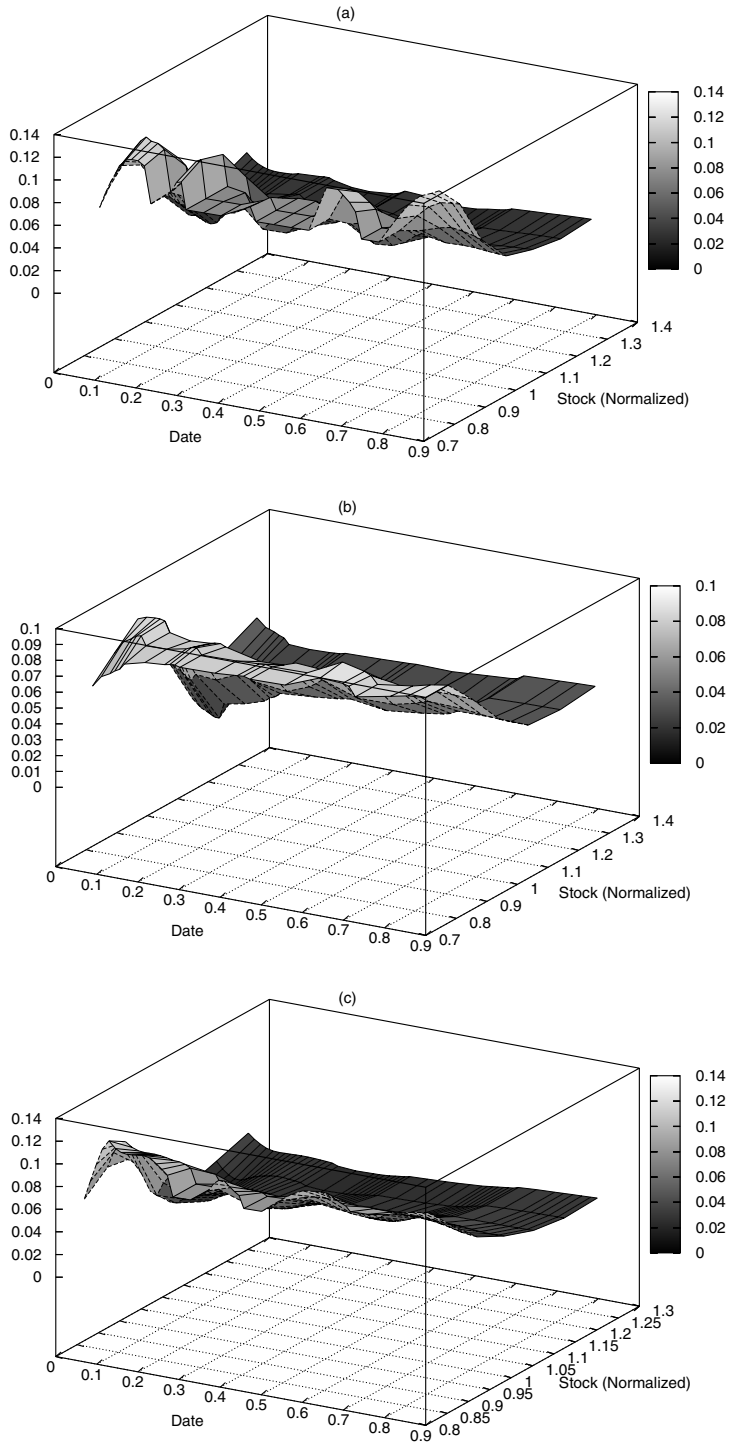

Figure 3. Squared local volatility at the end of the second computation stage, on the DAX data sets of May 2 (a), 3 (b) and 4 (c), 2001.

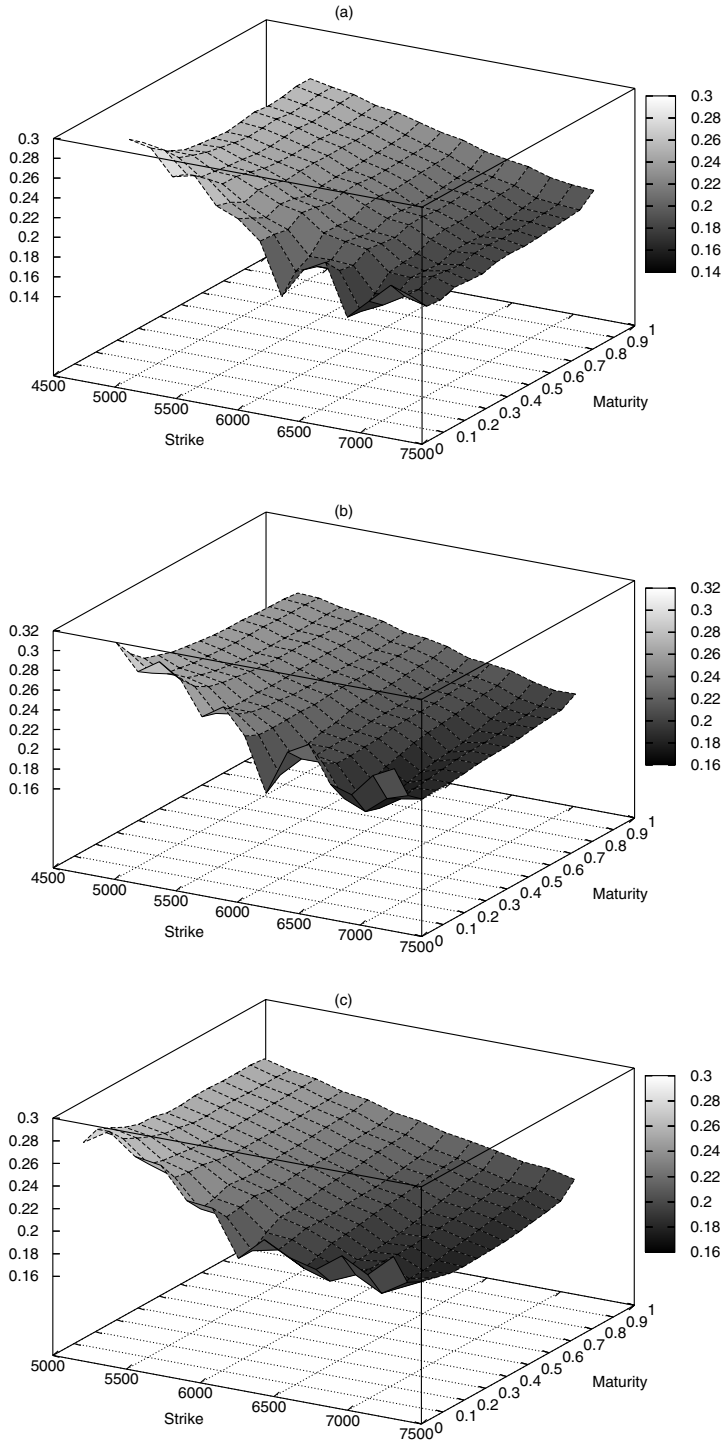

Figure 4. Implied volatility surface at the end of the second computation stage, on the DAX data sets of May 2 (a), 3 (b) and 4 (c), 2001.

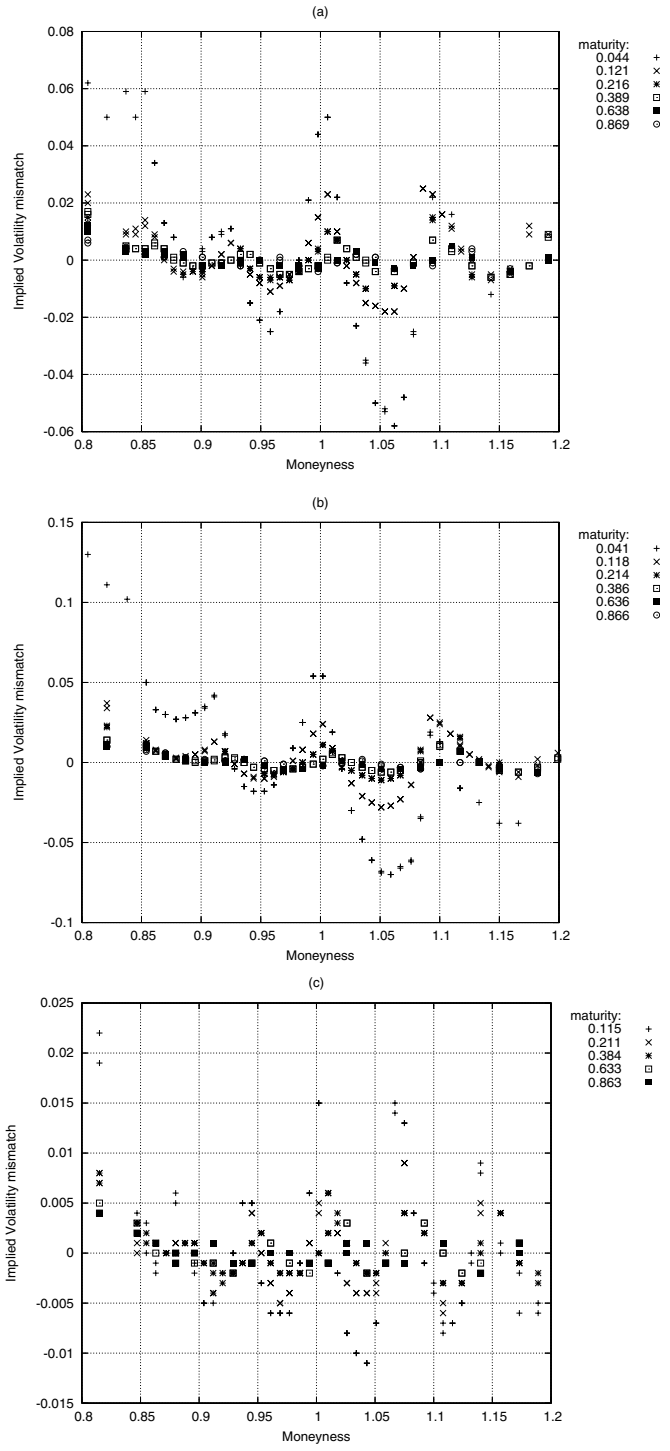

Figure 5. Calibration accuracy at the end of the second computation stage, on the DAX data sets of May 2 (a), 3 (b) and 4 (c), 2001.

