Identification of the local speed function in a Levy model for option pricing

S. Kindermann [a,c], P. Mayer [b,*], H. Albrecher [b,c], H. W. Engl [a,c]

[a] *Johannes Kepler University Linz, Altenbergerstrasse 69, A-4040 Linz, Austria*
[b] *Graz University of Technology, Steyrergasse 30, A-8010 Graz, Austria*

[c] *Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergerstrasse 69, A-4040 Linz, Austria*

Abstract

We propose a non-parametric stable calibration method based on Tikhonov regularization for the local speed function in a local Lévy model. The jump term in this model introduces an integral operator into the classic Black-Scholes partial differential equation such that the associated model calibration to observed option prices can be treated as a parameter identification problem for a partial integro-differential equation. This problem is shown to be ill-posed and thus requires regularization. It is proven that nonlinear Tikhonov regularization is a stable and convergent method for this problem. Furthermore, convergence rate results are established under an abstract source condition. Finally the theoretical results are underpinned by numerical illustrations including a real-world example.

1 Introduction

During the last decades, it became more and more obvious that the famous Black & Scholes model cannot adequately describe the stochastic behaviour of financial markets. Hence, much work has been devoted to find more appropriate models, which are able to reproduce the stylized facts of the observed asset price processes, as skewed log-returns and volatility clusters.
One idea of adapting the Black & Scholes model was pioneered by Dupire [13] and Derman & Kani [12], who modeled the volatility as a deterministic function of asset price and time (in contrast to assuming a stochastic process for the volatility leading to so-called stochastic volatility models, see e.g. [19] for diffusion-based or [5] for Lévy-type models). The motivation behind this idea was given by Gyöngy [21], who showed that the marginals of *any* Itô process can be recovered by such a model.
It is essential for the applicability of any model that it can be calibrated to traded market prices. The major task in the local volatility framework is to identify the local volatility function, which is well known to be an ill-posed problem and hence needs some regularization. The robustness and basic convergence rates for the Tikhonov-regularized inverse problem were proven e.g. by

*Supported by the Austrian Science Fund Project P-18392.

Crepey [11] and detailed convergence rates were obtained by Egger & Engl [14]. Although the Dupire model has some desireable properties (such as improving upon modelling the smile phenomenon, simple option pricing schemes and model completeness), there are also severe drawbacks. For instance, the normal distribution is symmetric and hence skewness in the local volatility model can only be introduced by means of a non-constant volatility function. This often results in a local volatility surface that tends to be steep for small t and to flatten out quite rapidly for larger values of time t. The resulting term structure of the local volatility function can lead to serious problems in pricing exotic options with the Dupire model (for details see e.g. Andersen & Andreasen [2] and references therein). Whereas the local volatility model is nevertheless widely used in practice as an interpolation scheme, it does not give an interpretation of the resulting market dynamics.
On the other hand, exponential Lévy models (with constant volatility) can explain stylized facts such as skewness and excess kurtosis through the presence of jumps in the asset price process, but usually fail to give a reasonable fit to liquid vanilla market quotes across all strikes and maturities (see e.g. [7, 36]).
To overcome these problems, Carr et al. [6] introduced the so-called *local Lévy model* as a generalization of both the exponential Lévy and the Dupire model. In this model, in addition to the local volatility concept of the Dupire model, the asset price process includes jumps driven by a Lévy process where the jump intensity is a parameter that evolves deterministically over time and is called *local speed function*. Just as in the Dupire model, where the log price process can be interpreted as a Brownian motion running at the speed of the square of the local volatility function, the log price process in the local Lévy model (in the absence of the diffusion component) can be interpreted as a (fixed) Lévy process running at a space-time dependent speed. The local Lévy model has, in addition to being capable of fitting the whole European option price surface, also a persistent skew, if an asymmetric jump size distribution is used for the jump part. In this way, it incorporates the advantages of the Dupire model and at the same time provides richer risk-neutral dynamics. In order to use the local Lévy model one needs to identify the Lévy measure of the jump part (together with the parameter of the continuous part) and the local speed function.
The first problem - a non-parametric estimation of an exponential Lévy market - has attracted much attention lately. For instance, Cont & Tankov [8, 9] proposed a robust non-parametric method to recover the Lévy measure from observed option prices, if the Lévy measure is assumed to be bounded. Another approach was proposed by Belomnesty & Reis [3, 4], who showed that the Lévy measure can be identified using the semi-closed formula for option prices involving the characteristic function of the asset price and an asymptotic analysis. In financial practice it is also quite common to fix a certain type of Lévy measure beforehand and then calibrate the involved parameters from market data. Cont & Tankov [9] observed that the stably calibrated exponential Lévy market model, while leading to good results for a single maturity, cannot adequately fit the observed option prices across several maturities. Furthermore they recognized that the Lévy measure changes structurally over time. By introducing the local speed function this behavior can be incorporated into the model. Hence by identifying this function, a given exponential Lévy process can be adapted to a more appropriate model for the asset price.
In this paper we are concerned with this question, i.e. given an exponential Lévy

model, how can this model be adapted by means of the local speed function to fit the observed European option prices across all maturities. Extending the analysis of Carr et al. [6], we show a way to robustly identify the local speed function by using the observed prices of European options and prove convergence results for this inverse problem. Also, within regularization theory the usual problem of bid-ask spreads of market option prices can be interpreted as a problem of data noise, which allows to incorporate the degree of liquidity into the calibration procedure.

The structure of the paper is as follows: Section 2 introduces the local Lévy model in more detail and defines the inverse problem to be solved. To deal with the ill-posedness of the problem (see Section 5), a regularization is needed to obtain robust calibration results. Section 3 proves that the corresponding forward operator is well-defined and Section 4 establishes some of its properties. These are used in Section 5 to prove stability of the Tikhonov-regularized inverse problem and to obtain convergence rates of the regularized solutions. Finally, in Section 6 we give numerical illustrations of the theoretical results.

2 The local Lévy model and the inverse problem

We will assume the asset price S to have the following risk-free dynamics:

$$\begin{aligned} S_t &= S_0 + \int_0^t (r-\eta) S_{s_-}\, ds + \int_0^t \sigma_0(S_{s_-}, t) S_{s_-}\, dW_s \\ &\quad + \int_0^t \int_{\mathbb{R}} (e^x - 1) \left(m_{(S_{s_-}, s)}(dx, ds) - \nu_{(S_{s_-}, s)}(dx, ds) \right), \qquad (1) \end{aligned}$$

where r is the riskless interest rate, η is the dividend yield, m is the jump-count-measure and ν is its compensator. The above dynamics is equivalent to the following representation:

$$S_t = e^{(r-\eta)t} e^{X_t}, \qquad (2)$$

where e^{X_t} is a martingale.

Following Carr et al. [6] we may introduce the local speed function $a_0(S_t, t)$ governing the speed at which the jumps arise, i.e., we set:

$$\nu_{(S_{t_-}, t)}(dx, dt) = a_0(S_{t_-}, t)\nu(dx, dt). \qquad (3)$$

Note that the above setting implies that the arrival rate of the jumps depends on time and state, while the jump size distribution itself remains unchanged over time.

We want to calibrate the above model to the observed option prices. As the call option payoff is a convex function, the Tanaka-Meyer formula (see [32]) can be used to derive a partial integro-differential equation for the European call price $C(K, T)$ with strike K and maturity T. This was done in [6], and the partial integro-differential equation (PIDE) reads as follows:

$$\begin{aligned} C_T &= -\eta C - (r-\eta) K C_K + \frac{\sigma_0^2(K,T)}{2} K^2 C_{KK} \\ &\quad + \int_0^\infty Y C_{KK}(Y,T) a_0(Y,T) \psi_e \left(\log\left(\frac{K}{Y}\right)\right) dY \quad \text{on } \mathbb{R}^+ \times [0, T^*] \qquad (4) \end{aligned}$$

where r is the riskless interest rate, η is the dividend yield, σ_0 is the volatility, a_0 is local speed function defined in (3), T^* is the finite planning horizon (e.g. the largest option maturity available in the market) and ψ_e is the double-exponential tail of the Lévy measure given by

$$\psi_e(z) = \begin{cases} \int_{-\infty}^{z} (e^z - e^x)\,\nu(dx) & \text{for } z < 0 \\ \int_{z}^{\infty} (e^x - e^z)\,\nu(dx) & \text{for } z > 0. \end{cases}$$

Note that the function ψ_e is similar to the price of out-of-the money puts and calls, respectively. In fact, for $z > 0$, it is the expected value of $(e^x - e^z)^+$ under the Lévy-measure ν, which governs the jumps of the asset price process. As we shall see later on, $\psi_e(z)$ plays a crucial role in the calibration of the local speed function as well as in the degree of ill-posedness of the inverse problem.

