www.oeaw.ac.at

Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

Tikhonov Regularization Applied to the Inverse Problem of Option Pricing: Convergence Analysis and Rates

H. Egger, H.W. Engl

RICAM-Report 2004-10

www.ricam.oeaw.ac.at

Tikhonov Regularization Applied to the Inverse Problem of Option Pricing: Convergence Analysis and Rates‡

Herbert Egger[1] **and Heinz W. Engl**[2]

[1] Spezialforschungsbereich F013, Johann Radon Institute for Computational and Applied Mathematics, Altenbergerstr. 69, A-4040 Linz, Austria
[2] Johann Radon Institute for Computational and Applied Mathematics and Johannes Kepler University Linz, Altenbergerstr. 69, A-4040 Linz, Austria

E-mail: `herbert.egger@oeaw.ac.at, heinz.engl@jku.at`

Abstract. This paper investigates the stable identification of local volatility surfaces $\sigma(S,t)$ in the Black-Scholes/Dupire equation from market prices of European Vanilla options. Based on the properties of the parameter-to-solution mapping, which assigns option prices to given volatilities, we show stability and convergence of approximations gained by Tikhonov regularization. In case of a known term-structure of the volatility surface, in particular if the volatility is assumed to be constant in time ($\sigma(S,T)=\sigma(S)$), we prove convergence rates under simple smoothness and decay conditions on the true volatility. The convergence rate analysis sheds light onto the importance of an appropriate *a-priori* guess for the unknown volatility and the nature for the ill-posedness of the inverse problem, caused by smoothing properties and the nonlinearity of the direct problem. Finally, the theoretical results are illustrated by numerical experiments.

1. Introduction

Financial derivatives are contracts between two or more parties, where payment is derived from some underlying asset like a stock, bond, commodity, interest or exchange rate. In the famous Black-Scholes model [3], the stochastic behaviour of the underlying asset S is modeled by a geometric Brownian motion

$$dS(t) = \mu S dt + \sigma S dW(t). \tag{1.1}$$

Here, W_t denotes a standard Wiener process. The parameters μ and σ are called drift rate and *volatility* of the underlying asset. By Ito's rule [20], the stochastic behaviour of a derivative security $V(S,t)$ is governed by the stochastic differential equation

$$dV(S,t) = \left(V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + \mu S V_S\right) dt + \sigma S dW(t). \tag{1.2}$$

In their seminal paper [3], Black and Scholes constructed a dynamic portfolio consisting of a derivative security and a variable amount of the underlying completely eliminating risk in

‡ supported by the the Austrian National Science Foundation FWF under grant SFB F013/08.

(1.2). In the absence of arbitrage the instantaneous return of this portfolio equals the return of a riskless investment, denoted by the interest rate r, which together with (1.2) yields the famous Black-Scholes equation

$$V_t + \frac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S = rV. \tag{1.3}$$

As an important consequence of no-arbitrage arguments (see, e.g., [28]), the drift rate μ does not enter (1.3). The payoff, i.e. the amount of money received when exercising the option at maturity, enters via the terminal condition, e.g.,

$$V^{K,T}(S, t = T) = \max(S - K, 0) \tag{1.4}$$

for a European call option with strike K and maturity T. In the case of constant parameters the value of a European call option is then given by (the Black-Scholes formula)

$$C = S_0\mathcal{N}(d_1) - e^{-rT}K\mathcal{N}(d_2)$$

with

$$d_1 = \frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}, \qquad d_2 = d_1 - \sigma\sqrt{T}.$$

The price of a European call option as a function of volatility $C(\sigma)$ is strictly monotone and the unique value of σ corresponding to a given option price $C(K, T; S, t)$ for strike K and maturity T, current stock price S and time t, is called (Black-Scholes-) *implied volatility*. In contrast to the constant parameter assumptions of the Black-Scholes model, implied volatilities show a distinct dependence on strike and maturity, which is referred to as the *smile*- respectively *term-structure*. One possibility to explain the volatility smiles in the (extended) Black Scholes world is to use a deterministic volatility function $\sigma(S, t)$ in (1.1). For other generalizations like stochastic volatility and jump-diffusion models, see, e.g., [15, 24]).

Once, $\sigma(S, t)$ and r have been specified a variety of options can be priced using equations similar to (1.3).

Volatility identification in extended Black-Scholes models, also referred to as calibration to market prices, has recently been investigated by several authors, e.g., in [5, 6, 14, 19, 22, 23] and it has been observed to be an ill-posed problem in the sense that reconstruction of local volatility is unstable with respect to errors in the data.

In [19, 22] emphasis has been put on certain aspects of a numerical implementation. In both papers, motivated by heuristic arguments, Tikhonov regularization is applied to stabilize the inverse problem. A different, computationally cheap estimation of local volatilities has recently been proposed in [13].

The authors of [5] consider a linearized problem and derive a uniqueness result for the purely state dependent case $\sigma = \sigma(S)$. Uniqueness for state dependent volatilities and the relation of optimal control problems corresponding to time discrete and time continuous observation are investigated in [23]. In [6], a quite general existence result for the pricing equation (1.3) is derived and compactness of the corresponding solution operator is shown for an reasonable choice of function spaces. Stability and convergence of the Tikhonov-regularized solutions follows directly by applying the standard convergence theory of Tikhonov regularization for non-linear inverse problems [9].

The case of purely time dependent volatilities is extensively studied in [14], and the ill-posedness of the inverse problem is proven.

In the first part of this work, we consider the operator connecting volatilities with corresponding option prices (the *parameter-to-solution map*) and investigate its properties in detail. Then we state the *inverse problem of option pricing* and discuss its stable solution via Tikhonov regularization and convergence of the approximations for various observation spaces.

In the second part, we focus on derivation of convergence rate results: The question of convergence rates of Tikhonov regularization applied to the inverse problem of option pricing has been investigated in [6] only in the general framework of [9], Section 10 so far; in particular, no interpretable sufficient conditions for application of the convergence rate result of [9] have been given. In Section 4, we prove convergence rates for the case of time-independent volatilities $\sigma = \sigma(S)$ under simple smoothness assumptions. The results can be generalized to a class of local volatility functions, where the term structure is assumed to be known, more precisely, $\sigma(S,t) = \sigma(S)\rho(t)$ and the function $\rho(t)$ is known. As an implication of the convergence rates considerations we obtain uniqueness of solutions to the inverse problem under the given assumptions. Furthermore, we show that it is possible to weaken the assumptions in the case of observations of first derivatives of the data (which amount to prices of digital call options in our case, which are also observable on the markets).

Finally, we discuss some aspects of an efficient implementation and present numerical tests, which confirm the theoretical results.