Table 2. Execution times on a PC Cluster with *nproc* processors.

$N \times nproc$	1	3	6
54	25 s	9 s	10 s
101	4 min 30 s	1 min 57 s	1 min 36 s

6.3. Stability with respect to the size of the tree

Figures 6–8 illustrate another kind of stability, by showing what happens when one increases the number N of time steps in the tree on the first data set of section 6.2. Observe how the implied volatility surfaces become more regular (figure 7), while the cloud of implied volatility mismatches narrows around 0 (figure 8: beware of the scale of the ordinate axis), as N increases. On a standard serial Pentium PC, execution times corresponding to these experiments vary from a few seconds (section 6.2, $N = 31$) to a few minutes ($N = 101$). Furthermore, table 2 shows the computation times by the MPI-parallel implementation described in section 5.1 on a `PC Cluster` with a variable number of active processors (1.3 GHz each). Notice that, because of the communication times between processors, the optimal number of active processors does not exceed a few units in these experiments.

6.4. Sensitivity with respect to the moneyness of the data

An interesting parameter is the width of the moneyness spread of the input data. In the previous experiments, we used data with moneyness K/S_0 ranging from 0.8 to 1.2. Figure 9 displays, for instance, the results of the calibration on the subset of the options in the first data set of section 6.2 with moneyness contained between 0.9 and 1.1, using $N = 74$ time steps in the tree. The local volatility surface is still much more regular, the implied volatility surface is also more regular, and the calibration is more accurate, than in the analogous experiments with moneyness bounds 0.8 and 1.2 (see parts (b) in figures 6–8).

6.5. Use of a nonconstant prior

If one wishes to reduce the dimension of the minimization program, our algorithm can be combined with a parametrization for the local volatility function, as for instance with splines in Jackson *et al* [17]. To do so, one computes the derivatives of the cost criterion along the basis vectors h of the parametrized family, through the chain rule

$$d\Pi_{T,k}(t_0, y_0; a) \cdot h = \sum_{(t_n, y_m) \in \mathcal{I}} h(t_n, y_m) d\Pi_{T,k}(t_0, y_0; a) \cdot \delta(t_n, y_m).$$

In this manner, we computed, on the DAX data set of 2 May 2001, a solution of the least square unregularized calibration problem in which a is to be found in the subclass of the local volatility functions that are constant on each slice $]T_{i-1}, T_i]$ of a tree with $N = 76$ time steps. The interesting point to note is that, although we obtained a time functional a_0 with values contained between 0.004 34 and 0.0358, this hardly improved the data fit with respect to the starting point in the algorithm, taken as a constant equal to our usual implied vegas weighted average of the input implied volatilities. Indeed the corresponding quadratic residues were equal to 0.002 57 and 0.002 31, respectively, while in the previous calibration experiments, the quadratic residues decreased by one or two orders of magnitude between the beginning and the end of the calibration. This illustrates *a posteriori* the need to look for a with a term and strike structure, as $a \equiv a(t, S)$. A mere term structure for a is unable to explain the volatility smile.

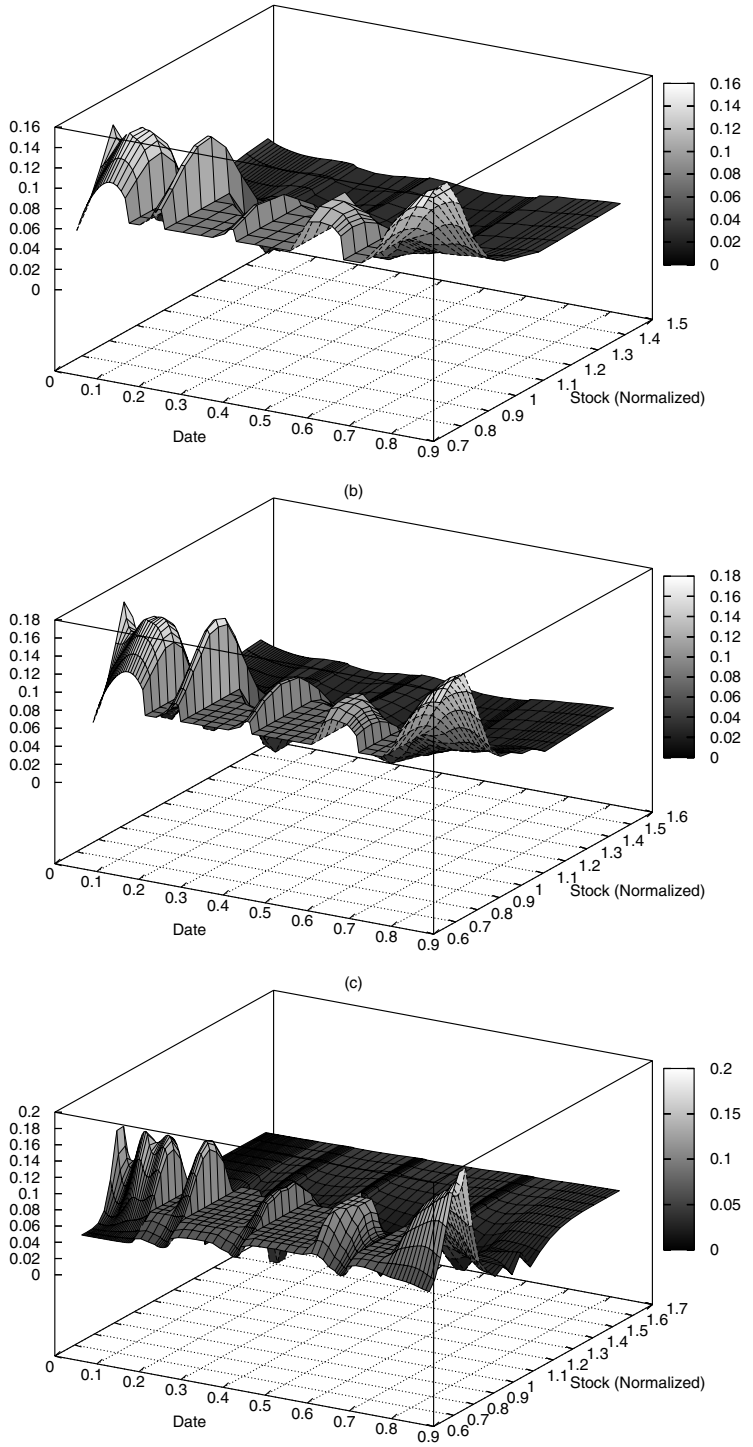

Figure 6. Squared local volatility at the end of the second computation stage, on the DAX data set of 2 May 2001 with $N = 54$ (a), 75 (b) and 101 (c) time steps in the tree.