Remark 1. *For notational convenience the following analysis is based on the assumption that r and η are constants. The analysis can, however, easily be extended to the case where r and η are deterministic functions of time.*

In order to solve the partial integro-differential equation (4) uniquely, it has to be supplemented with initial and boundary values: The initial value of the price of a European call option is given by

$$C(K, 0) = (S - K)^+ = \max(S - K, 0). \tag{5}$$

Moreover, we impose the following boundary conditions

$$C(0, T) = e^{-\eta T} S, \quad C(\infty, T) = 0. \tag{6}$$

A change of variables $k = \ln K$, $y = \ln Y$ and

$$c(k, T) = e^{\eta T} C(e^k, T) \quad \Leftrightarrow \quad C(K, T) = e^{-\eta T} c(\log(K), T)$$

leads to the equation

$$\begin{aligned} c_T(k,T) + \left(r - \eta + \frac{\sigma(k,T)^2}{2}\right) c_k(k,T) - \frac{\sigma^2(k,T)}{2} c_{kk}(k,T) \\ = \int_{-\infty}^{\infty} (c_{kk} - c_k)(y,T) a(y,T) \psi_e(k-y)\, dy \quad \text{on } \mathbb{R} \times [0, T^*], \end{aligned} \tag{7}$$

with

$$a(y,T) := a_0(e^y, T) \quad \text{and} \quad \sigma(k,T) := \sigma_0(e^k, T).$$

The initial condition and the boundary conditions now read

$$c(k, 0) = (S - e^k)^+, \tag{8}$$

$$c(\infty, T) = 0, \tag{9}$$

$$c(-\infty, T) = S. \tag{10}$$

Our aim is to identify the function a from the observed values of $c(k,T)$ (the liquid European call option prices in the market) given fixed values for the other model parameters, in particular a local volatility function σ. We formulate the problem as an abstract operator equation, which allows us to use a standard regularization approach for solving it.

We define a forward operator F mapping the local speed function a to the option price $c(k,T)$ for those time values t and log-strikes k, where data are given. Since we do not know the price for all times t and all strikes k, it is convenient to split F into an observation operator O and a parameter-to-solution operator $\tilde{F}$.
The mapping $\tilde{F}$

$$\tilde{F} : a \to c(k,T) \quad k \in \mathbb{R}, T \in [0, T^*]$$

assigns to a given local speed function a the corresponding transformed option price $c(k,T)$, which is a solution to (7) - (10) for all k,T in $\mathbb{R} \times [0,T^*]$.
The linear observation operator O is simply the restriction of c to the set of points $\Omega_{dat} \subset \mathbb{R} \times [0,T^*]$ for which data are given:

$$O : c \to c(k,T) \quad (k,T) \in \Omega_{dat}.$$

In total, we have a forward operator

$$F(a) := O\tilde{F}(a) \tag{11}$$

and the calibration problem can be written as the operator equation

$$F(a) = y, \tag{12}$$

where y are the given data of option prices at Ω_{dat}.
We cannot expect that equation (12) can be solved stably for a, in fact it is well-known that the similar identification problem of the volatility in the Black-Scholes equation is ill-posed in reasonable function spaces [14].
In Section 5 we show that the local Lévy problem above is ill-posed as well. It is well-known that in this situation a standard algorithm for solving (12) (e.g. just least squares minimization) might fail (because of instability) and one should instead apply regularization methods to this equation. The main idea of these methods is that, instead of the ill-posed equation (12), a related (well-posed) problem is solved. By solving this regularized problem the instability of the original problem is removed, at the cost of introducing an additional approximation error. In regularization theory one usually deals with a parameterized family of regularized problems. Within this family of problems the free regularization parameter has to be chosen in such a way to find a compromise between accuracy of the approximation of the original problem and stability. This choice of the regularization parameter (*parameter choice rule*) has to be done depending on the level of noise in the data y.
A very common regularization method, mainly for linear, but also for nonlinear problems is Tikhonov regularization, by which an approximate solution is found by minimizing the so-called Tikhonov functional

$$J(a) := \|F(a) - y_\delta\|^2 + \alpha \|a - a^*\|_s^2. \tag{13}$$

Here F is the forward operator defined in (11), y_δ is the data (the observable option prices), possibly contaminated with noise, a^* is an initial guess, $\alpha > 0$ the regularization parameter and $\|.\|_s$ denotes an appropriate norm in a Sobolev space. Convergence (and convergence rates) of Tikhonov regularization for the linear case can be found in [20, 30] and for the nonlinear case in [17, 28, 31]. For further information concerning Tikhonov regularization we refer to the monograph [16].

Using the general and well-known theory of Tikhonov regularization, in Section 5 we will show that under certain conditions nonlinear Tikhonov regularization yields a convergent regularization method also for the calibration problem (12). The main building blocks for the applicability of the general theory are proving continuity and differentiability properties of F, which we will be concerned with in Section 4.
However, we first have to show that F is a well-defined operator, i.e. we have to show that the PIDE (7)-(10) has a unique solution c. This is the topic of the next section.

3 Existence and Uniqueness of a solution to the PIDE

To show the existence of a solution to the forward problem (7) together with (8), (9) and (10), we will use some techniques (especially subtracting the payoff from the call price to get homogeneous boundary conditions) developed in Matache et al. [29], who showed solvability of the backward problem in the exponential Lévy market case (see also [10]). Here we extend this solvability result to the local Lévy market case.
We denote by $\mathcal{I}$ the integral operator, defined as the convolution with ψ_e:

$$\mathcal{I} : v \to \int_{-\infty}^{\infty} v(y,T)\psi_e(k-y)dy$$

and by $\mathcal{L}_a$ the integro-differential operator

$$\begin{aligned}\mathcal{L}_a : v \to &\left(r-\eta+\frac{\sigma^2(k,T)}{2}\right)v_k(k,T)-\frac{\sigma^2(k,T)}{2}v_{kk}(k,T)\\ &-\left(\mathcal{I}\left[a(v_{kk}-v_k)\right]\right)(k,T),\end{aligned}$$

where the functions c, a, σ may depend on k, T.
Then equation (7) can be written as

$$c_T + \mathcal{L}_a c = 0.$$

Since the boundary conditions are not homogeneous following [29] we subtract the payoff function from c and consider

$$\hat{c}(k,T) := c(k,T) - (S-e^k)^+. \tag{14}$$

The derivatives of $(S-e^k)^+$ in the sense of distributions are as follows:

$$g_0(k) := (S-e^k)^+$$

$$g_1(k) := \frac{d}{dk}(S-e^k)^+ = \begin{cases} -e^k & k \le \log(S) \\ 0 & \text{else} \end{cases} \tag{15}$$

$$g_2(k) := \frac{d^2}{dk^2}(S-e^k)^+ = \begin{cases} -e^k & k \le \log(S) \\ 0 & \text{else} \end{cases} + S\delta_{\log(S)}. \tag{16}$$

Now, by definition (14), $\hat{c}$ has to satisfy (g_0 does not depend on time)

$$\hat{c}_T + \mathcal{L}_a\hat{c} = -\mathcal{L}_a g_0 \tag{17}$$

with homogeneous boundary conditions.
We consider the following Hilbert spaces with the obvious inner products:

$$H := L^2(\mathbb{R}) = \left\{ f \,:\, \int_{\mathbb{R}} |f(x)|^2 dx < \infty \right\}$$

$$V := H^1(\mathbb{R}) = \left\{ f \in L^2(\mathbb{R}) \,:\, \int_{\mathbb{R}} |f(x)|^2 dx + \int_{\mathbb{R}} |f'(x)|^2 dx < \infty . \right\}$$

It is well known, that V is the closure of $C_0^\infty(\mathbb{R})$ in the norm H^1, i.e. any element in V can be approximated by $C_0^\infty(\mathbb{R})$ functions.
We will show existence of (17) in the space $L^2([0,T^*],V)$. Since the space V can be embedded continuously into H, $V \to H \to V'$ (V' being the dual space of V) forms a Gelfand triple. The crucial point for an existence proof of (17) is the Gårding inequality (22) and the fact that $\mathcal{L}_a g_0 \in L^2([0,T^*],V')$. To show these properties we need the following assumptions:

Assumptions

$$\sigma(k,T) \in L^\infty([0,T^*], W^{1,\infty}(\mathbb{R})) \tag{18}$$

$$\sigma(k,T) \geq c_0 > 0 \quad \text{for } (k,T) \in \mathbb{R} \times [0,T^*] \tag{19}$$

$$\mathbb{E}\left[S_t \ln S_t\right] < \infty, \quad 0 \leq t \leq T^*, \text{ for } S_t \text{ defined in (1)} \tag{20}$$

$$a(k,T) \in L^\infty(\mathbb{R} \times [0,T^*]), a_k(k,T) \in L^\infty([0,T^*], L^2(\mathbb{R})). \tag{21}$$