2. The inverse problem of option pricing

The inverse problem of option pricing under consideration is the identification of a local volatility surface $\sigma(S,t)$ such that the solutions $C(S,t;K,T) = C^{K,T}(S,T)$ of

$$C_t + \frac{1}{2}\sigma^2 S^2 C_{SS} + (r-q)\, S\, C_S - r\, C = 0 \tag{2.1}$$

$$C(S,T;K,T) = (S-K_i)^+, \qquad \text{for} \quad S > 0. \tag{2.2}$$

match quoted market prices $C^*(K,T)$. Each observation is linked with the solution of equation (2.1) with different terminal conditions (2.2). Under the assumption that the interest and dividend yield r, q are functions of time only, $C(S,t;K,T)$ also satisfies the following dual (Dupire-) equation (see [5, 7])

$$-C_T + \frac{1}{2}\,\sigma^2(K,T)\, K^2 C_{KK} + (q-r)\, K\, C_K - q\, V = 0 \tag{2.3}$$

with initial value

$$C(S,t;K,T=t) = (S-K)^+, \qquad \text{for} \quad K > 0. \tag{2.4}$$

The substitutions $K = e^y$, $\tau = T - t$, $u(y,\tau) = e^{\int_t^T q(s)\,ds}\, C(K,T)$, $a(y,\tau) = \frac{1}{2}\sigma^2(K,\tau)$ and $b(\tau) = q(\tau) - r(\tau)$ yield

$$-u_\tau + a(y,\tau)\,(u_{yy} - u_y) + b(\tau)\, u_y = 0 \tag{2.5}$$

$$u(y,0) = (S^* - e^y)^+. \tag{2.6}$$

Next, we show well-posedness of the direct problem, i.e. stable solution of (2.5), (2.6) for the following set of admissible parameters:

Definition 2.1 *Let* $\underline{a}, \overline{a} > 0$ *and* $\underline{a} \le a^* \le \overline{a}$ *with* $\nabla a^* \in (L_2(Q))^2$ *and* $Q := \mathbb{R} \times (0,T)$. *Define*

$$\mathcal{K}(a^*) := \{a \in a^* + W_2^1(Q) : \underline{a} \le a \le \overline{a}\}. \tag{2.7}$$

By $W_p^{2,1}$ we denote the usual Sobolev spaces, cf. [1]. The condition $a \in \mathcal{K}(a^*)$ is equivalent to σ being uniformly bounded from above and below and $\nabla\sigma \in (L_2(Q))^2$.

The main properties of the parameter to solution map $F : a \to u(a)$, where $u(a)$ is the solution of (2.5), (2.6) are summarized in the following:

Theorem 2.1 *There exist a $p^* > 2$ such that $F : \mathcal{K}(a^*) \to W^{2,1}_{p,loc}(Q)$ is continuous and compact for $2 \le p < p^*$. Moreover, F is weakly (sequentially) continuous and thus weakly closed.*

The proof can be found in the Appendix.

3. Stability and convergence of Tikhonov regularization

Continuity, compactness and weakly closedness of the parameter to solution map $F : \mathcal{K}(a^*) \to W^{2,1}_{2,loc}(Q)$ immediately imply local ill-posedness (see [17]) of the inverse problem of option pricing: Let a^n be a bounded sequence in $\mathcal{K}(a^*)$ that has no convergent subsequence. Since a^n is bounded, it has a weakly convergent subsequence a^{n_k}. Now by compactness and weak closedness of the operator F, the sequence $F(a^{n_k})$ converges. This means that similar option values may be linked to completely different volatilities, i.e., identification of a volatility surface corresponding to perturbed option prices is ill-posed.

For a mathematical analysis of the inverse problem, we assume that continuous observation of option prices $u(y,\tau)$ is available on Q (or at least on a strip $Q_\Delta := \mathbb{R} \times (T-\Delta, T)$ if $a(y,t) = a(y)$ does not depend on time). Stability and convergence results are stated for $L_2(Q)$ as observation space and $\mathcal{K}(a^*)$ as set of admissible parameters. Nevertheless, the convergence analysis below carries over to the case of a large class of observation spaces, as long as $L_2(Q)$ can be continuously embedded. In particular the results are applicable to discrete observations and state-continuous observations for certain maturities T_i.

We assume in the sequel that the "real" data $u^\dagger$ are attainable, i.e., there exists an $a^\dagger \in \mathcal{K}(a^*)$ with $u^\dagger = u(a^\dagger)$, and that an upper bound δ for the noise level

$$\|u^\dagger - u^\delta\| \le \delta. \tag{3.1}$$

of the observation is known *a-priori.* It is well-known that regularization without such a knowledge is not possible in general (cf. [2, 9]).

For a stable solution of the inverse problem, we follow the standard Tikhonov approach for regularization of nonlinear ill-posed problems: An approximate solution a^δ_β of the inverse problem $u(a) := F(a) = u^\delta$, is obtained by minimizing the Tikhonov functional (cf., e.g., [9])

$$J^\delta_\beta(a) = \|u(a) - u^\delta\|^2_{L_2(Q)} + \beta\,\|a - a^*\|^2_{W^1_2(Q)}. \tag{3.2}$$

The applicability of the standard theory for Tikhonov regularization (see, e.g. [9]), is granted by Theorem 2.1.

Theorem 2.1 naturally extends to more general observations, as long as a continuous embedding of $W^{2,1}_{2,loc}(Q)$ is available, i.p. to continuous observation in time/space and discrete observations in space/time and to fully discrete data. The standard results for Tikhonov regularization for nonlinear ill-posed problems [9] yield for fixed $\beta > 0$ existence of a minimizer, stability of solutions a^δ_β and convergence as $\beta \sim \delta \to 0$. We summarize the main results and refer to [9] for details and proofs. The first result states well-posedness of the regularized ($\beta > 0$) problems:

Proposition 3.1 (Stability) *Let $\beta > 0$ and $\{a_k\}$, $\{u_k\}$ be sequences where $u_k \to u^\delta$ and a_k is a minimizer of (3.2) with u^δ replaced by u_k. Then there exists a convergent subsequence of $\{a_k\}$ and the limit of every convergent subsequence is a minimizer of (3.2).*

The second result shows that the regularization parameter β can be chosen in dependence of the noise level δ such that minimizers of the approximating problems converge (in a set valued sense) to an a^*-minimum-norm-solution, i.e., a least squares solution of $F(a) = u^\delta$ with minimal distance to a^* (see, e.g.,[9]).

Proposition 3.2 (Convergence) *Let $\|u - u^\delta\| \le \delta$ and $\beta(\delta)$ be such that $\beta(\delta) \to 0$ and $\delta^2/\beta(\delta) \to 0$. Then every sequence $\{a^{\delta_k}_{\beta_k}\}$, where $\delta_k \to 0$, $\beta_k := \beta(\delta_k)$ and $a^{\delta_k}_{\beta_k}$ is a solution of (3.2) has a convergent subsequence. The limit of every convergent subsequence is an a^*-minimum-norm-solution. If the a^*-minimum-norm-solution $a^\dagger$ is unique, then*

$$\lim_{\delta\to 0} a^\delta_{\beta(\delta)} = a^\dagger.$$

It shall be mentioned, that the above general convergence theory for Tikhonov regularization for nonlinear ill-posed problems in particular applies to the case of discrete observations. However, the corresponding convergence results are rather weak, as they guarantee only convergence of subsequences to some a^*-minimum-norm-solution, which, in case of finitely many data, will not be unique in general. To study the structure of the inverse problem in more detail, we will investigate the rate of convergence under additional assumptions below.

4. Convergence rates

In general convergence stated in proposition 3.2 will be arbitrarily slow. A rate of convergence can only be derived under additional assumptions, e.g. in [9] a convergence rate of

$$\|a^\delta_\beta - a^\dagger\| = \mathcal{O}(\sqrt{\delta}) \quad \text{and} \quad \|F(a^\delta_\beta) - u^\delta\| = \mathcal{O}(\delta) \tag{4.1}$$

is proven for general inverse problems $F(a) = u^\delta$ under the following assumptions:

(i) F is Fréchet differentiable and the derivative satisfies a Lipschitz condition

$$\|F'(a) - F'(a^\dagger)\| \le \gamma \, \|a - a^\dagger\| \tag{4.2}$$

(ii) A source condition

$$a^* - a^\dagger = F'(a^\dagger)^* w \tag{4.3}$$

holds, with $\|w\|$ sufficiently small.