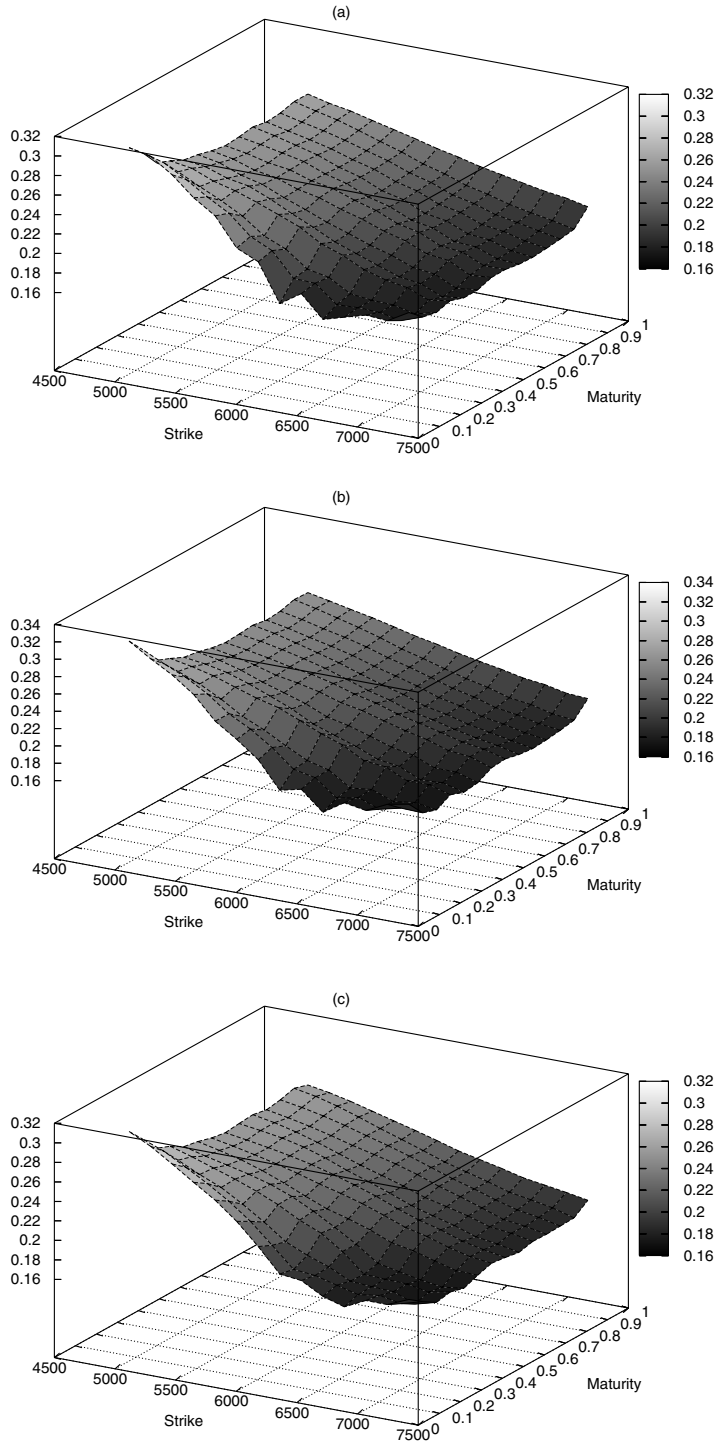

Figure 7. Implied volatility surface at the end of the second computation stage, on the DAX data set of 2 May 2001 with $N = 54$ (a), 75 (b) and 101 (c) time steps in the tree.

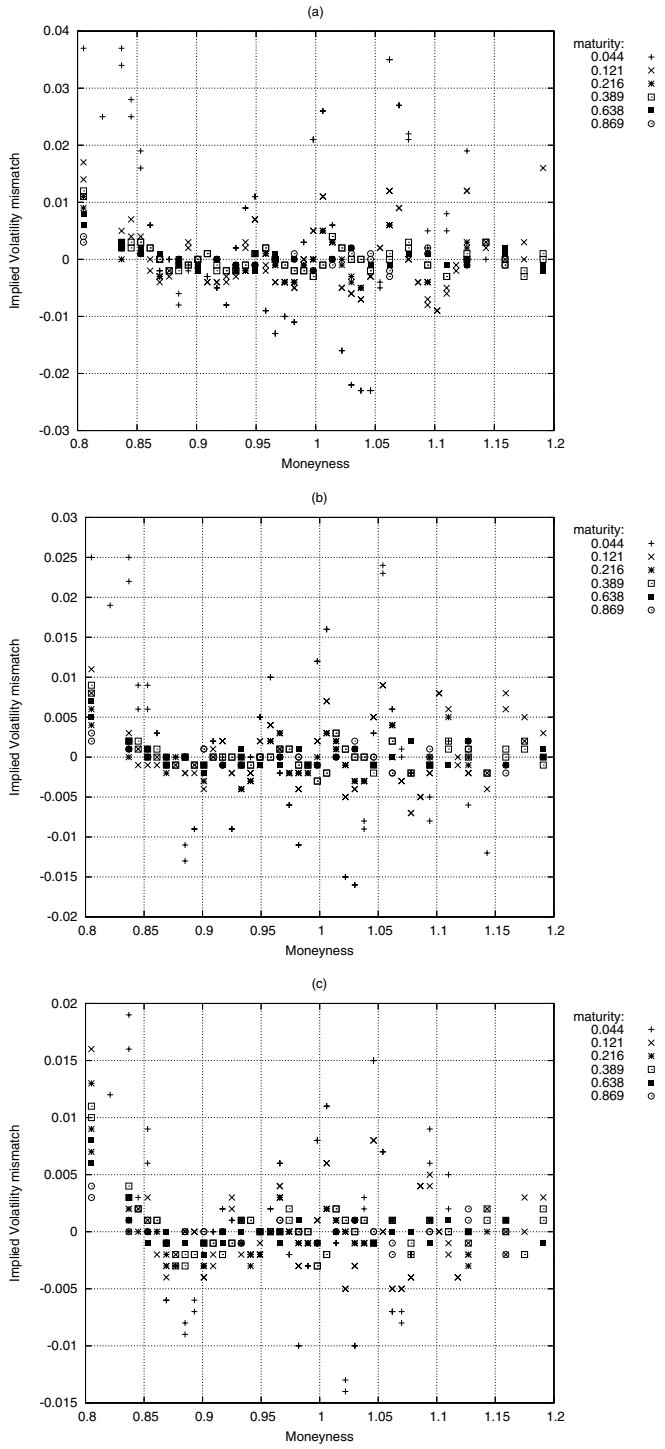

Figure 8. Calibration accuracy at the end of the second computation stage, on the DAX data set of 2 May 2001 with $N = 54$ (a), 75 (b) and 101 (c) time steps in the tree.

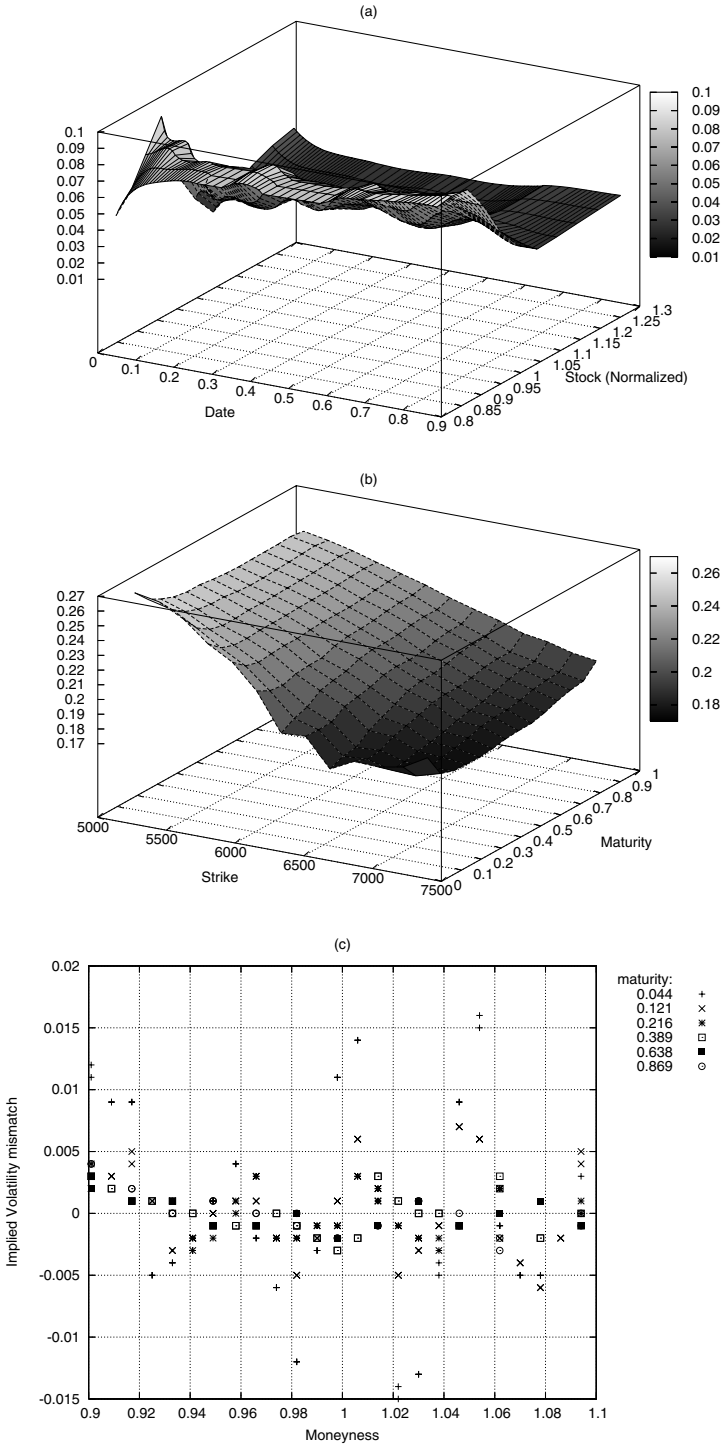

Figure 9. Squared local volatility (a), implied volatility surface (b) and calibration accuracy (c) at the end of the second computation stage, on a filtered DAX data set at 2 May 2001, with $N = 74$ time steps in the tree.

Another interesting feature of our Tikhonov regularized calibration procedure is that any local volatility function a_0 can be used as a prior. Recall that, in the continuous setting of section 3, a_0 was only bound to be measurable and positively bounded. In a context of daily calibration, a good choice for the prior at the current day could thus be the local volatility function calibrated the day before, so as to soften still more the daily variations of the calibrated model. A related further experiment consisted of using the time functional a_0 obtained above as a nonconstant prior on the DAX data set of 2 May 2001. The results are displayed in figure 10. Of course, the resulting local volatility function does not appear as regular as the ones before. Indeed the regularized function is $a - a_0$, which does not ensure that a is regular, for nonconstant a_0. However, the theoretical properties shown previously guarantee that the local volatility function a thus obtained, though irregular, has all the desired stability and convergence features as developed in section 2.