First we prove the Gårding inequality for the integro-differential operator $\mathcal{L}_a$ and continuity of the corresponding bilinear form:

Proposition 1. *Let assumptions* (18)-(21) *hold, then there exist some constants* $b > 0$, γ *and* B *such that*

$$(\mathcal{L}_a u, u)_{V',V} \geq b\|u\|_V^2 - \gamma \|u\|_{L^2}^2 \tag{22}$$

and

$$(\mathcal{L}_a u, v)_{V',V} \leq B\|u\|_V \|v\|_V . \tag{23}$$

Proof. Let $u, v \in C_0^\infty$. Multiplication of $\mathcal{L}_a u$ by v and integration gives a bilinear form A_T

$$\begin{aligned} A_T(u,v) &:= (\mathcal{L}_a u, v) \\ &= \int_{\mathbb{R}} \left(\left(r - \eta + \frac{\sigma^2(k,T)}{2} \right) u_k(k) - \frac{\sigma^2(k,T)}{2} u_{kk}(k) \right) v(k) dk \\ &\quad - \int_{\mathbb{R}} \left(\mathcal{I}\left[a(u_{kk} - u_k)\right] \right)(k,T)) \, v(k) dk. \end{aligned}$$

Integrating by parts we get

$$A_T(u,v) = I_1(u,v) + I_2(u,v) + I_3(u,v),$$

where

$$\begin{aligned} I_1(u,v) &= \int_{\mathbb{R}} \left(r - \eta + \frac{\sigma^2(k,T)}{2} + \sigma(k,T)\sigma_k(k,T) \right) u_k(k) v(k)\, dk \\ I_2(u,v) &= \int u_k(k) \frac{\sigma^2(k,T)}{2} v_k(k)\, dk \\ I_3(u,v) &= - \int (\mathcal{I}\left[a(u_{kk} - u_k)\right])(k,T) v(k)\, dk \end{aligned}$$

If we assume (18), we obtain with some constants B_0, B_1

$$
\begin{aligned}
|I_1(u,v)| &\le (B_0 + \|\sigma(.,T)\|^2_{L^\infty} + \|\sigma(.,T)\|_{L^\infty}\|\sigma_k(.,T)\|_{L^\infty})\|u_k\|_{L^2}\|v\|_{L^2} \\
&\le B_1\|u_k\|_{L^2}\|v\|_{L^2}.
\end{aligned}
$$

From (19) we obtain

$$I_2(u,u) \ge c_0\|u_k\|^2_{L^2}$$

and again with (18)

$$I_2(u,v) \le B_2\|u_k\|_{L^2}\|v_k\|_{L^2}.$$

Now let us look at the term I_3: substituting $k = k + y$, integrating by parts and interchanging the order of integration (applying Fubini's theorem) yields

$$
\begin{aligned}
|I_3(u,v)| &= \left| \int_{\mathbb{R}} \int_{\mathbb{R}} \psi_e(k-y) a(y,T)(u_{kk}(y) - u_k(y))\; dy\; v(k)\; dk \right| \\
&\le |i_1| + |i_2| + |i_3|
\end{aligned}
$$

with

$$
\begin{aligned}
|i_1| &= \left| \int_{\mathbb{R}} \psi_e(k) \int_{\mathbb{R}} v(k+y)(a(y,T))u_k(y)\; dy\; dk \right|, \\
|i_2| &= \left| \int_{\mathbb{R}} \psi_e(k) \int_{\mathbb{R}} v(k+y)(a_k(y,T))u_k(y)\; dy\; dk \right|, \\
|i_3| &= \left| \int_{\mathbb{R}} \psi_e(k) \int_{\mathbb{R}} v_k(k+y) a(y,T) u_k(y)\; dy\; dk \right|.
\end{aligned}
$$

Before discussing these integrals in detail we show that $|\psi_e|$ is integrable:

$$
\begin{aligned}
\int_{-\infty}^{\infty} |\psi_e(x)| dx &= \int_0^{\infty} \int_z^{\infty} (e^x - e^z)\; \nu(dx)\, dz \\
&\quad + \int_{-\infty}^{0} \int_{-\infty}^{z} (e^z - e^x)\; \nu(dx)\, dz.
\end{aligned}
$$

For the sake of brevity we will just consider the first integral, as for the second one the arguments are similar.
With the help of Tonelli's theorem we can interchange the order of integration to find

$$
\begin{aligned}
\int_0^{\infty} \int_z^{\infty} (e^x - e^z)\; \nu(dx)\, dz &= \int_0^{\infty} \int_x^{\infty} (e^x - e^z)\; dz\, \nu(dx) \\
&= \int_0^1 (xe^x - e^x + 1)\; \nu(dx) \quad (*) \\
&\quad + \int_1^{\infty} (xe^x - e^x + 1)\; \nu(dx) \quad (**).
\end{aligned}
$$

The integral $(*)$ exists, since

$$\lim_{x \to 0} \frac{xe^x - e^x + 1}{x^2} = \frac{1}{2}$$

and $\int_0^1 x^2 \nu(dx) < \infty$ for every Lévy measure. The integral $(**)$ is finite, if (20) is met (cf. Sato [35, Cor.25.8]).

Turning to the estimates for the integrals we find with the help of (21) and the Hölder inequality for the first integral:

$$|i_1| \leq \|a(.,T)\|_{L^\infty}\|v\|_{L^2}\|u_k\|_{L^2}\int_{\mathbb{R}}|\psi_e(k)|\ dk.$$

For i_2 using the Sobolev embedding theorem results in:

$$\begin{aligned} |i_2| &\leq \|v\|_{L^\infty}\|a_k(.,T)\|_{L^2}\|u_k\|_{L^2}\int_{\mathbb{R}}|\psi_e(k)|\ dk \\ &\leq B_4\|v\|_{H^s}\|a_k(.,T)\|_{L^2)}\|u_k\|_{L^2}\int_{\mathbb{R}}|\psi_e(k)|\ dk, \end{aligned}$$

with $1/2 < s < 1$.
Hence we are left to deal with the H^s-norm of v. To that end we consider the interpolation inequality

$$\|v\|_{H^s} \leq B_5\|v\|_{H^1}^s\|v\|_{L^2}^{1-s}.$$

Then by Young's inequality we get that for any $\epsilon > 0$ a $B_6(\epsilon)$ exists with

$$|i_2| \leq \left(\frac{\epsilon}{4}\|v\|_{H^1} + B_6(\epsilon)\|v\|_{L^2}\right)\|a_k(.,T)\|_{L^2}\|u_k\|_{L^2}\int_{\mathbb{R}}|\psi_e(k)|\ dk.$$

Let us now turn to the integral i_3. For arbitrary δ, applying partial integration to the integral over $\{|k| > \delta\}$, it can be bounded by

$$\begin{aligned} |i_3| &\leq \left|\left(\psi_e(\delta)+\psi_e(-\delta)\right)\int_{\mathbb{R}} v(k+\delta)a(y,T)u_k(y)\ dy\ dk\right| \\ &\quad + \left|\int_{\{|k|>\delta\}}\frac{\partial}{\partial k}\psi_e(k)\int_{\mathbb{R}} v(k+y)a(y,T)u_k(y)\ dy\ dk\right| \\ &\quad + \left|\int_{\{|k|\leq\delta\}}\psi_e(k)\int_{\mathbb{R}} v_k(k+y)a(y,T)u_k(y)\ dy\ dk\right| \\ &\leq B_6\|a(.,T)\|_{L^\infty}\|v\|_{L^2}\|u_k\|_{L^2}\left(\psi_e(\delta)+\psi_e(-\delta)+\int_{\{|k|>\delta\}}\left|\frac{\partial}{\partial k}\psi_e(k)\right|\ dk\right) \\ &\quad + B_7\|a(.,T)\|_{L^\infty}\|v_k\|_{L^2}\|u_k\|_{L^2}\int_{-\delta}^{\delta}|\psi_e(k)|\ dk \\ &\leq 2B_6\|a(.,T)\|_{L^\infty}\|v\|_{L^2}\|u_k\|_{L^2}(\psi_e(\delta)+\psi_e(-\delta)) \\ &\quad + B_7(\delta)\|a(.,T)\|_{L^\infty}\|v_k\|_{L^2}\|u_k\|_{L^2}\int_{-\delta}^{\delta}|\psi_e(k)|\ dk, \end{aligned}$$

where the fact was used, that ψ_e is differentiable everywhere except in zero, $\frac{\partial}{\partial k}\psi_e(k)$ is non-positive for $k > 0$ and non-negative for $k < 0$ and furthermore $\lim_{k\to\infty}\psi_e(k) = 0$. Hence for any $\delta > 0$ the integral $\int_{\{|k|>\delta\}}\left|\frac{\partial}{\partial k}\psi_e(k)\right|\ dk$ is bounded.
Since $\int_{-\delta}^{\delta}|\psi_e(k)|\ dk \to 0$ for $\delta \to 0$ there is a $\delta > 0$, such that $\int_{-\delta}^{\delta}|\psi_e(k)|\ dk < \epsilon/4$ and hence:

$$|I_3(u,v)| \leq B_8(\epsilon)\|u_k\|_{L^2}\|v\|_{L^2} + \frac{\epsilon}{2}\|u_k\|_{L^2}\|v_k\|_{L^2}.$$