For the inverse problem of option pricing under consideration, the Fréchet differentiability of the parameter-to-solution map F follows by

Proposition 4.1 *Let $F : \mathcal{K}(a^*) \to W^{2,1}_{p,loc}(Q)$, $2 \le p < p^*$ be the parameter-to-solution map, i.e. $u(a) = F(a)$ satisfy (2.5), (2.6). Then F is Fréchet-differentiable and satisfies a Lipschitz condition (4.2).*

Proof. We set for simplicity $r = q = 0$. By linearity, the directional derivative in direction p, $u'h$, satisfies

$$-(u'h)_\tau + a\left((u'h)_{yy} - (u'h)_y\right) = -h\left(u_{yy}(a) - u_y(a)\right)$$

with homogeneous initial condition. Proposition A.1 yields existence and uniqueness of a solution $u'h \in W^{2,1}_p(Q)$ for $2 \le p < p^*$. Note, that $h \in W^1_2(Q) \Rightarrow h \in L_q$ for $1 \le q < \infty$. Linearity of $u'h = F'(a)h$ in h and continuity in a imply Fréchet differentiability. By the same arguments we get $w := (v'h - u'h) = (F'(b) - F'(a))h$ solves

$$-w_\tau + a\left(w_{yy} - w_y\right) = -h\left((v-u)_{yy} - (v-u)_y\right) - (b-a)\left((v'h)_{yy} - (v'h)_y\right)$$

and hence by Proposition A.1 $\|w\|_{W^{2,1}_p} \le c\|h\|_{W^2_1}\|(b-a)\|_{W^2_p}$. ■

For continuous observation in $L_2(Q)$ and time and space dependent volatility $a \in \mathcal{K}(a^*)$, the source condition (4.3) requires the existence of a source function w, such that

$$\int_Q p(u_{yy} - u_y) W d(y,t) = (a^\dagger - a^*, p)_1, \qquad \forall p \in H^1(Q),$$

where W is a solution to the adjoint equation

$$W_t + (aW)_{yy} + (aW)_y = w,$$

with homogeneous boundary and terminal condition (cf., e.g., [6]). A similar representation holds in the case of purely space dependent volatility and one maturity. It is not clear how to derive reasonable sufficient conditions for the existence of such a source-function w.

In order to derive convergence rates under interpretable conditions, we use a problem adapted theory, cf. [10] in the context of parameter identification in heat conduction: First, note that data are usually very sparse in time direction (e.g. four maturities per year), and one cannot expect to identify the term-structure of volatility accurately from such data. Therefore, we make an additional assumptions on the term structure of volatility, e.g. $\sigma = \sigma(K)$ or, more generally, $\sigma(K,T) = \rho(T)\sigma(K)$ (see [4]), where in the second case the term structure $\rho(T)$ is assumed to be known. For brevity, we restrict ourselves to case of one maturity and time independent volatility (cf. e.g. [4, 5, 23]) in the sequel and assume continuous observations on an arbitrarily small time interval $(T-\Delta, T)$. For $a \in \mathcal{K}_2(a^*) := \{q \in \mathcal{K}(a^*) : q - a^* \in W_2^2(\mathbb{R}), \quad a^* := a^*(y)\}$, the Tikhonov functional now reads

$$J_\beta^\delta(a) = \int_{Q_\Delta} |u(a) - u^\delta|^2 d(y,t) + \beta \, \|a - a^*\|^2_{W_2^2(\mathbb{R})}, \tag{4.4}$$

where $u(a)$ is solution of (2.5), (2.6). Here $Q_\Delta = \mathbb{R} \times (T-\Delta, T)$. Stability for $\beta > 0$ with respect to data noise and convergence are guaranteed by Proposition 3.1, 3.2. Following the ideas in [10], we derive the convergence rates (4.1) under interpretable conditions.

Lemma 4.1 *(compare Th. 6.1. in [10])*
Let $a^\dagger \in \mathcal{K}_2(a^*)$ and $u = u(a^\dagger)$. Additionally, let $a_\beta^\delta \in K_2(a^*)$ denote a minimizer of (4.4) and $u(a_\beta^\delta)$ the corresponding solution to (2.5), (2.6) with a replaced by a_β^δ. If there exists a function

$$\varphi \in W_2^{2,1}(Q_\Delta) \cap \dot{W}_2^1(Q_\Delta)$$

such that

$$(a^\dagger - a^*, p)_{W_2^2(\mathbb{R})} = \int_{Q_\Delta} p\,(u_{yy} - u_y)\,\varphi\, d(y,\tau) \qquad \forall p \in W_2^2(\mathbb{R}),$$

then with $\beta \sim \delta$ the estimates

$$\|a_\beta^\delta - a^\dagger\|_{W_2^2(\mathbb{R})} = \mathcal{O}(\sqrt{\delta}) \quad \textit{and} \quad \|u(a_\beta^\delta) - u^\delta\|_{L_2(Q_\Delta)} = \mathcal{O}(\delta) \tag{4.5}$$

hold.

Proof. The proof uses ideas from [10]. We assume again for simplicity that $b = 0$. Let $u = u(a^\dagger)$, then

$$\|u^\delta - u\|^2_{L_2(Q_\Delta)} + \beta \, \|a^\delta - a^\dagger\|^2_{W_2^2(\mathbb{R})} \le \delta^2 + 2\,\beta\,(a^\dagger - a^*, a^\dagger - a^\delta)_{W_2^2(\mathbb{R})} \tag{4.6}$$

Now, $\beta\,(a^\dagger - a^*, a^\dagger - a^\delta)_{W_2^2(\mathbb{R})}$ can be estimated as follows:

$$\begin{aligned}
\beta\,(a^\dagger - a^*, a^\dagger - a^\delta)_{W_2^2(\mathbb{R})} &= \beta \int_{Q_\Delta} \left[(a^\dagger - a^\delta)\,(u_{yy} - u_y))\right]\,\varphi\, d(y,\tau) \\
&= \beta \int_{Q_\Delta} \left[-(u^\delta - u)_t + a^\delta\,(u^\delta - u)_{yy} - a^\delta\,(u^\delta - u)_y\right]\,\varphi\, d(y,\tau) \\
&= \beta \int_{Q_\Delta} \left((u^\delta - u)\right)\left\{\varphi_t + (a^\delta\varphi)_{yy} + (a^\delta\varphi)_y\right\} d(y,\tau) \\
&\le \epsilon\,\|u - u^\delta\|^2_{L_2(Q_\Delta)} + \frac{\beta^2}{4\,\epsilon}\left(1 + 2\,\|a^\delta\|_{W_2^2(\mathbb{R})}\right)\|\varphi\|_{W_2^{2,1}(Q)}.
\end{aligned}$$

The first term can be shifted to the left side of (4.6) which completes the proof. ■

Lemma 4.1 immediately yields global uniqueness of the inverse problem under the given source condition. Uniqueness for a linearized problem has been proven in [4, 5] with different methods.

With additional smoothness assumptions on $a^\dagger - a^*$, a function φ satisfying the requirements of Lemma 4.1 can be constructed explicitly, which leads to the following

Theorem 4.1 *Let the assumptions of Lemma 4.1 hold. Assume that* $w := (a^\dagger - a^*)$ *and* $g := \partial_y^4 w - \partial_y^2 w + w \in W_2^2(\mathbb{R})$ *satisfy*

$$|g_{yy}|, |g_y|, |g| = \mathcal{O}(e^{-B' y^2})$$

for some B' *large enough and that* $b(t)$ *is Hölder continuous (f.i.* $b \in W_2^1(0,T)$*). Then the rates (4.5) hold.*

Proof. Let

$$\varphi(y) = \int_{T-\Delta}^{T} (T-\tau)\,(\tau - T + \Delta)\,a^\dagger\,(u_{yy} - u_y)\,d\tau$$

and define

$$\Phi(y,\tau) = (T-\tau)\,(\tau - T + \Delta)\,a^\dagger\,\frac{g}{\varphi(y)}.$$