7. Comparison with other methods

7.1. Use of Dupire's formula

Figure 11 displays the local volatility surface obtained by the method of differentiation from discrete data implemented in the `CalibrationEngine` (see section 1.2) on the DAX data sets of May 2 (a), 3 (b) and 4 (c), 2001. The accuracy of the calibration is perfect, by construction. Moreover, the computation is almost instantaneous. But, as expected, the local volatility surface thus calibrated exhibits a highly irregular and unstable behaviour.

7.2. Constrained stochastic control

Figure 12 displays the results obtained on the DAX data set of 2 May 2001, using the constrained stochastic control method implemented in the `CalibrationEngine` (see section 1.2). There were $N = 75$ time steps in the tree. The accuracy of the calibration is still very good and the computation is somewhat faster than by Tikhonov regularization. Indeed the size of the minimization program is proportional to the number of tree nodes in the latter case, while in the former dual reformulation of the problem it is proportional to the number of input data. These numbers are typically of the order of a few thousands and a few hundreds, respectively. But the squared local volatility surface calibrated by the constrained stochastic control method is much less regular that those obtained by Tikhonov regularization. In particular it presents characteristic spikes, with values up to 0.3, in the neighbourhood of the pairs (T, K) with observed prices. By contrast, the maximum value of the squared local volatility obtained by Tikhonov regularization with $N = 75$ time steps in the tree was of the order of 0.18 (see figure 6(b)). Recall that Tikhonov regularization penalizes the fluctuations of the *gradient* of $a - a_0$, instead of those of the *values* of $a - a_0$ for the constrained stochastic control method.

7.3. Advantages and disadvantages of the different methods

The Tikhonov variational regularization method is somewhat slower than the other ones, but, thanks to the reduced cost procedure for the gradient, the computation times are still very acceptable, especially in the parallel implementation. The accuracy of the resulting calibration is very good, still better in general than by the constrained stochastic control method. Above all, Tikhonov regularization appears to be the only method that gives acceptable local volatility surfaces, from the point of view of regularity and stability. It is actually the only method to demonstrate that there is no intrinsic trade-off between stability and accuracy in the calibration problem.

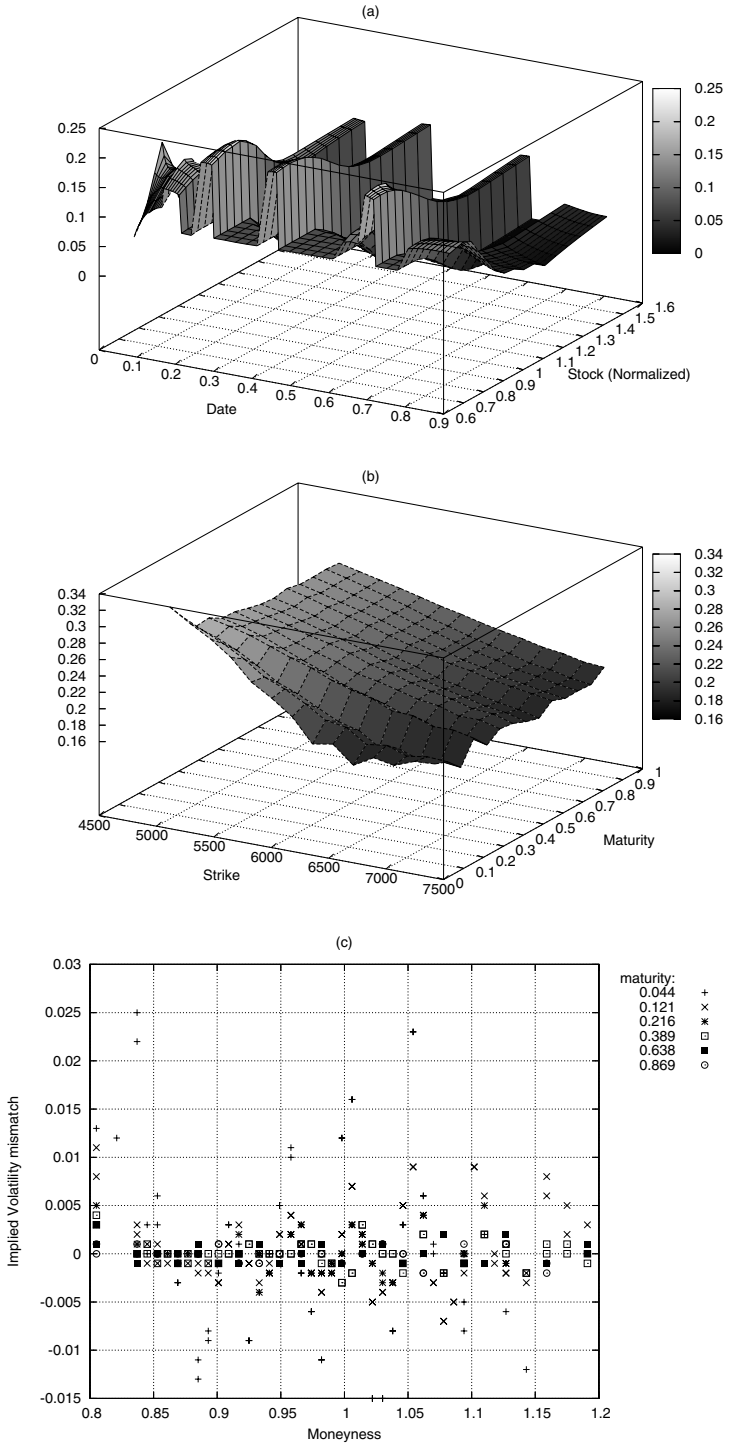

Figure 10. Squared local volatility (a), implied volatility surface (b) and calibration accuracy (c) at the end of the second computation stage, on the DAX data set of 2 May 2001 with nonconstant prior and $N = 76$ time steps in the tree.

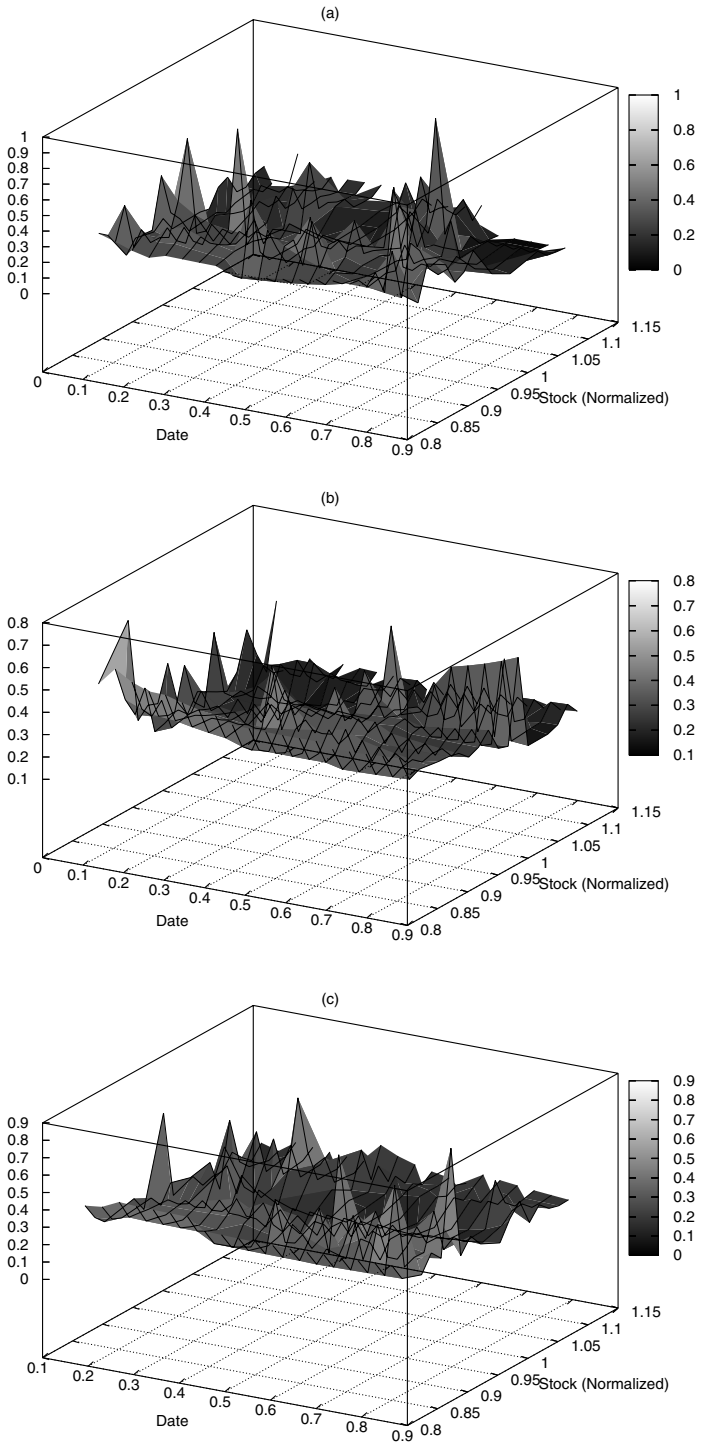

Figure 11. Squared local volatility obtained by application of Dupire's formula on the DAX data sets of May 2 (a), 3 (b) and 4 (c), 2001.