For later reference we state the main estimate of $|I_3|$:

$$|I_3(u,v)| \leq R\Big(\|a(.,T)\|_{L^\infty} + \|a_k(.,T)\|_{L^2}\Big)\|\psi_e\|_{L^1}\|u_k\|_{L^2}\|v\|_V \tag{24}$$

for some constant $R > 0$.
Thus $A_T(u,u)$ fulfills

$$A_T(u,u) \geq (c_0 - \epsilon/2)\|u_k\|^2_{L^2} - B_9\|u_k\|_{L^2}\|u\|_H - B_{10}\|u\|_{L^2}\|u\|_{L^2}.$$

with a constant B_9 depending on the norms of σ and a. By Young's inequality, we have

$$\|u_k\|_{L^2}\|u\|_H \leq \frac{\epsilon}{2}\|u_k\|^2_{L^2} + \frac{1}{2\epsilon}\|u\|^2_H$$

and

$$A_T(u,u) \geq (c_0 - \epsilon)\|u\|^2_V - \gamma(\epsilon)\|u\|^2_H. \tag{25}$$

Since ϵ can be arbitrarily small, A_T satisfies the inequality (22). Moreover, by the same estimates we obtain

$$|A_T(u,v)| \leq B\|u\|_V\|v\|_V,$$

which proves continuity of the bilinear form. □

Now we can prove the main existence theorem:

Theorem 1. *Let a and σ be such that* (18)-(21) *hold. Then there exists a unique weak solution $\hat{c} \in L^2([0,T^*],V)$ of* (17)*, i.e. a solution satisfying*

$$(\hat{c}_T, \phi)_{L^2} + (\mathcal{L}_a\hat{c}, \phi)_{L^2} = (-\mathcal{L}_a g_0, \phi)_{L^2} \quad \forall \phi \in V, \tag{26}$$

$$\hat{c}(0) = 0 \tag{27}$$

and there are constants B, B' such that

$$\|\hat{c}\|_{L^2([0,T],V)} \leq B\|\mathcal{L}_a g_0\|_{L^2([0,T],V')} \leq B'. \tag{28}$$

Proof. The existence and uniqueness of a weak solution results from standard parabolic theory (e.g. [27, 18]) using Proposition 1, if the right hand side of (17) is in $L^2([0,T],V')$. But this follows directly with the estimates for I_1, I_2, I_3 in the proof of Proposition 1: Indeed, since $g_1 \in L^2(\mathbb{R})$ we get

$$|(\mathcal{L}_a g_0, v)| \leq B_0\|g_1\|_{L^2}(\|v\|_{L^2} + \|v_k\|_{L^2}),$$

for any $v \in H^1$ and hence $\mathcal{L}_a g_0 \in L^2([0,T],V')$. □

Note that using the same techniques as in Carr et al [6] it can be shown that $e^{(\eta-r)T}\mathbb{E}[(S_T-K)^+] - (S_0-K)^+$ fulfills equation (26) and hence corresponds in fact to the solution of the weak formulation of the PIDE.

Remark 2. *Let us discuss the assumptions on the problem in more detail. The smoothness and positivity condition* (18), (19) *on the volatility σ are common for such problems, similar assumptions are also used for the local volatility problem in the Black-Scholes setting. They are certainly satisfied, if σ is taken to be a positive constant. Condition* (21) *is a smoothness condition on the local speed function and it defines the space in which we search for a.*

Condition (20) *is a sufficient condition such that the double exponential tail is integrable. It translates the integrability condition into a condition of the stochastic process modelling the underlying* S.
By definition, $\mathbb{E}[S_t] = e^{rt}S_0 < \infty$ *for all arbitrage-free market models so that* (20) *is not very restrictive (for instance, it holds if* $\mathbb{E}[S_t^{1+\delta}] < \infty$ *for some* $\delta > 0$*, which is in particular the case whenever the variance of the asset price is finite).*

Remark 3. *The main difficulty in the proof of Theorem 1 was the estimate concerning the convolution terms. It is based on the special form of the double exponential kernel such that its derivative is in* $L^1_{loc}(\mathbb{R}\setminus\{0\})$ *and the fact that* ψ_e *is integrable. Let us mention that the proof holds under alternative conditions on* ψ_e*, which do not take into account the specific form of the kernel. For instance, if* ψ_e *is such that* $\mathcal{I}$ *acts as a smoothing operator in Sobolev spaces* $H^{-s} \to L^2$*, for any* $s > 0$ *the term* $|i_2|$ *could be estimated from above by* $\|a\|_{L^\infty}\|v\|_{H^{1-s}}\|u_k\|_{L^2}$ *and with an interpolation inequality and Young's inequality we still arrive at* (25)*. Alternatively, if* ψ_e *is such that* $\mathcal{I}$ *is a compact operator from* $L^2 \to L^2$ *one could use Ehrling's lemma [34] with the same conclusion.*

4 Properties of the parameter to solution map

For the application of regularization theory [16] we have to prove some basic continuity and – for convergence rates results [16, 17] – differentiability properties of the forward operator F defined in (11). This is the topic of this section. At first we focus on continuity of the parameter-to-solution map $\tilde{F}$ in a reasonable space, i.e., the well-posedness of the forward problem. Note that $\tilde{F}$ was defined as a mapping of the local speed function a to c (which is a solution to (7)). In Section 3 it was convenient to subtract g_0 from c and state the integro-differential equation for $\hat{c}$ defined in (14). For investigating the continuity properties of $\tilde{F}$ we will again subtract g_0 from c and will analyze the operator

$$G : a \to \hat{c} = \tilde{F}(a) - g_0. \tag{29}$$

Of course, since g_0 is a known function, which does not depend on a, continuity of G will also imply continuity of $\tilde{F}$, if g_0 lies in the same space as $G(a)$. The only difficulty we face here is that $g_0(k)$ is not an $L^2(\mathbb{R})$ function, since $\lim_{k\to-\infty} g_0(k) = S$ is a nonzero constant. Hence $\tilde{F}(a)$ is not in $L^2(\mathbb{R})$. One possibility to prove continuity of $\tilde{F}$ would be to use a weighted L^2-norm for the image space of $\tilde{F}$ (compare [1]). Another possibility, which we will follow here, is to subtract g_0 from $\tilde{F}$ and also from the data to state an equivalent identification problem with G as parameter-to-solution operator. This is convenient, because in this case we can work with the usual L^2 space.

4.1 Continuity

We now show the continuity of the modified parameter-to-solution operator (29)

$$G : a \to \hat{c},$$

with $\hat{c}$ solution to (17).
We will use the following space for the local speed function a:

$$X := \{a(k,T) \in L^\infty(\mathbb{R} \times [0,T^*]) \,|\, a_k(k,T) \in L^\infty([0,T^*], L^2(\mathbb{R}))\} \tag{30}$$

with norm

$$\|a\|_X := \|a\|_{L^\infty(\mathbb{R}\times[0,T^*])} + \sup_{t\in[0,T^*]} \|a_k(\cdot,T)\|_{L^2(\mathbb{R})}.$$

For the norm on the image space of G we use $L^2([0,T],V)$.