Then $\int_{T-\Delta}^{T} (u_{yy} - u_y)\,\Phi\,d\tau = g$ by construction. Note that $G(y,\tau) := u_{yy} - u_y$ is a fundamental solution of

$$-G_\tau + (a\,G)_{yy} + (a\,G)_y + b\,G_y = 0$$

and satisfies (see [11], Corollary 1, Appendix)

$$G(y,\tau) \ge A\,e^{-B(d)y^2} \qquad \text{for} \quad \tau \ge d > 0$$

for some $B(d) > 0$. Hence $\frac{1}{\varphi(y)} \le C\,e^{By^2}$. It remains to be shown that $\varphi \in W_2^{2,1}(Q_\Delta)$:

$$\begin{aligned}
\varphi(y) &= \int_{T-\Delta}^{T} (T-\tau)\,(\tau - T + \Delta)\,a^\dagger\,(u_{yy} - u_y)\,d\tau \\
&= \int_{T-\Delta}^{T} (T-\tau)\,(\tau - T + \Delta)\,(u_\tau - b\,u_y)\,d\tau \\
&= \int_{T-\Delta}^{T} (2\,T - \Delta - 2\,\tau)u - (T-\tau)\,(\tau - T + \Delta)\,b\,u_y\,d\tau
\end{aligned}$$

If $b(t)$ is Hölder continuous with exponent α, then $u(a;\cdot,\tau) \in C^{2+\alpha}(\mathbb{R})$ for any $\tau > 0$ (see [11]) and it follows (see proof of Proposition A.1) that $u_y \in W_2^{2,1}(Q_\Delta)$. ■

Theorem 4.1 makes assumptions on the derivatives of $(a^\dagger - a^*)$ up to order 6. These smoothness conditions can be relaxed if also derivatives of the data are available. Assume that observation is taken in $W^{1,0}_{2,loc}(Q_\Delta)$ and define the Tikhonov functional

$$J^\delta_\beta(a) = \|u(a) - u^\delta\|^2_{W_2^{1,0}(Q_\Delta)} + \beta\,\|a - a^*\|^2_{W_2^1(\mathbb{R})} \tag{4.7}$$

for parameters $a \in \mathcal{K}_1(a^*) := \{q \in \mathcal{K}(a^*) : q - a^* \in W_2^1(\mathbb{R}),\quad a^* = a^*(y)\}$. By Theorem 2.1 $F : \mathcal{K}_1(a^*) \to W^{1,0}_{2,loc}(Q_\Delta)$ is compact and the stability and convergence results of Proposition 3.1, 3.2 are applicable. If there exists a function $\Phi \in \dot{W}_2^1(Q_\Delta)$ such that

$$(a^\dagger - a^*, p)_{W_2^1(\mathbb{R})} = \int_{Q_\Delta} p\,(u_{yy} - u_y)\,\Phi\,d(y,\tau) \qquad \forall p \in W_2^1(\mathbb{R}),$$

then integration by parts only once in the proof of Lemma 4.1 yields

$$\begin{aligned}\beta\,&(a^\dagger - a^*, a^\dagger - a^\delta)_{W_2^1(\mathbb{R})}\\ &\le \epsilon\,\|u - u^\delta\|^2_{W_2^{1,0}(Q_\Delta)} + \frac{\beta^2}{4\,\epsilon}\left(1 + 2\,\|a^\delta\|_{W_2^1(\mathbb{R})} + \|b\|_{L_\infty(0,T)}\right)\|\Phi\|_{W_2^1(Q)}.\end{aligned}$$

Thus, if we claim that $\tilde{g} := -w_{yy} + w \in W_2^1(\mathbb{R})$ satisfies

$$|\tilde{g}_y|, |\tilde{g}| = \mathcal{O}(e^{-B' y^2})$$

and follow the construction in the proof of Theorem 4.1, we get

$$\|a^\delta_\beta - a^\dagger\|_{W_2^1(\mathbb{R})} = \mathcal{O}(\sqrt{\delta}) \quad \text{and} \quad \|F(a^\delta_\beta) - u^\delta\|_{W_2^{1,0}(Q_\Delta)} = \mathcal{O}(\delta). \tag{4.8}$$

This yields the following result:

Theorem 4.2 *Let $a^\delta_\beta \in \mathcal{K}_1(a^*)$ denote a minimizer of (4.7) and $u(a^\delta_\beta)$ be the corresponding solution of (2.5), (2.6). If $w := (a^\dagger - a^*)$ and $g := -w_{yy} + w \in W_2^2(\mathbb{R})$ satisfy*

$$|g_y|, |g| = \mathcal{O}(e^{-B' y^2})$$

for some B' large enough and if $b(t)$ is Hölder continuous (for instance, $b \in W_2^1(0,T)$), then the rates (4.8) hold.

Observation of derivatives of $u(a)$ can be realized, for instance, by smoothing L^2 observations. Since our choice of parameters ($a \in W_2^1$) implies $(u(a) - u(b))|_{Q_\Delta} \in C^{4+\alpha,2+\alpha/2}(\overline{Q_\Delta})$, we immediately obtain convergence rates, e.g. for Tikhonov regularization, for the pre-smoothing process, i.e. we can give W_2^1 error estimates of the smoothed data in terms of L_2 error bounds for the initial data. A second way is to utilize market prices of digital call options: deriving (2.3), (2.4) once by K it follows immediately that $C_K = e^{-y}u_y$ are the values of digital calls, which are quoted on the markets and observable, too.

The given convergence rate results immediately imply uniqueness of solutions to the inverse problem.

In reality, only discrete data (option prices for a discrete set of maturities and strikes) are available, which suffices to apply the stability and convergence results of Section 3. The assumption of availability of observations over a small time interval used for the derivation of convergence rates is probably the more restrictive one: In [23], the relation of the optimal control problems related to observations over a small time interval $(T-\Delta), T)$ and time-discrete observations at $t = T$ have been investigated: In fact, for $\beta > 0$ the corresponding solutions $a^\delta_\beta(\Delta)$ converge to a solution of the problem with terminal observation only as $\Delta \to 0$. However, we were not able to show that this is true, if β and Δ converge to 0 simultaneously. Note, that a convergence rate result for β, $\Delta \to 0$ would imply uniqueness of the inverse problem with only terminal observation, which is not proven for general inverse parabolic problems (see, e.g., [18]).

5. Numerical implementation and test examples

For an illustration of the theoretical results derived in the last section, we investigate the the numerical realization in the case of one maturity and a time independent volatility. For simplicity we further assume the interest and dividend rates to be $q = r = 0$.

Several numerical techniques are frequently used for pricing options, for instance, bi- or trinomial trees, Monte-Carlo methods, finite difference and finite element methods (see, e.g. [26]). For a validation of our convergence results, we will utilize a finite element discretization of the dual (Dupire-) equation in logarithmic variables, (2.5), (2.6) similar to [19]. The approach via the dual equation allows to price a whole set of options via solution of only one pde.

As approximate solutions for our problem, we look for minimizers of the Tikhonov functional

$$J(a) = \|u(a) - u^\delta\|_u^2 + \beta\,\|a - a^*\|_a^2 \tag{5.1}$$

for a given regularization parameter β and measured (noisy) data u^δ. The norms $\|\cdot\|_u$, $\|\cdot\|_a$ will be chosen in dependence of the available data, in particular, we investigate the choices

$$\|u\|_u^2 = \int |u(y,T_0)|^2\,dy \quad \text{or} \quad \|u\|_u^2 = \sum_i |u(y_i,T_0)|^2$$

and choose for the regularization term

$$\|a\|_a^2 = \int |a|^2 + |a_y|^2 dy,$$

motivated by the convergence considerations of the previous sections.