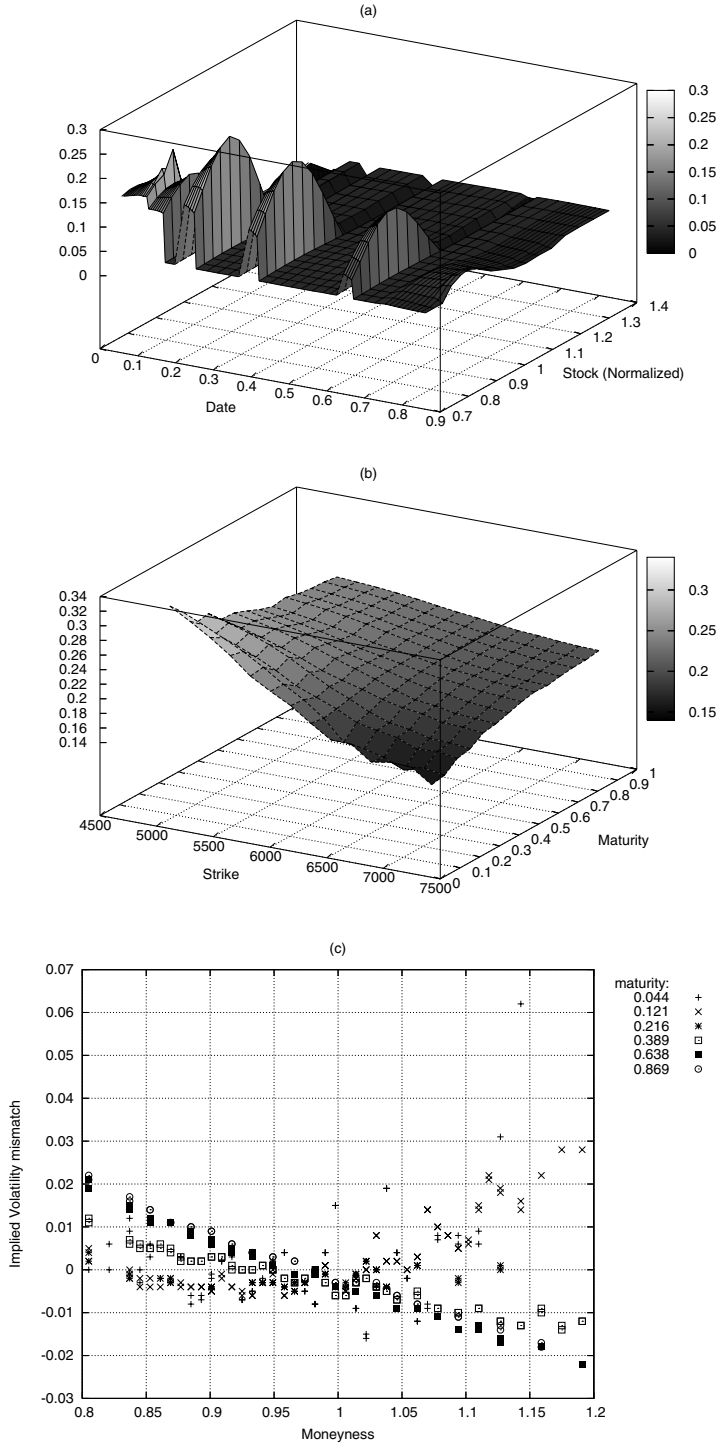

Figure 12. Squared local volatility (a), implied volatility surface (b) and calibration accuracy (c) obtained by the constrained stochastic control method on the DAX data set of 2 May 2001, with $N = 75$ time steps in the tree.

A further distinguishing feature of the variational approach, to which we now come, is the possibility of extending it to the problem of calibration with *American* option prices, at least in the discrete setting, giving rise to an effective algorithm. This is all the more important since, as already mentioned in section 1.2, many of the equity index option markets to which the generalized Black–Scholes model is more specifically dedicated are American options markets.

8. The American calibration problem

8.1. Theoretical results

In the continuous setting of the generalized Black–Scholes model, it is an open question whether or not the theoretical results summed up in section 3 can be extended to the American calibration problem. In fact, given a measurable and positively bounded local volatility function a, we do not even know whether the well known free boundary problem for American call prices (see for instance Huang and Pang [24, problem (3)]) is well posed. Therefore we will move directly to the discrete setting of section 4. Let us denote by $\overline{\Pi}$ the call payoffs, namely

$$\overline{\Pi}_{T,k}(t_n, y_m) \equiv (S_0 \mathrm{e}^{(y_m+(r-q-\underline{a})(t_n-t_0))} - \mathrm{e}^k)^+, \qquad (T,k) \in \mathcal{F}.$$

We thus consider the pair

$$(\Pi, \Pi') \equiv (\Pi_{T,k}(t_n, y_m; a), \Pi'_{T,k}(t_n, y_m; a)), \qquad (T,k) \in \mathcal{F},$$

jointly defined by the terminal condition $\Pi(T, \cdot; a) \equiv \overline{\Pi}(T, \cdot)$, and for $n = N, \ldots, 1$:

(a) $\Pi'(t_{n-1}, \cdot; a)$ defined from $\Pi(t_n, \cdot; a)$ in the same way as $\Pi(t_{n-1}, \cdot; a)$ from $\Pi(t_n, \cdot; a)$ in the European case.
(b) $\Pi(t_{n-1}, \cdot; a) = \max(\Pi'(t_{n-1}, \cdot; a), \overline{\Pi}(t_{n-1}, \cdot))$.

We will then denote by $\mathcal{C}^{n-1}_{T,k}(a)$ the set of all y_m such that $\Pi'_{T,k}(t_{n-1}, y_m; a) \geqslant \overline{\Pi}(t_{n-1}, y_m)$. Moreover, we will qualify the local volatility function a as *regular*, if this inequality is strict for every $n = 1, \ldots, N$ and $y_m \in \mathcal{C}^{n-1}_{T,k}(a)$.

Remark 8.1. Since the non-regular local volatility functions are to be sought among the solutions of one out of a finite number of equations, namely one for each node of $\mathcal{I}$, it is reasonable to expect that most local volatility functions are regular or, more precisely, that the regular volatility functions form a full Lebesgue measure subset of $\mathcal{A}$.

An important difference between European and American call and put option prices is that, in the exercise zone of the American options, their prices are locally constant with respect to the local volatility function. Moreover, the following theorem regarding American option prices shows that these are less regular than the European prices.

Theorem 8.2. *For every $(T,k) \in \mathcal{F}$, the American call or put price $\Pi \equiv \Pi_{T,k}(t_0, y_0; a)$ is Lipschitz continuous and directionally differentiable with respect to the local volatility function $a \in \mathcal{A}$. Moreover, if a is regular, then Π is differentiable and the partial derivative of Π with respect to the value of the local volatility function a at node $(t_n, y_m) \in \mathcal{I}$ with $t_n \leqslant T$ admits the following Feynman–Kac representation:*

$$d\Pi_{T,k}(t_0, y_0; a) \cdot \delta(t_n, y_m) = \Gamma_{T,k}(t_n, y_m; a) \gamma^{T,k}_{t_0,y_0}(t_{n-1}, y_m; a) \exp(-r\tau)\tau, \tag{19}$$

where

$$\Gamma_{T,k}(t_n, y_m; a) \equiv \left(\frac{1}{\varepsilon^2} + \frac{1}{2\varepsilon}\right)\Pi_{T,k}(t_n, y_{m-1}; a) - \frac{2}{\varepsilon^2}\Pi_{T,k}(t_n, y_m; a) + \left(\frac{1}{\varepsilon^2} - \frac{1}{2\varepsilon}\right)\Pi_{T,k}(t_n, y_{m+1}; a),$$

and $\gamma^{T,k}_{t_0,y_0}(t_{n-1}, y_m; a)$ *means the probability, discounted at rate* r*, that the Markov chain* y_n *goes from* (t_0, y_0) *to* (t_{n-1}, y_m) *through* $\mathcal{C}_{T,k}(a)$.

In particular $\gamma^{T,k}_{t_0,y_0}(t_{n-1}, y_m; a)$ is equal to 0 if $y_0 \notin \mathcal{C}^0_{T,k}(a)$ or $y_m \notin \mathcal{C}^{n-1}_{T,k}(a)$.