Proposition 2. *G is Lipschitz continuous from $X \to L^2([0,T],V)$, i.e. there exists a constant B such that for all a_1, $a_2 \in X$*

$$\|G(a_1) - G(a_2)\|_{L^2([0,T],V)} \leq B\|a_1 - a_2\|_X.$$

Proof. According to the definition we have

$$G(a_1) - G(a_2) = \hat{c}_1 - \hat{c}_2,$$

where $\hat{c}_1, \hat{c}_2$ solve (17) with parameters a_1 and a_2, respectively. Now consider the difference

$$v := \hat{c}_1 - \hat{c}_2.$$

Note that $v \in H^1$ satisfies

$$\begin{aligned} v_T + \mathcal{L}_{a_1} v &= \mathcal{L}_{a_2} g_0 - \mathcal{L}_{a_1} g_0 + \mathcal{L}_{a_2}\hat{c}_2 - \mathcal{L}_{a_1}\hat{c}_2 \\ &= \big(\mathcal{I}a_1 - \mathcal{I}a_2\big)(g_2 - g_1) \\ &\quad + \big(\mathcal{I}a_1 - \mathcal{I}a_2\big)\left(\frac{d^2}{dk^2}\hat{c}_2 - \frac{d}{dk}\hat{c}_2\right), \end{aligned} \tag{31}$$

with homogeneous boundary conditions and g_1, g_2 as in (15) and (16), respectively. Since v solves the same equation as $\hat{c}$ with a different right-hand side, we can apply Theorem 1 to obtain an estimate with some constant B

$$\begin{aligned} &\|v\|_{L^2([0,T],V)} \leq \\ &\qquad B\left\|\big(\mathcal{I}a_1 - \mathcal{I}a_2\big)(g_2 - g_1) + \big(\mathcal{I}a_1 - \mathcal{I}a_2\big)\left(\frac{d^2}{dk^2}\hat{c}_2 - \frac{d}{dk}\hat{c}_2\right)\right\|_{L^2([0,T],V')} \end{aligned} \tag{32}$$

and hence, we have to estimate the right hand side of (31) in the norm of $L^2([0,T],V')$.
Let $u \in V$, with the use of (24) we then obtain

$$\begin{aligned} &\big|\big(\mathcal{I}a_1(g_2 - g_1) - \mathcal{I}a_2(g_2 - g_1), u\big)\big| \leq \\ &\left(\|a_1(.,T) - a_2(.,T)\|_{L^\infty(\mathbb{R})} + \|\big(a_1(.,T) - a_2(.,T)\big)_k\|_{L^2(\mathbb{R})}\right)\|g_1\|_{L^2(\mathbb{R})}\|u\|_V \leq \\ &\qquad B_1\left(\|a_1(.,T) - a_2(.,T)\|_{L^\infty(\mathbb{R})} + \|\big(a_1(.,T) - a_2(.,T)\big)_k\|_{L^2(\mathbb{R})}\right)\|u\|_V \end{aligned}$$

where the last inequality follows from $\|g_1\|_{L^2(\mathbb{R})} < \infty$. With $\hat{c}_2$ in place of g_0 we find with the same arguments:

$$\begin{aligned} &\left|\left(\big(\mathcal{I}a_1 - \mathcal{I}a_2\big)\left(\frac{d^2}{dk^2}\hat{c}_2 - \frac{d}{dk}\hat{c}_2\right), u\right)\right| \leq \\ &B_2\left(\|a_1(.,T) - a_2(.,T)\|_{L^\infty(\mathbb{R})} + \|\big(a_1(.,T) - a_2(.,T)\big)_k\|_{L^2(\mathbb{R})}\right)\|u\|_{H^1}\|\frac{d}{dk}\hat{c}_2\|_{L^2}. \end{aligned}$$

Now we let u depend on time as well and integrate over time to get

$$\left| \int_0^T \big((\mathcal{L}_{a_2} g_0 - \mathcal{L}_{a_1} g_0), u \big) \; dt \right| \quad \leq \quad B_1 \| (a_1 - a_2) \|_X \| u \|_{L^2([0,T],V)}$$

and

$$\left| \int_0^T \big((\mathcal{L}_{a_2} \hat{c}_2 - \mathcal{L}_{a_1} \hat{c}_2), u \big) \; dt \right| \leq \| (a_1 - a_2) \|_X \| u \|_{L^2([0,T],V)} \| \hat{c}_2 \|_{L^2([0,T],V)}.$$

Because $\| \hat{c}_2 \|_{L^2([0,T],V)}$ is bounded by (28) we get

$$\| \left(\mathcal{L}_{a_2} g_0 - \mathcal{L}_{a_1} g_0 \right) + \left(\mathcal{L}_{a_2} \hat{c}_2 - \mathcal{L}_{a_1} \hat{c}_2 \right) \|_{L^2([0,T],V')} \leq B_3 \| (a_1 - a_2) \|_X.$$

Together with (32) this proves the Lipschitz continuity of G. □

4.2 Frechet-differentiability

Now by similar means we can compute the Frechet-derivative of G.

Proposition 3. *G is Frechet-differentiable as a mapping from $X \to L^2([0,T],V)$ and the Frechet derivative $G'(a)$ is given by*

$$G'(a) : h \to v,$$

where v solves the equation

$$v_T + \mathcal{L}_a v = \mathcal{I} \left(h \left(\frac{d^2}{dk^2} c_a - \frac{d}{dk} c_a \right) \right), \tag{33}$$

with homogenous boundary and initial conditions and

$$c_a = \hat{c}_a + g_0 = G(a) + g_0 = \tilde{F}(a).$$

Moreover, the Frechet derivative $G'(a)$ is Lipschitz continuous.

(Note that the subscript a in c_a, $\hat{c}_a$ does not denote the derivative with respect to a, but just indicates their dependence on a.)

Proof. We first have to show that (33) is well-defined, i.e. it has a unique solution for any $h \in X$. Since v solves an equation similar to $\hat{c}$, but with different right-hand side, we can apply Theorem 1 with the right-hand side as in (33). To show the existence and uniqueness of a weak solution we have to bound this term in the $L^2([0,T],V')$-norm. However, this follows similarly as in the proof of Lipschitz-continuity (Proposition 2) using (24), $c_a = g_0 + \hat{c}_a$ and (28). Following the proof in detail it can even be shown that

$$\left\| -\mathcal{I} \left(h \left(\frac{d^2}{dk^2} c_a - \frac{d}{dk} c_a \right) \right) \right\|_{L^2([0,T],V')} \leq B_0 \| h \|_X. \tag{34}$$

The next step is to show that the formal derivative $G'(a)$ defined in (33) is really the Frechet-derivative, i.e.

$$\| G(a+h) - G(a) - G'(a) h \|_{L^2([0,T],V)} = o(\| h \|_X). \tag{35}$$

For any a, $a+h \in X$ let $\hat{c}_a = G(a)$, $\hat{c}_{a+h} = G(a+h)$ be the corresponding solutions to (17) and $v = G'(a)h$ defined in (33). The difference $u := \hat{c}_{a+h} - \hat{c}_a - v$ satisfies the equation

$$u_T + \mathcal{L}_a u = \mathcal{I}\left(h\left(\frac{d^2}{dk^2} - \frac{d}{dk}\right)(\hat{c}_{a+h} - \hat{c}_a)\right).$$

Again we can use the main estimate (28) to find that

$$\|u\|_{L^2([0,T],V)} \leq B_1 \|\mathcal{I}h(\hat{c}_{a+h} - \hat{c}_a)\|_{L^2([0,T],V')} \leq \\ B_2\|h\|_X \|(\hat{c}_{a+h} - \hat{c}_a)\|_{L^2([0,T],V)} \leq B_3\|h\|_X^2,$$

which implies (35). The first inequality in the last line can be found by following the proof of Proposition 2, while the second one stems from the Lipschitz continuity of G.
Finally the Lipschitz continuity of the Frechet derivative can be derived by observing that $w = v_a - v_{\tilde{a}}$ (v_a is the Frechet derivative of the parameter to solution map G at point a) solves the following PIDE:

$$w_T + \mathcal{L}_a w = (\mathcal{L}_{\tilde{a}} - \mathcal{L}_a)\, v_{\tilde{a}} + \mathcal{I}\left(h\left(\frac{d^2}{dk^2} - \frac{d}{dk}\right)(\hat{c}_a - \hat{c}_{\tilde{a}})\right).$$

As in the proof of Proposition 2 and using the continuity of $G'(a)$ (i.e. the bound $\|v_{\tilde{a}}\|_{L^2([0,T],V)} \leq C\|h\|_X$) the first term on the right hand side can be estimated by

$$\begin{aligned} \|(\mathcal{L}_{\tilde{a}} - \mathcal{L}_a)\, v_{\tilde{a}}\|_{L^2([0,T^*],V')} &\leq B_4\|a - \tilde{a}\|_X \|v_{\tilde{a}}\|_{L^2([0,T],V)} \\ &\leq B_5 \|h\|_X \|a - \tilde{a}\|_X. \end{aligned}$$

Similarly as in (34) together with the Lipschitz continuity of $\tilde{F}$ in Proposition 2 ($\hat{c}_a$ is Lipschitz in a) we obtain

$$\begin{aligned} &\mathcal{I}\left(h\left(\frac{d^2}{dk^2} - \frac{d}{dk}\right)(\hat{c}_a - \hat{c}_{\tilde{a}})\right)\|_{L^2([0,T^*],V')} \\ &\leq\ B_6\|h\|_X \|\hat{c}_a - \hat{c}_{\tilde{a}}\|_{L^2([0,T^*],V)} \leq B_6\|h\|_X\|a - \tilde{a}\|_X \end{aligned}$$