5.1. Gradient evaluation

For minimizing the Tikhonov functional we employ gradient type (e.g., steepest descent or quasi-Newton) techniques. In any of these methods, a gradient evaluation of (5.1) has to be performed, which can be done efficiently via an adjoint approach. The directional derivative of (5.1) is given by

$$J'(a)[p] = 2(u(a) - u^\delta, u'(a)[p])_u + 2\beta(a - a^*, p)_a.$$

Assuming measurements at time T in L^2, we have with $r := u(a) - u^\delta$, $w := u'(a)[p]$

$$\begin{aligned}
(u'(a)[p], u(a) - u^\delta)_u &= \int w(y,T) r(y) dy \\
&= \int_0^T \frac{d}{dt}\int w(y,t)R(y,t)dydt = \int_0^T \int w_t R + w R_t dydt \\
&= \int_0^T \int [a(w_{yy} - w_y) - p(u_{yy} - u_y)]R + w R_t dydt \\
&= \int_0^T \int [R_t + (aR)_{yy} + (aR)_y]w - p(u_{yy} - u_y)R dydt
\end{aligned}$$

$$= \int_0^T \int p(u_{yy} - u_y) R \, dy dt,$$

where we have chosen R to be a solution to

$$R_t + (aR)_{yy} + (aR)_y = 0$$

with terminal condition

$$R(y, T) = r(y).$$

In case the state domain is approximated by a finite interval, R has to satisfy the additional boundary condition

$$R(y, t) = 0, \quad (y, t) \in \partial\Omega \times (0, T).$$

Alternatively, one can also solve for $S := aR$, which solves

$$S_t + a(S_{yy} + S_y) = 0$$

with

$$S(y, t) = a(y) r(y)$$

and homogeneous Dirichlet boundary conditions. Hence, the same solver as for the forward problem can be utilized. A gradient direction is then given by solution of the variational problem

$$\begin{aligned} (g, p)_a &= 2(u(a) - u^\delta, u'(a)[p])_u + 2\beta(a - a^*, p)_a \\ &= 2 \int_0^T \int p(u_{yy} - u_y) R \, dy dt + 2\beta(a - a^*, p)_a. \end{aligned} \tag{5.2}$$

Here, the bilinear form $(\cdot, \cdot)_a$ depends on the regularization norm, e.g., (5.2) amounts to solving the variational problem

$$-g_{yy} + g = -2(u_{yy} - u_y) R + 2\beta(-da_{xx} + da),$$

with $da = a - a^*$, when regularizing in H^1. Summarizing, one step in a gradient or quasi-Newton algorithm requires two numerical solutions of parabolic pde problems and one solution of the variational problem above (5.2).

5.2. Approximations and incomplete data

For a numerical treatment, the state domain is restricted to a finite interval, e.g., we solve (2.5) on the domain $Q_{M,T} := \Omega_M \times (0, T)$, with $\Omega_M := (-M, M)$, and introduce artificial boundary conditions $u(-M, t) = g_0(t)$, $u(M, t) = g_1(t)$. Appropriate values for $g_i(t)$ can be deduced by studying the decay properties of solutions u. E.g, for the choice $g_0(t) = 1$, $g_1(t) = 0$, the error introduced in this manner is of the order of $e^{-|M|}$, see [23]. Alternatively, approximate Neumann boundary conditions can be derived, i.e. $u_y(-M) \sim u_y(M) \sim e^{-|M|}$.

In the test example below, the boundary values $g_i(t)$ are determined in the following manner: First we use the Black-Scholes formula to calculate Black-Scholes implied volatilities for $u(-M) = z_0$ and $u(M) = z_1$, where z_0 and z_1 are the data for the lowest and highest possible strike. Then these volatilities are used to calculate option values $g_0 = u(-M, T)$ and $g_1 = u(M, T)$ for $T \in [0, T_0]$ again by the Black-Scholes formula.

5.3. Choice of the regularization parameter

In the previous sections we showed that convergence of the approximations a_β^δ obtained by Tikhonov regularization can be expected for $\delta \to 0$, if $\delta^2/\beta \to 0$. Especially, the *a-priori* choice $\beta \sim \delta$ will lead to a converging regularization algorithm. The right choice of the regularization parameter β is crucial for numerical approximations at finite noise levels. We propose the following strategy, motivated by *a-posteriori* parameter choice rules (e.g., Morozov's discrepancy principle, see e.g. [9, 25]): In order to stay within the converging regime, we define a sequence of admissible regularization parameters $\beta_0 = 10^2\delta < \beta_1 < ... < \beta_n = 10^{-3}\delta$. We start minimizing the Tikhonov functional J_{β_0}. Then we subsequently decrease the regularization parameter $\beta = \beta_i$, starting the minimization of the Tikhonov functional J_{β_i} at the previous minimizer $a^\delta_{\beta_{i-1}}$. The procedure is stopped, if the minimal regularization parameter is active, or if a discrepancy principle $\|u(a_\beta^\delta) - u^\delta\| \leq \tau\delta$ for some $\tau > 1$ is met. In our numerical test, with noise levels ranging from 10^{-1} to 10^{-5}, the discrepancy principle for a choice of $\tau = 1.5$ could be met for typical values of $\beta \sim 10^{-2}\delta$.

5.4. Test examples

In order to illustrate the stability and convergence results of Sections 3 and 4, and to discuss some additional matters, e.g., approximation of discrete data, choice of regularization norms, restriction to bounded domains and the influence of data noise, we consider the following numerical test examples:

Example 1 *Constant Volatility:*
Let $S = 100$, $r = q = 0$ *and* $T = 1$ *year to expiry. We set* $M = 3$, *and thus take into account possible strikes* $K \in (5, 2000)$. *The first example consists of identifying the (unknown) constant parameter* $a = 0.15$.

For the first set of examples, we choose the initial guess $a_1^* = 0.15 - 0.05\mathrm{erf}(-y^2)$, which has the correct asymptotic behaviour for $|y|$ large (compare Theorems 4.1 and 4.2). In a first simulation (A), we use a complete set of option prices calculated via the Black-Scholes formula. Data noise is only due to discretization and restriction to a bounded domain, hence can be assumed to be very small. We choose the regularization parameter to be $\beta = 10^{-6}$.

In a second test (B), we investigate the influence of incomplete data given for 20 strikes only. Note, that the problem is now highly under determined, and we expect the resulting minimizer to depend more strongly on the a-priori guess a^*.

The bad identifiability of volatilities far away from the spot and the importance of the decay behaviour of a^* is considered in a third test run (C), where we choose the a-priori guess $a_2^* = 0.1$ and $\beta = 10^{-6}$ as in the previous examples.

In the last test (D), we investigate the influence of data noise by adding 0.1% uniformly distributed noise to the data. In case of an underlying worth 1000$, this amounts to an "uncertainty" of ±1$ in the option prices, which is about the size of typical the bid-ask spreads, e.g., for the S&P500 index option.

Finally, we compare the results of the reconstructions with the option prices for the very good guess $a = 0.16$.

The error-free data in this example are constructed by the Black-Scholes formula: A comparison of the Black-Scholes data with the option values (optimal) corresponding to the true parameter $a = 0.15$ indicates that the restriction to a finite domain is reasonable.