Proof. Following Huang and Pang [24, theorem 1], Π is Lipschitz continuous and directionally differentiable with respect to a. Huang and Pang work with general θ-schemes and assume a condition to ensure the positive definiteness of a matrix Q. But this is always verified in the particular case of the explicit scheme in which Q is, up to a positive factor, equal to the identity matrix. Moreover, still following Huang and Pang [24, theorem 1], Π is differentiable if a is regular, and the partial derivative

$$\delta\Pi(t_i, y_j) \equiv d\Pi_{T,k}(t_i, y_j; a) \cdot \delta(t_n, y_m); \qquad i = 0, \ldots, n-1,\ j = -i, \ldots, i$$

satisfies the usual European-like dynamic programming equation if $y_j \in \mathcal{C}^i_{T,k}(a)$, with boundary condition $\delta\Pi(t_i, y_j) = 0$ otherwise. Moreover, $\delta\Pi$ satisfies the terminal condition

$$\delta\Pi(t_{n-1}, y_j) = \begin{cases} \mathbf{1}_{\{j=m\}}\Gamma_{T,k}(t_n, y_m; a)\exp(-r\tau)\tau, & y_j \in \mathcal{C}^{n-1}_{T,k}(a) \\ 0, & \text{else.} \end{cases}$$

Now, (19) is nothing but the Feynman–Kac representation for $\delta\Pi(t_0, y_0)$. □

Remark 8.3. As a matter of fact, it is rather obvious, in our trinomial tree setting, that Π is Lipschitz continuous and that Π is differentiable if a is regular, with the dynamic programming characterization as in the proof above for the partial derivatives. This could have been shown directly in a similar way to the analogous statements in the European case (see the theorem 4.1 and its proof), without referring to the more general results of Huang and Pang. Anyway, as in the European case, the *explicit* finite differences, or trinomial tree setting, is interesting because of the probabilistic representation (19) for these derivatives, in which the $\gamma^{T,k}$ can be computed forward in the tree, using Fokker–Planck discrete equations that express the composition of discounted probabilities in $\mathcal{C}_{T,k}(a)$. This results in a reduced cost procedure for these derivatives. Since γ now depends on T and k, one must solve one Fokker–Planck equation *by option*, instead of one Fokker–Planck equation as a whole in the European case. But this is still one or two orders of magnitude faster than solving one equation with source term *by option and mesh node* if one uses the dynamic programming characterization for the partial derivatives directly. Moreover, our computation for the $\gamma^{T,k}$ can be parallelized in the same way as the one for the option prices (see section 5.1).

Since Π is continuous, the stability and convergence theorems 2.5 and 2.9 are applicable (see remark 2.12). But as Π is not everywhere differentiable, we can no longer apply theorem 2.11.

8.2. Calibration algorithm and numerical results

Now the calibration algorithm is the same as in the European case, except that we provide the minimization descent routines with the right-hand side of (19) instead of the Euclidean gradient of Π, knowing that both of them coincide, at the (presumably) full set of regular local volatility functions in $\mathcal{A}$. Moreover, by American call/put *implied volatility*, for given data r, q, t_0, y_0, T, k and price π, we mean the unique (or largest, in the exercise zone) constant volatility such that π is the price of the American call/put in the classic, nonlocal trinomial tree with 200 time steps corresponding to these data. There is indeed no closed pricing formula in the American case. By remark 8.3, the overall computation cost of the calibration does not exceed twice that required in the European case, both in the serial and in the MPI-parallel implementation.

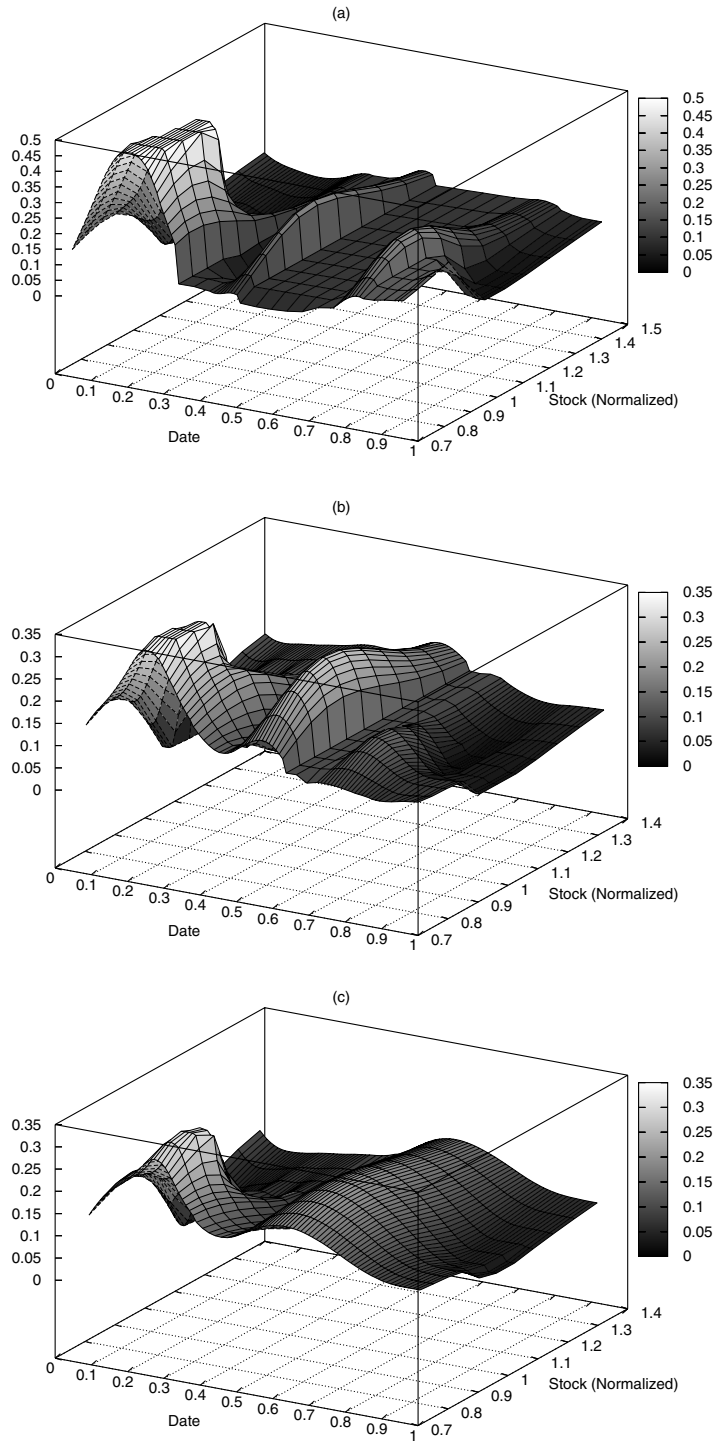

Figure 13. Squared local volatility at the end of the second computation stage, on the FTSE American data sets of January 4 (a), 5 (b) and 6 (c), 1999.

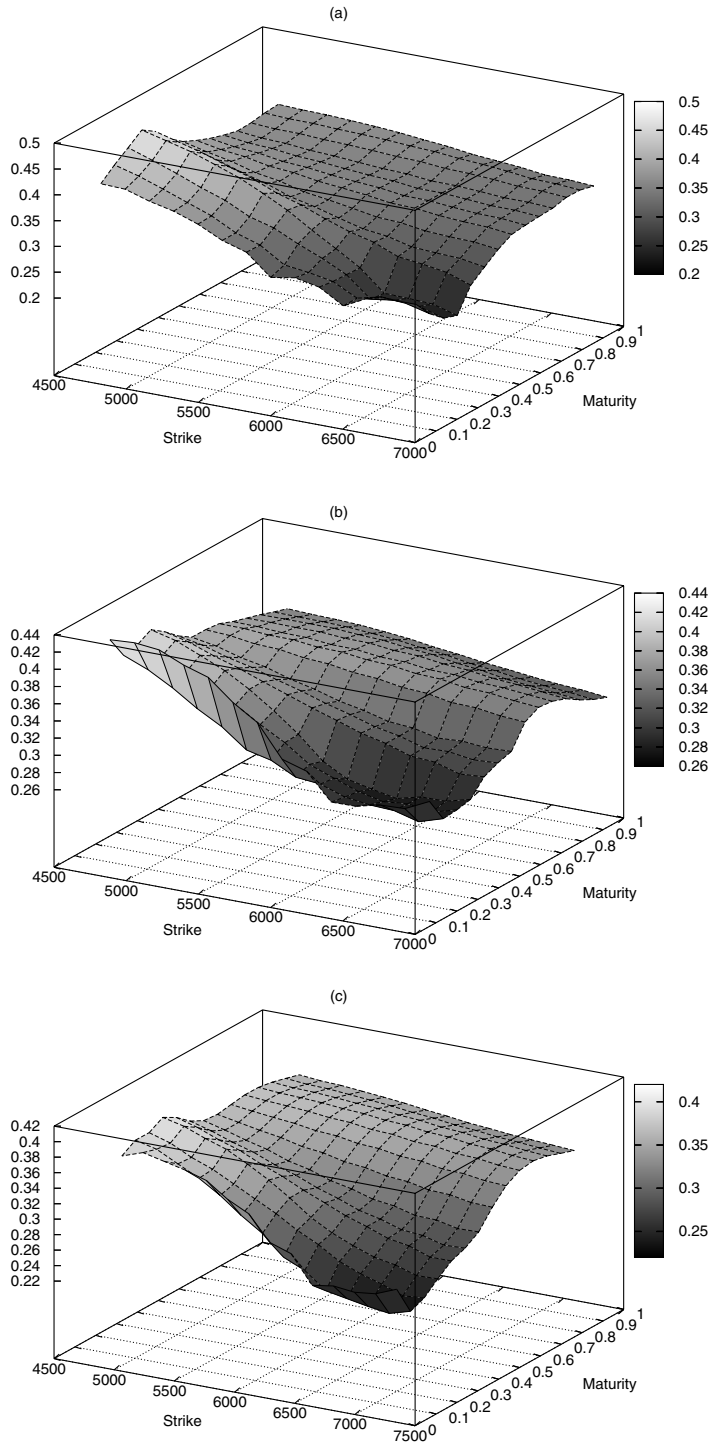

Figure 14. Implied volatility surface at the end of the second computation stage, on the FTSE American data sets of January 4 (a), 5 (b) and 6 (c), 1999.