Now applying the main estimate as in (28) we find

$$\|w\|_{L^2([0,T],V)} \leq B\|a - \tilde{a}\|_X \|h\|_X.$$

By definition of the operator norm this establishes Lipschitz continuity of the Frechet-derivative. □

4.3 The Adjoint Operator

For convergence rate results as well as for the numerical implementation we also have to calculate the adjoint of $G'(a)$. The usual convergence theory for nonlinear Tikhonov regularization [17] is formulated in Hilbert spaces, whereas up to now we used the Banach space X (30) for the parameter space of the local speed function. To apply the abovementioned regularization theory as well as to find the adjoint operator $G'(a)$ we have to use Hilbert spaces. The previous results (Proposition 2, Proposition 3 in Section 4) still hold if X is replaced by a Hilbert space H^s with norm $\|.\|_s$ which is embedded into X, for later purposes we will also need that H^s is compactly embedded, i.e. a bounded sequence in H^s has a strongly convergent subsequence in X:

$$\|a\|_X \leq C_1\|a\|_s \; \forall a \in H^s \quad \wedge \quad \|a_n\|_s \leq C_2 \Rightarrow \exists_{a_{n_k}} a_{n_k}. \text{ convergent in } X \tag{36}$$

Our notation indicates that we will use Sobolev spaces H^s of some order s since for those embedding theorems are available and their computation is well known (e.g. using finite elements or finite differences).
So we now treat G as an operator from $H^s \to Y$ with $Y = L^2([0,T] \times L^2(\mathbb{R}))$ and H^s embedded into X as in (36). $G'(a)$ is continuous from H^s to Y, hence an adjoint operator $G'(a)^*$ exists, mapping from $Y \to H^s$. This linear operator is defined by the identity

$$(G'(a)h, \phi)_{L^2[0,T]\times L^2(\mathbb{R})} = (h, G'(a)^*\phi)_{H^s}$$

for all $h \in H^s$ and $\phi \in Y$.
The computation of the adjoint is split into several steps: First we have to consider the adjoint equation of (33). We start with arbitrary ϕ in $L^2[0,T] \times L^2(\mathbb{R})$ and consider a (weak) solution of

$$-u_T + \mathcal{L}_a^* u = \phi, \quad u(T^*,\cdot) = 0 \tag{37}$$

where $\mathcal{L}_a^*$ is the formal adjoint differential operator to $\mathcal{L}_a$

$$\begin{aligned} \mathcal{L}_a^* u &= -\frac{\partial}{\partial k}\left(\left(r - \eta + \frac{\sigma(k,T)^2}{2}\right)u(k,t)\right) - \frac{\partial^2}{\partial k^2}\left(\frac{\sigma(k,T)^2}{2}u(k,T)\right) \\ &\quad - \left(\frac{\partial^2}{\partial k^2} + \frac{\partial}{\partial k}\right)(a\mathcal{I}^* u). \end{aligned}$$

$\mathcal{I}^*$ is the adjoint integral operator of $\mathcal{I}$ in the L^2-inner product. It is well-known in the theory of integral equations (see e.g. [15, 26]), that $\mathcal{I}^*$ has the same form as $\mathcal{I}$ (a convolution operator) but with kernel $\psi(y-k)$ instead of $\psi(k-y)$.
$\mathcal{L}_a^*$ is the formal adjoint to $\mathcal{L}_a$, i.e. it follows by integration by parts that

$$(\mathcal{L}_a v, u)_{L^2} = (v, \mathcal{L}_a^* u)_{L^2} \quad \forall u, v \in V.$$

This immediately implies that the Gårding inequality (22) and continuity (23) also hold for $\mathcal{L}_a^*$. Thus, using a change of the time variable $s \to T^* - T$, (37) can be rewritten as a standard parabolic initial value problem, for which the existence and uniqueness of a solution in $L^2([0,T^*], V)$ holds.
With integration by parts with respect to the time-variable we get for any $h \in H^2$ and u a solution to (37)

$$\begin{aligned} (G'(a)h, \phi)_{L^2[0,T]\times L^2(\mathbb{R})} &= \int_0^T \int_{\mathbb{R}} \mathcal{I}\left(h\left(\frac{d^2}{dk^2}c_a - \frac{d}{dk}c_a\right)\right) u \, dkdt \\ &= \int_0^T \int_{\mathbb{R}} h(k,t)\left(\frac{d^2}{dk^2}c_a - \frac{d}{dk}c_a\right)\mathcal{I}^* u =: (h, \tilde{G}'(a)^*\phi)_{H^s, H^{s\prime}}. \end{aligned}$$

$\tilde{G}'(a)^*$ is now the adjoint using the dual pairing of H^s and $H^{s\prime}$. For the adjoint in H^s we have to compose it with the Riesz isomorphism $\mathcal{R}_s : H^{s\prime} \to H^s$, which is a smoothing operator and for many simple cases involves solving a partial differential equation (for instance if $s = 1$ the isomorphism involves solving the Poisson equation). Altogether we get the adjoint $G'(a)^*$ by the mapping

$$\phi \to \mathcal{R}_s \circ \left[\left(\frac{d^2}{dk^2} c_a - \frac{d}{dk} c_a \right) \mathcal{I}^* u \right], \tag{38}$$

where u solves (37). This was the derivation of the adjoint for the case that the observation operator O is the identity, i.e. when prices for all log strikes and times are avaliable. If this is not the case, then the adjoint is given in a similar way as in (38), but where u solves

$$-u_T + \mathcal{L}_a^* u = O^* \phi, \quad u(T^*, \cdot) = 0$$

instead of (37). The L^2-adjoint O^* of the observation operator is the extension operator of ϕ by 0 to the whole space $\mathbb{R} \times [0, T^*]$.
The computation of the adjoint is important for at least two reasons. Firstly, an efficient numerical method for computing a minimizer of the Tikhonov functional requires the adjoint operator of the Frechet-derivative. The essence of the so-called adjoint method (compare [14]) is that for computing a descent direction of the Tikhonov functional we do not have to compute the full Frechet-derivative or its adjoint, but only the application of this operator to a given element.
Secondly, the adjoint is also important in the theory of Tikhonov regularization, because a sufficient condition for obtaining convergence rates - the so-called source condition - is formulated via the adjoint operator. The range of the adjoint determines in an essential way the degree of ill-posedness of the problem. We will come back to this issue in the next section in the discussion of the convergence rates result in Theorem 4.

5 Regularization

We now turn to the regularization of the inverse problem of identifying the local speed function. For the problem described, there are at least two objectives to be achieved: At first we want to calibrate the problem to data, i.e. the discrepancy between the model and the data should be small. On the other hand, we want to find a model which is robust, even if the data are noisy. For this reason we phrased the problem as a parameter identification problem, where we not only try to make the discrepancy small, but try to find one local speed function - in a stable way - which reproduces the data well. This shifts the focus of convergence in the data space to convergence in the parameter space. If we consider the identification problem as solving equation (12) (or equivalently (29)) in the space H^s such that (36) holds, it follows immediately that the forward operator is compact and hence by a well-known result [16] the problem is locally ill-posed in this space. As it is typical for ill-posed problems, a good fit to the data does not necessarily imply that the computed parameters are close to the "real" parameters. For that reason we have to use regularization. With this tool we can compute a parameter in a stable way even in the presence of errors in the data. In addition, we then have a convergence theory which basically tells us that for small data noise the computed parameters will still be close to the real parameters. Moreover, it is even possible to give bounds on

the error between computed and real parameters, which is the essence of the convergence rates result later. We make use of the well-known regularization theory in Hilbert spaces, which can be found in detail in [16].
To state the result, we need some definitions and assumptions. First we assume that the data are attainable, i.e. there is a real speed function a which creates the exact data y via (12). In practice, exact data are usually not available, instead what can be observed are noisy data y_δ. The noise level δ is the distance between the exact and the noisy data:

$$\|y - y_\delta\|_Y \leq \delta, \tag{39}$$

where y is in the range of the forward operator F. Using Tikhonov regularization we compute a regularized solution $a_{\alpha,\delta}$ from the noisy data y_δ by minimizing the Tikhonov functional in (13):

$$a_{\alpha,\delta} = \operatorname{argmin}_{a \in H^s} \left(\|F(a) - y_\delta\|_Y^2 + \alpha \|a - a^*\|_s^2 \right). \tag{40}$$