In all examples, the identified volatilities are very close to the true value on the range of strikes given. However, in case C, the volatilities deviate significantly far away from the spot

Table 1. Reconstructed option values after calibration - Example 1: (A) complete exact data, (B) exact incomplete (20 strikes) data, (C) bad initial guess $a^* = 0.1$ and complete exact data, (D) 0.1% additional data noise, (BD) incomplete (20 strikes) data with added 0.1% noise; (optimal) option prices corresponding to true parameter $a = 0.15$ calculated numerically .

strike	true value	optimal	A	B	C	D	BD	noisy data	a=0.16
600.00	439.20	439.15	439.20	439.20	439.20	439.20	438.98	439.49	442.70
700.00	369.61	369.51	369.60	369.60	369.61	369.48	369.36	369.09	374.30
800.00	309.65	309.51	309.65	309.65	309.65	309.42	309.41	308.69	315.34
900.00	258.70	258.54	258.70	258.70	258.70	258.39	258.46	258.23	265.13
1000.00	215.81	215.63	215.80	215.81	215.81	215.44	215.57	215.82	222.70
1100.00	179.93	179.76	179.93	179.93	179.93	179.56	179.70	180.18	187.06
1300.00	125.24	125.09	125.23	125.23	125.24	124.90	125.01	125.85	132.27
1500.00	87.57	87.48	87.57	87.57	87.57	87.32	87.37	87.68	94.05
1800.00	51.88	51.86	51.88	51.88	51.88	51.78	51.82	51.16	57.18

Table 2. Reconstructed volatilities: (A) complete exact data, (B) exact incomplete (20 strikes) data, (C) bad initial guess $a^* = 0.1$ and complete exact data, (D) 0.1% additional data noise, (BD) incomplete (20 strikes) data with added 0.1% noise.

strike	true value	A	B	C	D	BD
600.00	0.1500	0.1500	0.1502	0.1499	0.1504	0.1489
700.00	0.1500	0.1502	0.1502	0.1502	0.1499	0.1495
800.00	0.1500	0.1503	0.1502	0.1503	0.1501	0.1499
900.00	0.1500	0.1503	0.1503	0.1503	0.1495	0.1500
1000.00	0.1500	0.1503	0.1502	0.1503	0.1495	0.1501
1100.00	0.1500	0.1503	0.1503	0.1503	0.1496	0.1501
1300.00	0.1500	0.1502	0.1502	0.1502	0.1498	0.1500
1500.00	0.1500	0.1500	0.1500	0.1500	0.1497	0.1490
1800.00	0.1500	0.1497	0.1497	0.1497	0.1500	0.1500

($K = 1000$) by sticking to the "wrong" *a-priori* guess $a^* = 0.1$, which corresponds to a violation of the source condition in Theorems 4.1 and 4.2.

The large deviation in option prices corresponding to $a = 0.16$ indicates high sensitivity of option values on volatility around the spot, where vice versa good reconstructions can be expected.

Example 2 *Let the S, r, q, T, M be chosen as in the previous example and the true parameter $a = \sqrt{2\sigma}$ be given by*

$$a(y) = 0.15 + 0.05\ \exp[-(y + 0.3)^2] \sin(2\pi y) + 0.05\,\mathrm{erf}(20y).$$

The data are constructed numerically on the domain $\Omega_M = (-5, 5)$ with a fine discretization and a different time integration scheme as used for the reconstruction, to avoid inverse crimes.

Again, we compare the reconstructed volatilities and corresponding option prices for several test runs: (A) The case of complete, exact data and initial guess $a_1^* = 0.15 + 0.05\mathrm{erf}(2y)$ with the correct asymptotic behaviour for $|y|$ large (compare Theorem 4.1 and 4.2). In (B) we investigate the influence of only partial data (for 20 strikes). The influence of a "wrong" *a-priori* guess $a^* = 0.15$ is reported as test (C). In (D) we add additional 0.1% noise. In this example, the true parameter has some oscillating features with decaying amplitude for increasing $|y|$ and is increasing sharply around $y = 0$ ($K = 1000$).

Table 3. Reconstructed option values after calibration - Example 2: (A) complete exact data, (B) exact incomplete (20 strikes) data, (C) bad initial guess $a^* = 0.15$ and complete exact data, (D) 0.1% additional data noise, (BD) incomplete (20 strikes) data with added 0.1% noise; (optimal) option prices corresponding to true parameter $a = 0.15$ calculated numerically .

strike	true value	opt	A	B	C	D	BD	data
600.00	416.95	417.54	416.96	416.94	416.93	416.42	415.73	416.32
700.00	336.69	337.59	336.68	336.76	336.71	336.72	336.41	336.73
800.00	271.59	272.58	271.59	271.77	271.66	272.48	272.81	271.71
900.00	225.42	226.53	225.43	225.71	225.65	226.54	226.91	226.14
1000.00	193.67	194.07	193.66	193.50	193.64	193.40	193.84	194.57
1100.00	168.70	169.18	168.71	168.70	168.52	167.45	167.95	169.19
1300.00	128.42	128.78	128.42	128.40	128.35	127.19	127.72	128.26
1500.00	97.45	97.73	97.45	97.45	97.45	97.01	97.23	97.92
1800.00	64.30	64.48	64.31	64.31	64.34	64.61	64.27	63.37

Table 4. Reconstructed volatilities: (A) complete exact data, (B) exact incomplete (20 strikes) data, (C) bad initial guess $a^* = 0.15$ and complete exact data, (D) 0.1% additional data noise, (BD) incomplete (20 strikes) data with added 0.1% noise.

strike	true value	A	B	C	D	BD
600.00	0.1020	0.0983	0.1000	0.0983	0.0873	0.0872
700.00	0.0696	0.0680	0.0673	0.0679	0.0654	0.0624
800.00	0.0554	0.0540	0.0548	0.0530	0.0561	0.0578
900.00	0.0701	0.0655	0.0713	0.0716	0.0814	0.0838
1000.00	0.1500	0.1609	0.1514	0.1483	0.1360	0.1351
1100.00	0.2273	0.2223	0.2269	0.2168	0.1889	0.1874
1300.00	0.2434	0.2414	0.2433	0.2438	0.2305	0.2399
1500.00	0.2201	0.2185	0.2194	0.2202	0.2318	0.2346
1800.00	0.1869	0.1864	0.1870	0.1880	0.2023	0.1894

As in the previous example option values can reproduced very accurately by the reconstructed volatility. The behaviour of the volatility smile around the spot can be resolved very accurately, while the values for large $|y|$ are highly determined by the *a-priori* guess a^*.

Summarizing, all the numerical test show that volatility smiles can be identified stably by Tikhonov regularization. The decay condition on the unknown part of the volatility $a^\dagger - a^*$ needed for the proof of convergence rates (see Theorems 4.1, 4.2) is reflected in the numerical examples, in which good identifiability is observed only near the spot $y = 0$. Thus, for a reasonable reconstruction of volatilities for very large/small strikes, different techniques have certainly to be used. The numerical tests also confirm that the restriction of the computational domain to a finite interval does not influence the reconstruction of volatilities near the spot noticeably.

Further test examples and details on the influence or the regularization parameter choice on the reconstruction can be found in [8].

6. Concluding Remarks

Tikhonov regularization yields a stable and convergent method for volatility estimation by calibrating a modified Black-Scholes model to market prices of European Vanilla options. The regularized least squares formulation of the inverse problem allows to treat various kinds of observations (continuous, discrete) in a uniform framework.

The question of convergence rates under interpretable assumptions is answered for the case when the term structure of volatility is known, in particular, if volatility is piecewise constant in time (see Theorems 4.1, 4.2). The convergence rate results also imply uniqueness of solutions to the inverse problem. The conditions on differentiability and decay of the unknown part of volatility $(a^* - a^\dagger)$ illustrate two sources of ill-posedness in the problem:

- the differential part corresponds to the smoothing properties of the parameter-to-solution map;
- the decay condition for $|y| \to \infty$ reflects the nonlinearity of the forward problem.

Even though the source condition may seem rather strong, one should bear in mind that market prices of options depend on markets expectation of the future values of volatility: thus it seems reasonable to assume that these expected volatilities are rather smooth. On the other hand, for high/low strikes, volatility cannot be reconstructed stably and therfore additional knowledge on the behaviour of the volatility function for high/low strikes has to be incorporated into the *a-priori* guess a^*.