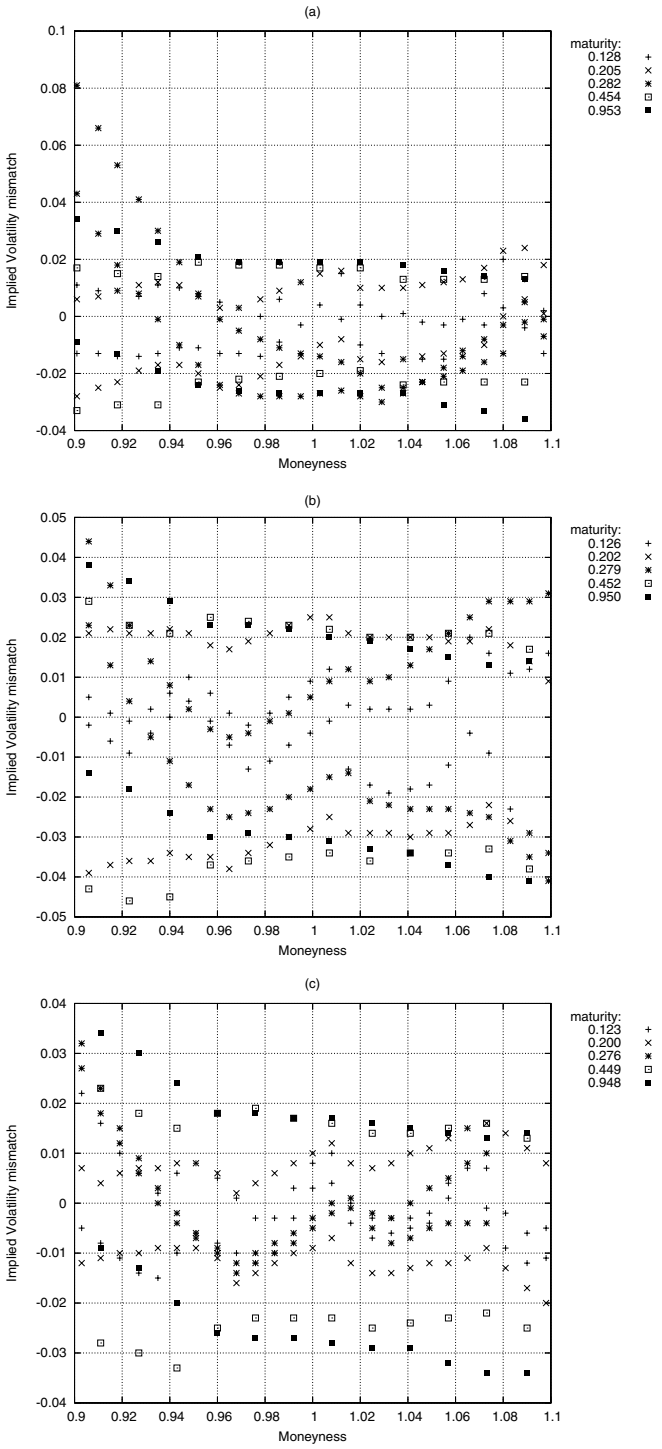

Figure 15. Calibration accuracy at the end of the second computation stage, on the FTSE American data sets of January 4 (a), 5 (b) and 6 (c), 1999.

Remark 8.4. Huang and Pang have proposed a similar heuristic, but without the reduced cost procedure for the partial derivatives, and called it *an implicit programming algorithm* [24, section 4.1]. This consists of a descent method with optimal step, in which the descent step is positively bounded below to ensure that the algorithm does not loop in a nondescent direction. Since there is no reason why the calibration algorithm should drive the local volatility function a outside the full set of regular local volatility functions, this latter precaution seemed a bit superfluous to us. If a problem occurs on a specific data set, it is still possible to perturb the data so as to bring the problem back to a regular case.

Figures 13 and 14 show the local (squared) and implied volatility surfaces obtained on data sets of vanilla American call and put option prices on the FTSE index at January 4 (a), 5 (b) and 6 (c), 1999. We filtered out the options with prices outside the American arbitrage bounds, as well as the less liquid options with moneyness smaller than 0.9 or larger than 1.1. All dividends were taken equal to 0. There were $N = 66$, 90 and 88 time steps in the trees. Figure 15 displays the accuracy of the calibrations. The results are fairly comparable in nature to those obtained in the European case.

Of course nothing guarantees that the subset $\mathcal{E} \subseteq \mathcal{F}$ of the options to be exercised in the market coincides with the one in the calibrated model. This could be guaranteed by adding to the minimization problem the following non-convex constraints:

$$\begin{cases} \Pi_{T,k}(t_0, y_0; a) = \overline{\Pi}_{T,k}(t_0, y_0), & (T,k) \in \mathcal{E} \\ \Pi'_{T,k}(t_0, y_0; a) \geqslant \overline{\Pi}_{T,k}(t_0, y_0), & (T,k) \in \mathcal{F}\backslash\mathcal{E}. \end{cases}$$

9. Conclusion

In Crépey [2, 3], we obtained results regarding stability, convergence and convergence rates for applying Tikhonov regularization to the inverse problem of calibrating a local volatility function from observed vanilla option prices, in a generalized Black–Scholes model. Here, we have shown that the analogous results hold true in a suitable trinomial tree or explicit finite differences discretization. We have also presented efficient serial and parallel implementations of the method in the discrete setting, as well as their extension to the American calibration problem.

Numerical experiments on real data sets involving several hundreds of input prices support strong evidence that there is no trade-off between stability and accuracy when using this regularized calibration procedure, both in the European and in the American case. Except for those corresponding to the shortest maturities, for which calibration is irrelevant, most prices are calibrated up to a few tens of basis points of implied volatility, while the local and implied volatility functions thus calibrated exhibit satisfactory regularity and stability properties. The stability of the results contrasts with the instability that occurs if elementary reconstruction procedures are used for the local volatility function.

Acknowledgments

The author wishes to thank all the Artabel Finance team for helpful insights and fruitful discussions.

Appendix. Explicit finite differences, implicit finite differences and the pricing of vanilla options

We discuss in this appendix a point raised at the end of section 5.1. In the special case of pricing a vanilla call or put option, an explicit scheme may converge as fast in practice as a

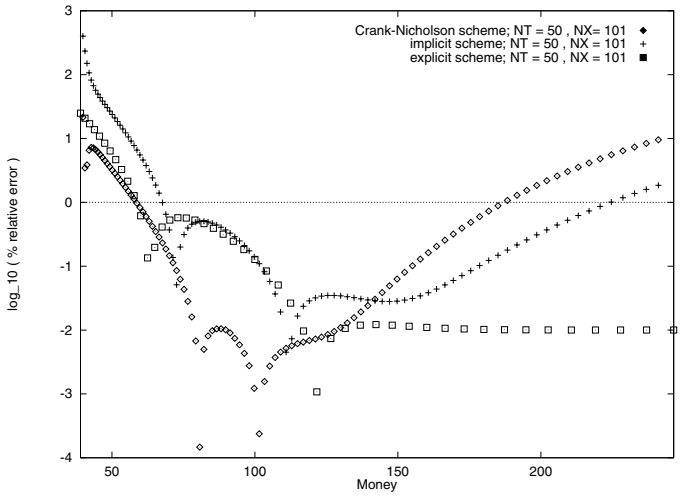

Figure A.1. Error in the pricing of a European call with one year time-to-maturity and strike 100.

Table A.1. Percentage of average calibration error, European case.

N	30	54	72	97
Explicit scheme	1.02	0.66	0.36	0.38
Implicit scheme	0.79	0.63	0.77	1.03

fully implicit one. Figure A.1 (drawn from my PhD thesis [19, p 147]) displays the error in the pricing of a European call in a classic, nongeneralized Black–Scholes model, computed numerically using explicit, fully implicit and Crank–Nicholson θ-schemes. Here by *error* we mean the basis 10 logarithm of the percentage of relative error, computed by using the closed formulae. For instance, the level 0 curve on the figure corresponds to a relative error of 1%. The Crank–Nicholson scheme is more accurate, but this scheme is at the limit of stability with respect to the high frequencies modes entailed by the Dirac masses present in the solutions of our calibration equations.

Further use of the calibrated local volatility function may require more elaborate numerical schemes. This may be the case if one wants to investigate the exercise boundary of an American option, or to deal with barriers. It is then possible to interpolate the local volatility function calibrated in the tree and fit it into another scheme. Table A.1 shows that an implicit scheme obtained in this way is still calibrated to the data. These consist of a set of vanilla European option prices on the DAX index at 2 May 2001. Here by *average calibration error* we mean

$$\frac{\sum_{(T,k)\in\mathcal{F}} |\Pi_{T,k}(t_0, y_0; a) - \pi^{\delta}_{T,k}|}{\sum_{(T,k)\in\mathcal{F}} \pi^{\delta}_{T,k}},$$

where $\Pi_{T,k}(t_0, y_0; a)$ is the price in the tree with local volatility function a and N time steps, or the price in the implicit scheme with 100 time steps, 200 space steps and a local volatility function interpolated from a. By interpolated value for a at the mesh node (t, y) in the implicit scheme, we have taken the value of a at a node (t_n, y_m) in the tree as close as possible to (t, y). Table A.2 displays the analogous results for data consisting of American vanilla option prices in the FTSE index at 4 January 1999.