Utilization of the dual (Dupire) equation (2.3) is essential for fast numerical algorithms, since it allows to valuate a range of options, and hence calculation of the Tikhonov functional, by solving only one differential equation. Minimization of the Tikhonov functional can be done by standard algorithms for nonlinear optimization, e.g., Quasi-Newton-type methods. Gradients, which involve adjoint Frèchet-derivatives of the parameter-to-solution map, can be calculated effectively by using an adjoint approach (5.2), hence again only one additional solution of a parabolic differential equation is necessary for a gradient evaluation of the Tikhonov functional. Using these two techniques, the numerical effort for the reconstruction of a volatility surface can be kept rather small enabling an "online" calibration to market prices.

The non uniqueness of solutions to the inverse problem respectively special assumptions on the term structure (here: piecewise constant) should be kept in mind when pricing path dependent options by means of the calibrated model: If the volatility smile is flat $(\sigma \neq \sigma(S))$, then it is well known that European Call prices expiring at time T depend on the aggregated volatility $\int_0^T \sigma^2(t)dt$ only, i.e., a variety of term structures will perfectly calibrate the model to the market, whereas the corresponding prices of exotic derivatives may vary a lot. Thus, for a reasonable reconstruction of the term structure of volatility, additional data have to be used.

Appendix

Extending standard results on existence and regularity of solutions to parabolic equations (cf. e.g. [11, 21]) by the ideas in [12, 16, 27] one can derive existence and uniqueness results for solutions to (2.5), (2.6) in $W_p^{2,1}$-spaces. A similar existence result for the case $p \in]2, p^*[$ has been derived in a probabilistic framework in [6].

Proposition A.1 *Let $a \in \mathcal{K}(a^*)$, $b \in L_\infty(Q)$ and $f \in L_p(Q) \cap L_2(Q)$. Then there exist $p^* > 2$ such that*

$$-v_\tau + a\, v_{yy} + b\, v_y = f \tag{A.1}$$

$$v(y, 0) = 0 \tag{A.2}$$

has a unique solution $v \in W_p^{2,1}(Q)$ satisfying the estimates

$$\|v\|_{W_p^{2,1}} \leq C_p\, \|f\|_p$$

for $2 \leq p < p^*$, with C_p depending only on the bounds of the coefficients and p.

Proof. According to [21] Th. IV.9.2, (A.1) has a unique solution $v^n \in W_p^{2,1}(Q)$ for smooth (Hölder) a^n with the estimate

$$\|v^n\|_{W_p^{2,1}} \le C(p)\,\|f\|_p \tag{A.3}$$

We establish existence of a solution to (A.1), (A.2) and a uniform bound (A.3) for arbitrary $a \in \mathcal{K}(a^*)$. We start with the case $p = 2$. By linearity $w^n = v_y^n$ solves

$$-w_\tau^n + (a^n\, w_y^n)_y + (b\, w^n)_y = f_y$$

and satisfies the estimate

$$\|w^n\|^2_{W_2^{1,0}(Q)} \le c \int_0^T \|f_y(\tau)\|^2_{W_2^{-1}(\mathbb{R})}\, d\tau \le C\,\|f\|_{L_2(Q)},$$

where the constants c, C now depend only on the lower bound of a and $\|a\|_{L_\infty}$, $\|b\|_{L_\infty}$. To see that $\|v^n\|_{L_2(Q)}$ is also bounded, consider

$$\begin{aligned}
\|v^n\|^2_{L_2(Q)} &= \int_Q \int_0^\tau \frac{d}{d\tau}|v^n|^2\, d\tau\, d(y,t) \\
&= 2\int_Q \int_0^\tau \left\{a\left(v_{yy}^n + b\, v_y^n\right) - f\right\} v d\tau d(y,t) \\
&\le c(T)\,\left\{\|v_{yy}^n\|_{L_2(Q)} + \|v_y^n\|_{L_2(Q)} + \|f\|_{L_2(Q)}\right\}\,\|v^n\|_{L_2(Q)} \\
&\le C(T)\,\|f\|_{L_2(Q)}\,\|v^n\|_{L_2(Q)}
\end{aligned}$$

Thus (A.3) holds uniformly for smooth $a^n \in \mathcal{K}(a^*)$. Now let $a^n \to a \in \mathcal{K}(a^*)$ (in W_2^1), then a (sub-)sequence v^{k_n} of the corresponding solutions v^n admits a weak limit $\hat{v} \in W_2^{2,1}(Q)$. For $\varphi \in W_2^{2,1}(Q)$ with compact support

$$\begin{aligned}
&\int_Q (-v_\tau + a(v_{yy} - v_y) + bv_y)\varphi d(y,\tau) \\
&\quad = \lim_{n\infty} \int_Q (-v_\tau^{k_n} + a^{k_n}(v_{yy}^{k_n} - v_y^{k_n}) + bv_y^{k_n})\varphi d(y,\tau) = (f,\varphi)_{L_2(Q)}
\end{aligned}$$

Hence, by regularity, $\hat{v} = v$ solves (A.1).

Now let $p \neq 2$ and set $a^* = (\overline{a} + \underline{a})/2$. By linearity v solves

$$-v_\tau^k + a^*\,(v_{yy}^k) = f - b\, v_y^n + (a^* - a^n)\, v_{yy}^n =: g^n$$

Following the proof of Th. IV.9.2 in [21], we get

$$\widehat{(v_{yy}^n)}(\xi,s) = \frac{\xi^2}{i\,s + a^*\,\xi^2}\,\widehat{g}(\xi,s)$$

and by Plancherel's theorem

$$\|v_{yy}^n\|_{L_2(Q)} \le \frac{1}{a^*}\|g^n\|_{L_2(Q)} \tag{A.4}$$

Now the Riesz-Thorin theorem [16] implies that the constants $C(p)$ of the estimate

$$\|v_{yy}^n\|_{L_p(Q)} \le C(p)\|g^n\|_{L_p(Q)}$$

depends continuously on p, i.e. there exist $p^* > 2$ such that $\frac{\overline{a}-\underline{a}}{2}C(p) < 1$ for $2 \le p < p^*$ and

$$\|v_{yy}^n\|_{L_p(Q)} \le C(p)\left(\|f\|_{L_p(Q)} + C\,\|v^n\|_{W_p^{1,0}(Q)} + \frac{\overline{a}-\underline{a}}{2}\|v_{yy}^n\|_{L_p(Q)}\right)$$

which by $\|v^n\|_{W_p^{1,0}(Q)} \le c\,\|v^n\|_{W_2^{2,0}(Q)}$ implies

$$\|v^n\|_{W_p^{2,1}(Q)} \le C(p)\ \left[\|f\|_{L_p(Q)} + \|f\|_{L_2(Q)}\right]$$

for $2 \le p < p^*$. Thus the solutions v^n are uniformly bounded in $W_p^{2,1}$. That v solves (A.1) follows as above. ∎

By standard arguments, the existence theorem can be extended to cover our equation (2.5).

Corollary A.1 *Let $a \in \mathcal{K}(a^*)$, $b \in L_\infty(Q)$. Then there exist $p^* > 2$ such that for $2 \le p < p^*$ the problem*

$$-v_\tau + a\,(v_{yy} - v_y) + b\,v_y = 0 \qquad in \quad Q \tag{A.5}$$

$$v(y,0) = (S^* - e^y)^+ \tag{A.6}$$

has a unique solution $v \in W^{2,1}_{p,loc}(Q)$ satisfying the estimates

$$|v| \le S^* \quad and \quad \|v_x\|_{W_p^{1,0}} \le C_p$$

The constant C depends only on the bounds of the parameters and p.

Proof. Observe that

$$\tilde v(y,\tau) := \int_{-\infty}^{y_0} \frac{1}{\sqrt{4\pi\tau}} e^{-\frac{(y-\nu)^2}{4\tau}} (S^* - e^\nu)\, d\nu \quad \in \quad W^{2,1}_{p,loc}(Q)$$

for $p < 3$, $y_0 = \ln(S^*)$. The difference $w = v - \tilde v$ then solves

$$-w_\tau + a\,(w_{yy} - w_y) + b\,w_y = \tilde v_\tau - a\,(\tilde v_{yy} - \tilde v_y) - b\,\tilde v_y$$

with homogeneous initial condition. Proposition A.1 completes the proof. ∎

Continuity of the parameter-to-solution map

$$F : a \mapsto v(a)$$

follows from linearity of equation (2.5) and application of corollary A.1.