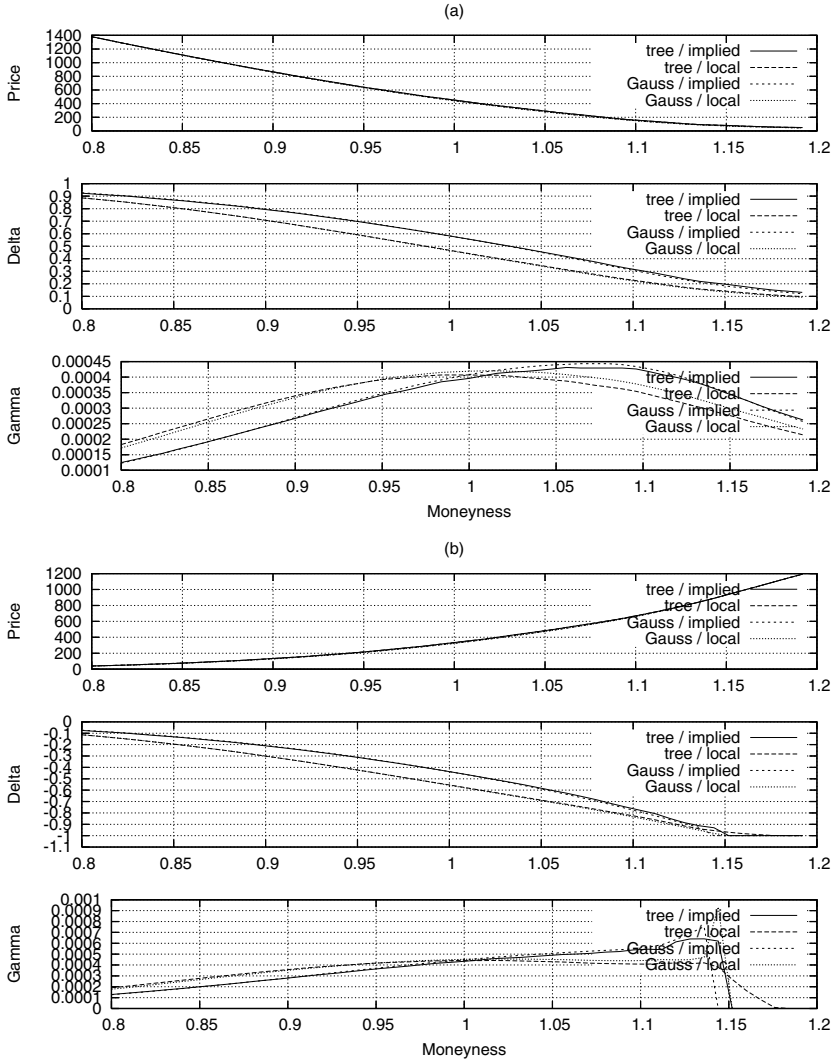

Figure A.2. Price and Greeks of options with half-year time-to-maturity: (a) European call, (b) American put.

Table A.2. Percentage of average calibration error, American case.

N	41	72	108
Explicit scheme	5.01	4.69	4.50
Implicit scheme	4.79	4.84	7.28

Moreover, figure A.2(a) displays the price, delta and gamma of a European call with a half-year time-to-maturity and variable moneyness K/S_0, computed in four different ways:

(a) Using a trinomial tree with local volatility function a and $N = 97$ time steps, calibrated on the set of vanilla option prices on the DAX index at 2 May 2001.

(b) Using an implicit scheme with 100 time steps and 200 space steps and a local volatility function interpolated from a as above.
(c) and (d) Computing the implied Black–Scholes volatilities corresponding to the prices obtained at points (a) and (b) and using the closed Black–Scholes formulae to obtain the associated Greeks.

Therefore points (a) and (b) correspond to computations in calibrated local volatility models, while points (c) and (d) correspond to a common practice, whereby the smile on option markets is handled by using the Greeks of the implied volatilities of the options. These points (a)–(d) correspond to the curves labelled on the figure as tree/local, Gauss/local, tree/implied and Gauss/implied, respectively. Figure A.2(b) displays the analogous results for an American put option, where a classic, nonlocal trinomial tree with 200 time steps is used to compute the implied Black–Scholes volatilities and the associated Greeks. There are indeed no closed pricing formulae in the American case. One can observe that the prices, deltas and gammas obtained at points (a) and (b) are very close to one another, as are also the ones obtained at points (c) and (d). In contrast, there is a clear discrepancy between the local deltas and gammas obtained at points (a) and (b) on the one hand, and the implied deltas and gammas obtained at points (c) and (d) on the other.

References

[1] Lagnado R and Osher S 1997 A technique for calibrating derivative security pricing models: numerical solution of an inverse problem *J. Comput. Finance* **1** 13–25
[2] Crépey S 2002 Tikhonov regularization and calibration of a local volatility in finance—stability, convergence and convergence rates issues *CMAP Internal Research Report* no 474
[3] Crépey S 2002 Calibration of the local volatility in a generalized Black–Scholes model using Tikhonov regularization *SIAM J. Math. Anal.* accepted
[4] Black F and Scholes M 1973 The pricing of options and corporate liabilities *J. Pol. Econ.* **81** 637–59
[5] Dupire B 1994 Pricing with a smile *Risk* **7** 18–20
[6] Derman E and Kani I 1994 Riding on a smile *Risk* **7** 32–9
[7] Derman E, Kani I and Chriss N 1996 Implied trinomial trees of the volatility smile *J. Derivatives* **4** 7–22
[8] Rubinstein M 1994 Implied binomial trees *J. Finance* **69** 771–818
[9] Andersen L and Brotherton-Ratcliffe R 1997 The equity option volatility smile: an implicit finite difference approach *J. Comput. Finance* **1** 5–37
[10] Dumas B, Fleming J and Whaley R 1998 Implied volatility functions: empirical tests *J. Finance* **53** 2059–106
[11] Coleman T, Li Y and Verma A 1999 Reconstructing the unknown volatility function *J. Comput. Finance* **2** 77–102
[12] Rebonato R 1999 *Volatility and Correlation in the Pricing of Equity, FX and Interest-Rate Options* (New York: Wiley)
[13] Bouchouev I and Isakov V 1999 Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets *Inverse Problems* **15** R95–116
[14] Bouchouev I, Isakov V and Valdivia N 2002 Recovery of volatility coefficient by linearization *Quant. Finance* **2** 257–63
[15] Avellaneda M, Friedman C, Holmes R and Samperi D 1997 Calibrating volatility surfaces via relative-entropy minimization *Appl. Math. Finance* **41** 37–64
[16] Samperi D 2002 Calibrating a diffusion pricing model with uncertain volatility: regularization and stability *Math. Finance* **12** 71–87
[17] Jackson N, Süli E and Howison S 1999 Computation of deterministic volatility surfaces *J. Comput. Finance* **2** 5–32
[18] Bodurtha J and Jermakyan M 1999 Nonparametric estimation of an implied volatility surface *J. Comput. Finance* **2** 29–60
[19] Crépey S 2001 Contribution à des méthodes numériques appliquées à la Finance et aux Jeux Différentiels *PhD Thesis* École Polytechnique, France
[20] Berestycki H, Busca J and Florent I 2002 Asymptotics and calibration of local volatility models *Quantum Finance* **2** 61–9

[21] Achdou Y and Pironneau O 2002 Volatility smile by least square *Int. J. Theor. Appl. Finance* at press
[22] Chriss N 1996 Transatlantic trees *Risk Mag.* **9** 45–8
[23] Barle S and Cakici N 1998 Growing a smiling tree *Hedging with Trees: Advances in Pricing and Risk Managing Derivatives* ed M Broadie and P Glasserman (RISK Books) pp 173–8
[24] Huang J and Pang J 2000 A mathematical programming with equilibrium constraints approach to the implied volatility surface of American options *J. Comput. Finance* **4** 21–56
[25] Derman E and Kani I 1998 Stochastic implied trees: arbitrage pricing with stochastic term and strike structure of volatility *Int. J. Theor. Appl. Finance* **1** 61–110
[26] Andersen L and Andreasen J 2000 Jump-diffusion processes: volatility smile fitting and numerical methods for option pricing *Rev. Derivatives Res.* **4** 231–62
[27] Cont R and Tankov P 2002 Non-parametric calibration of jump-diffusion models *CMAP Internal Research Report* no 490
[28] Derman E 1999 Regimes of volatility *Risk Mag.* **4** 55–9
[29] Cont R and da Fonseca J 2002 Dynamics of implied volatility surfaces *Quant. Finance* **2** 45–60
[30] Tikhonov M 1963 Regularization of incorrectly posed problems *Sov. Math.* **4** 1624–7
[31] Engl H W, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (Dordrecht: Kluwer)
[32] Engl H W, Kunisch K and Neubauer A 1989 Convergence rates for Tikhonov regularisation of nonlinear ill-posed problems *Inverse Problems* **5** 523–40
[33] Binder A, Engl H W, Groetsch C W, Neubauer A and Scherzer O 1994 Weakly closed nonlinear operators and parameter identification in parabolic equations by Tikhonov regularization *Appl. Anal.* **55** 13–25