Proposition A.2 $F : \mathcal{K}(a^*) \to W^{2,1}_{p,loc}(Q)$ *is continuous for* $2 \le p < p^*$.

Proof. Let $a_n \in \mathcal{K}(a^*)$, u^n denote the corresponding solutions of (2.5),(2.6) and $a_n \to a$ in W_2^1. By Corollary A.1 $u^n \in W^{2,1}_{\tilde p,loc}$ for any $2 \le \tilde p < p^*$. Employing linearity we get that $w^n := u^n - u(a)$ solves

$$-w^n_\tau + a\,(w^n_{yy} - w^n_y) + b\,w^n_y = -(a^n - a)\,(u^n_{yy} - u^n_y)$$

and satisfies the estimate

$$\|w\|_{W_p^{2,1}(Q)} \le C\,\|a^n - a\|_{L_q(Q)} \|u^n_y\|_{W^1_{\overline p}(Q)}$$

for $\frac{1}{\overline p} + \frac{1}{q} = \frac{1}{p}$ and $p < \overline p < p^*$. Sobolev embedding theorems [29] yield $a^n \to a$ in $L_q(Q)$ for $1 \le q < \infty$ and the assertion follows by $\|u^n_y\|_{W^1_{\overline p}(Q)} \le C_{\overline p}$ uniformly for all n. ∎

By compact Sobolev embedding and utilizing the asymptotic behaviour of solutions to (2.5), (2.6) we can even show compactness of this mapping, which is one source of the ill-posedness for the inverse problem of option pricing.

Proposition A.3 *$F : \mathcal{K} \to W^{2,1}_{p,loc}(Q)$ is compact for $2 \le p < p^*$ and weakly (sequentially) continuous.*

Proof. Let a_n, u^n as in the proof of proposition A.2, $a_n \rightharpoonup a$ in $W^1_2(Q)$ and $u^n \rightharpoonup u$ in $W^{2,1}_p(Q) \cap W^{2,1}_{\overline{p}}(Q)$ for $p < \overline{p} < p^*$. By compact Sobolev embedding $a^n \to a$ in $L_q(Q_c)$ for any $1 \le q < \infty$ and $Q_c \subset Q$ compact. On the other hand u satisfies

$$-u_\tau + a\,(u_{yy} - u_y) + b\,u_y = 0 \qquad \text{in} \quad L_p(Q)$$

and as in proposition A.2 the difference $w^n := u^n - u$ satisfies

$$-w^n_\tau + a^n\,(w^n_{yy} - w^n_y) + b\,w^n_y = -(a - a^n)\,(u_{yy} - u_y) =: f^n \tag{A.7}$$

We split $f^n = f^n_1 + f^n_2$ with

$$f^n_1 := f^n \chi_{[-M,M]}, \qquad f^n_2 := f^n - f^n_1$$

and get

$$\|w^n\|_{W^{2,1}_p(Q)} \le C\,\left\{\|a^n - a\|_{L_q(Q_C)}\|u_y\|_{W^1_{\overline{p}}(Q_C)} + \|a^n - a\|_{L_q(Q^c_C)}\|u_y\|_{W^1_{\overline{p}}(Q^c_C)}\right\}$$

for any $2 \le p < \overline{p}$, $\frac{1}{p} + \frac{1}{q} = \frac{1}{\overline{p}}$ and $Q_C := (-M, M) \times (0, T)$, $Q^c_C = Q \backslash Q_C$. Now $\|a^n - a\|_{L_q(Q_C)} \to 0$ by compact Sobolev embedding and $\|u_y\|_{W^1_{\overline{p}}(Q^c_C)} \to 0$ with $M \to \infty$. ■

References

[1] R. A. Adams. *Sobolev Spaces.* Academic Press, New York, 1975.

[2] A. B. Bakushinskii. Remarks on choosing a regularization parameter using the quasi-optimality and ration criterion. *USSR Comp. Math. MAth. Phys*, 24:181–182, 1984.

[3] F. Black and M. Scholes. The pricing of options and corporate liabilities. *J. of Political Economy*, 81:637–659, 1973.

[4] I. Bouchouev and V. Isakov. The inverse problem of option pricing. *Inverse Problems*, 13:L1–7, 1997.

[5] I. Bouchouev and V. Isakov. Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets. *Inverse Problems*, 15:R95–116, 1999.

[6] S. Crépey. Calibration of the local volatility in a generalized Black-Scholes model using Tikhonov regularization. *SIAM J. Math. Anal.*, 34:1183–1206, 2003.

[7] B. Dupire. Pricing with a smile. *RISK*, 7:18–20, 1994.

[8] H. Egger. Volatility estimation in the Black-Scholes model via Tikhonov regularization. Master's thesis, Johannes Kepler Universität Linz, 2002.

[9] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems.* Kluwer, Dordrecht, 1996.

[10] H. W. Engl and J. Zou. A new apporach to convergence rate analysis of Tikhonov regularization for parameter identification in heat conduction. *Inverse Problems*, 16:1907–1923, 2000.

[11] A. Friedman. *Partial Differential Equations of Parabolic Type.* Prentice-Hall, Englewood Cliffs, 1964.

[12] P. Grisvard. *Elliptic Problems in Nonsmooth Domains (Monographs and Studies in Math.) vol 24.* 1985.

[13] M. Hanke and E. Rösler. Computation of local volatilities from regularized Dupire equations. 2004. submitted.

[14] T. Hein and B. Hofmann. On the nature of ill-posedness of an inverse problem in option pricing. *Inverse Problems*, 19:1319–1338, 2003.

[15] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2), 1993.

[16] L. Hoermander. *The analysis of linear partial differential operators.* 1983.

[17] B. Hofmann and O. Scherzer. Local ill-posedness and source conditions of operator equations in Hilbert spaces. *Inverse Problems*, 14:1189–1206, 1998.

[18] V. Isakov. *Inverse Problems for Partial Differential Equations.* 1998.

[19] N. Jackson, E. Süli, and S. Howison. Computation of deterministic volatility surfaces. *J. Mathematical Finance*, 2(2):5–32, 1998.

[20] I. Karatzas and S. E. Shreve. *Methods of Mathematical Finance*. 1998.

[21] O. A. Ladyženskaya, V. A. Solonikov, and N.N. Ural'ceva. *Linear and quasilinear equations of parabolic type (Translations of Mathematical Monographs) vol 24*. AMS, Rhode Island, 1968.

[22] R. Lagnado and S. Osher. A technique for calibrating derivative security pricing models: numerical solution of the inverse problem. *J. Computational Finance*, 1:13–25, 1997.

[23] J. Lishang and T. Youshan. Identifying the volatility of underlying assets from option prices. *Inverse Problems*, 17:137–155, 2001.

[24] R. C. Merton. Option pricing when the underlying stock returns are discontinuous. *J. Financial Economics*, 3, 1976.

[25] V. A. Morozov. On the solution of functional equations by the method of regularization. *Soviet Math. Dokl.*, 7:414–417, 1966.

[26] W. Shaw. *Modelling financial derivatives with Mathematica*. Cambridge University Press, 1999.

[27] M. E. Taylor. *Partial Differential Equations III*. Springer, 1996.

[28] P. Wilmott, S. Howison, and J. Dewynne. *The Mathematics of Financial Derivatives*. Cambridge University Press, 1995.

[29] W. P. Ziemer. *Weakly Differentiable Functions (Graduate Texts in Math. vol 120)*. Springer, New York, 1989.