Here $\alpha > 0$ is the so-called regularization parameter, and a^* is an initial guess. As norm $\|.\|_Y$ we use the L^2-norm on the observation set Ω_{dat}. By applying the theory of nonlinear Tikhonov regularization [17, 16] we obtain the result that if the noise level δ goes to 0 and the regularization parameter is chosen appropriately (see Theorem 3) then the computed solution converges as well. Its limit is a so-called a^*-minimum norm solution, which reproduces the exact data (i.e. (12) holds) and has minimal distance to a^* under all other parameter choices for which (12) holds. In mathematical terms, an a^*-minimum norm solution (usually denoted by $a^\dagger$) is defined as

$$a^\dagger := \operatorname{argmin}_a \left\{ \|a - a^*\|_s \,\middle|\, F(a) = y \right\}. \tag{41}$$

The notion of minimum norm solution is only relevant if the problem (12) does not have a unique solution. If (12) is uniquely solvable, the minimum norm solution is 'the' solution to the problem. For the definition of $a^\dagger$ to make sense we have to assume that the set of solutions to (12) with $\|a - a^*\|_s < \infty$ is not empty. The initial guess a^* has to be chosen such that this assumption is satisfied. If, for instance, the exact parameter is a perturbation of a known constant $a = a_0 + h$ with $a_0 \in \mathbb{R}$ and $h \in H^s$, then a^* is conveniently chosen as this constant a_0.
Before stating the main results we have to include a continuity assumption of the observation operator: We impose that O is continuous from $L^2([0,T^*],V) \to L^2(\Omega_{dat})$, where Ω_{dat} are the set of points in k,T-space where option prices are available, and

$$\|Of\|_{L^2(\Omega_{dat})} \leq B_{11} \|f\|_{L^2([0,T^*],V)}. \tag{42}$$

This holds, for instance if Ω_{dat}, has positive Lebesgue measure in (k,T)-space.
At first, we have to prove that a minimizer of the Tikhonov functional exists. This can be done with help of the analysis in Section 4. Since there we showed several properties of G it is convenient to reformulate the minimization problem (40) in an equivalent way with G. This operator was obtained from the forward operator by subtracting the known function g_0 in (29); in a similar manner we can subtract it from the data to get a problem involving G. More precisely, with the noisy data $\hat{y}_\delta := y_\delta - Og_0$ and exact data $\hat{y} = y - Og$ and the notation

$$\tilde{G} := OG, \tag{43}$$

(40) is equivalent to

$$\text{argmin}_{a\in H^s}\|\tilde{G}(a)-\hat{y}_\delta\|_Y^2+\alpha\|a-a^*\|_s^2. \tag{44}$$

Moreover, (41) and (39) are equivalent to the corresponding definition involving $\hat{y}, \hat{y}_\delta, OG(a)$ instead of $y, y_\delta, F(a)$. We can now show that the Tikhonov functional has a minimizer:

Theorem 2. *Under the assumptions* (18)–(20) *on* σ, ψ, S, *and if* (36), (42) *hold, and a minimum norm solution* $a^\dagger$ *with* (41) *exists, then for a fixed noise-level* δ, *(cf.* (39)*), the functional* (44) *has a minimizer* $a_{\alpha,\delta}$.

Proof. By [17] and using formulation (44) we only have to show that the graph of $\tilde{G}$ is weakly sequentially closed. However, since H^s is weakly sequentially closed and compactly embedded into X by (36), a weakly convergent sequence in H^s converges strongly in X. By Proposition 2 G is continuous from X to $L^2([0,T^*],V)$, hence, with (42) $\tilde{G}$ is continuous from X to $L^2(\Omega_{dat})$. In total $\tilde{G}$ is continuous and compact from H^s to $L^2(\Omega_{dat})$. From this it follows that the graph fo $\tilde{G}$ is weakly sequentially closed. □

We can now state the main result on stability of the regularization and convergence.

Theorem 3. *Let the assumptions* (18),(19), (20) *on* σ, ψ *be satisfied,* (42) *hold and* δ *be as in* (39). *If the norm* $\|.\|_s$ *in* (40) *is chosen such that* $s > \frac{3}{2}$, *then the Tikhonov functional admits a global minimum* $a_{\alpha,\delta}$.
Moreover for fixed α *this regularized solution depends stably on the data* y_δ *in the sense that if* y_δ *is replaced by a sequence* y_k *converging to* y_δ *as* $k\to\infty$, *then the corresponding minimum* $a_{\alpha,k}$ *in* (40), *has a convergent subsequence with limit* $a_{\alpha,\delta}$.
The regularized solution $a_{\alpha,\delta}$ *converges in the following sense: For a sequence of data* y_{δ_k} *with noise levels* $\delta_k\to 0$ *and if* α_k *is chosen such that* $\alpha_k\to 0$ *and* $\delta_k^2/\alpha_k\to 0$, *then the sequence of regularized solutions* a_{α,δ_k} *corresponding to these data has a convergent subsequence, and the limit of any convergent subsequence is an* a^**-minimum norm solution of* (7). *If the solution to* (7) *is unique then* a_{α,δ_k} *itself converges.*

Proof. With the results of Section 4, (39) and (42), the result follows from the well-known theorems in [17, 16]. □

Theorem 3 summarizes the main results for Tikhonov regularization: Existence of a minimizer, stability of the regularized solution and convergence if the noise-level goes to 0 and the regularization parameter α is chosen as stated in the theorem. It is well-known in regularization theory that for ill-posed problems in order to obtain convergence, α cannot be chosen freely, but has to depend on the noise level. This choice of α is called a parameter choice rule.
The previous theorem establishes convergence of the regularized solutions to a minimum norm solution as $\alpha\to 0$ and $\delta\to 0$. The next question is, how fast this convergence takes place. It is known that for ill-posed problems, without additional conditions, the convergence speed can be arbitrarily slow. On the other hand, if the minimum norm solution and the initial guess satisfy some

abstract smoothness condition ("source condition"), convergence rates can be derived.
From the analysis in the previous section, together with the theory of nonlinear Tikhonov regularization [17, 16], we can conclude that the basic assumptions for an application of Tikhonov regularization are satisfied and that we get convergence and also convergence rates under the following source condition:

Theorem 4. *If in addition to the assumptions of Theorem 3 the conditions*

$$\exists_{w \in Y} : a^\dagger - a^* = \left(\tilde{G}'(a^\dagger)\right)^* w \tag{45}$$

$$\gamma \|w\| < 1 \tag{46}$$

hold, where γ is the Lipschitz constant of G', and α is chosen as $\alpha \sim \delta$ (as $\delta \to 0$), then the regularized solution converges to the minimum norm solution with rate

$$\|a_{\alpha,\delta} - a^\dagger\| \leq \mathcal{O}(\sqrt{\delta}).$$

Note that we can use the equivalent formulation of Tikhonov regularization in (45) instead of (40), which is why we used G in (45). For (40) the analogous condition would involve $F'(a^\dagger)^*$ instead of $(\tilde{G}'(a^\dagger)^*$.
We now give a detailed interpretation of the smallness condition (46) and the source condition (45). The first one can be interpreted in a way that our initial guess a^* has to be sufficiently close to the exact solution. The latter, however, is a stronger condition in the sense that we also have to know that the difference between initial guess and exact solution have to be in the range of $\tilde{G}'(a^\dagger)^*$. Since this is a smoother space than L^2, this implies that the non-smooth part of $a^\dagger$ has to be known, in order to get convergence rates.
The necessity of a source condition for obtaining convergence rates is a consequence of the ill-posedness of the problem. Note that equation (45) is a non-trivial condition on $a^\dagger - a^*$ because the equation cannot simply be solved for w by inverting $\tilde{G}'(a^\dagger)^*$; thus such a w may not exist. The source condition is also very useful in interpreting the ill-posedness of the problem. The difficulty of the problem can be classified in degrees of ill-posedness according to the smoothing property of $\tilde{G}'(a^\dagger)^*$, and hence how hard it is to fullfill condition (45). If the range of $\tilde{G}'(a^\dagger)^*$ consists of n-times differentiable functions, then we have to know the jumps in the nth derivative of the unknown parameter $a^\dagger$ in order to fulfill (45). Now the higher the degree of smoothness of the range $R(\tilde{G}'(a^\dagger)^*)$, the harder it is to satisfy the source condition. This gives a measure of the ill-posedness of the problem: If $R(\tilde{G}'(a^\dagger)^*)$ is the Sobolev space H^n, then the problem is as ill-posed as n-times differentiation.
We can take the source condition (45) and analyze the smoothing properties of $(\tilde{G}'(a^\dagger))^*$ in more detail, to identify those factors that make the problem ill-posed and to compare the local-Lévy identification problem with others such as the corresponding identification problem for the Dupire model.
From the structure of the adjoint we can observe several parameters influencing the ill-posedness. $\tilde{G}(a)^*$ maps ϕ to u, the solution of the adjoint equation with right-hand side $O^*\phi$. Then the integral operator $\mathcal{I}^*$ is applied to u and finally $\mathcal{I}^*(u)$ is multiplied with $\left(\frac{d^2}{dk^2}c_a - \frac{d}{dk}c_a\right)$.
If we assume for simplicity, that we have data on the whole space, i.e. $O = Id$, we can analyze the smoothing properties of $(\tilde{G}(a))^*$ as follows:

The step from $\phi \to u$ involves solving the adjoint equation, which is only mildly smoothing. In fact there is almost a 1-1 correspondence of u and ϕ: For any ϕ we have a solution u, and if u_t, u_x and u_{xx} are in L^2, then a ϕ exists that satisfies the adjoint equation. Thus solving the adjoint equation is smoothing in the sense that we gain one time-derivative and two space-derivatives. This ill-posedness does not depend very much on the choice of ψ and a.
The main difficulty in satisfying the source condition comes from the integral operator $\mathcal{I}$ and the multiplication with $\left(\frac{d^2}{dk^2}c_a - \frac{d}{dk}c_a\right)$. If ψ_e is a k-times differentiable kernel function, then the corresponding integral operator maps into a space of k-times differentiable functions. Hence in order to satisfy a source condition, $(a^* - a^\dagger)\left(\frac{d^2}{dk^2}c_a - \frac{d}{dk}c_a\right)^{-1}$ has to be k-times differentiable as well. In that case, we have to know the jumps in the $(k-1)$-th derivative of the solution to get convergence rates.
As a result we can state that the smoother the kernel function ψ, the more ill-posed the problem is. Note that for the corresponding identification problem of the volatility in the Dupire model, the adjoint has a similar structure [14], where $\mathcal{I}$ is the identity operator, which is not smoothing at all. In this case the range of the operator in (45) is mainly determined by the smoothness of the solution u to the adjoint equation (37). For the local Lévy model, additionally a smoothing integral operator acts on u, which makes (45) a more difficult condition than for the Dupire (or Black-Scholes) model. This shows that the local Lévy model is (depending on ψ_e) much more ill-posed than the corresponding Black-Scholes calibration model.
However, there is a second implication of the source condition, which comes from the multiplication operator with multiplier $\left(\frac{d^2}{dk^2}c_a - \frac{d}{dk}c_a\right)$. This quantity is proportional to the formal density of the asset price at time t. To see this, observe that $\left(\frac{d^2}{dk^2}c_a - \frac{d}{dk}c_a\right) = K^2 C^{(a)}_{KK}$, where $C^{(a)}$ denotes the European call prices with the local speed function a. Note that $C_{KK}(K,T) = f_{S_T}(K)$ with f_{S_T} denoting the density of S_T.
If this density is strictly positive for all t and K, then we can divide by the density and the source condition is essentially a smoothness condition determined by ψ_e. However, in degenerate cases (if $\sigma = 0$ for instance) it may happen that the multiplier is 0 for some points (k,T). In this case the source condition implies that $a^\dagger = a^*$ in these regions. Consequently we have to know the solution in this region (which is evident, since in regions with density 0, the value of a has no influence on the call price c_a, and hence cannot be reconstructed from any data there). The assumption $\sigma > 0$ used in this paper avoids this problem. Nevertheless, the problem of zero density should kept in mind for the study of more general problems (in particular, pure jump processes are often considered to be appealing models).
Furthermore the source condition gives a more precise quantitative interpretation: Wherever the density is very small, the initial guess has to be close to the exact solution in order to guarantee that $(a^* - a^\dagger)\left(\frac{d^2}{dk^2}c_a - \frac{d}{dk}c_a\right)^{-1}$ is in the range of $\mathcal{I}^*$.
So we have identified at least two major influence factors for the ill-posedness: a small value of the density function and the smoothness of ψ_e.
Finally we note that if O is not the identity operator, i.e. the data are given

only on a subset of $\mathbb{R} \times [0, T^*]$ the problem is also more difficult, since the right-hand side of the adjoint equation is only supported on the set of observation values. As a consequence, we have less freedom of choice for w in (45). This formally implies the natural result that less data make the calibration problem more difficult.

6 Numerical Illustration

We now turn to the numerical computations and results for the Tikhonov regularization of the identification problem. To get a regularized solution we have to compute the Tikhonov functional (13) and apply a minimization algorithm. The main computational work in computing the forward operator $\tilde{F}$ in the functional concerns the numerical solution of the integro-differential equation (4). For this we have to discretize the domain $\mathbb{R}^+ \times [0, T^*]$ and the governing PIDE (4) and the involved Hilbert space norms.
For the computation we replaced the unbounded domain $\mathbb{R}^+ \times [0, T^*]$ by a bounded one, $[0, K_0] \times [0, T^*]$, where K_0 was chosen to be $K_0 = 5S$, S being the spot price.
The computational domain is then discretized uniformly into n_K and n_T intervals for the K- and T-direction, respectively.
The derivatives with respect to K in (4) are replaced by finite differences, and the integral is discretized by a midpoint rule. On the interval $[0, K_0]$ we used the same Dirichlet boundary condition as in (6) and the initial condition (5). The parameters in the equation σ_0 and a_0 are discretized on the same uniform grid as the calls.
For the resulting discretized evolution equation of (4) we used a Crank-Nicholson type scheme of the form

$$(\tfrac{1}{\Delta T} I - \tfrac{1}{2} A) C_{n+1} = (\tfrac{1}{\Delta T} I + \tfrac{1}{2} A) C_n + B C_n,$$

where C_n is the vector of discretized call prices at time step n, ΔT is the time step size, the matrix A represents the discretized versions of the terms in equation (4) involving zeroth, first and second derivatives, while the matrix B comes from the discretization of the integral operator. A standard Crank-Nicholson scheme would also involve the matrix B on the left-hand side of the equation, but we choose this modification for efficiency reasons: A is a sparse matrix coming from the derivative terms, while B is a full matrix corresponding to the integral operator. Hence, if B appears on the left hand side, each time step would involve solving a linear equation for C_{n+1} with a full matrix. In our modification we only need to solve a sparse matrix equation, which can be done more efficiently. Our modified iteration can be seen as a mixture between a Crank-Nicholson scheme and an explicit Euler scheme (for the terms involving the integral operator). Of course, an implicit Euler scheme could be used as well and this would be more stable, but again, a fully implicit scheme has the drawback that one would have to invert a full system matrix due to the integral operator term.
Assuming that the set of observation points Ω_{dat} is discrete, the observation operator O is computed by a piecewise linear interpolation of the discretized solution to the observation points.

Since we replace the infinite domain $\mathbb{R}^+$ by a finite one, we do not have the problem that g_0 is not in L^2 in the discretized case and we do not have to introduce the operator $\tilde{G}$ in (29). Thus we can minimize the functional (13). We used the discretized L^2-norm for the error term $F(a)-y$. For the regularization norm $\|.\|_s$ we have to keep in mind that the embedding condition (36) has to hold. One possible choice for a norm is the discretized version of the $H^2(\mathbb{R}^+ \times [0,T^*])$, which involves all second derivatives. For the numerical computations we used a different one - the tensor product norm $H^1[0,K_0] \otimes H^1[0,T^*]$:

$$\|u\|^2_{H^1[0,K_0]\otimes H^1[0,T^*]} = \int_0^{T^*} (\|u(.,T)\|^2_{H^1([0,K_0])} + \left(\frac{d}{dT}\|u(.,T)\|_{H^1([0,K_0])}\right)^2 dT.$$

Two reasons are responsible for this choice: First of all, this norm is weaker than the H^2-norm, as only mixed second order derivatives $\frac{\partial^2}{\partial K \partial T}u$ appear in the highest order terms, hence the exact solution does not need to be in H^2 but may have slightly less regularity. Secondly, a discretization of the tensor product norm can be easily obtained by taking the Kronecker product of the matrices corresponding to a discretization of the one-dimensional H^1-norm. The embedding condition (36) for the tensor product norm follows immediately from the one-dimensional Sobolev embedding theorem $H^1(I) \to C(I)$ (I being an interval).

With the described discretization we can numerically compute the Tikhonov functional (13) for a given discretized speed function a. The minimization of this functional was done by a Gauss-Newton method. In each step, this involves the computation of the Frechet-derivative of F and its adjoint, as well as solving a linear equation for the update $a_{new} - a_{old}$ (cf. Section 4). The numerical computation of the corresponding equations was done in the same way as for the call price equation (4) and in a consistent way such that the discretization of the derivative equals the derivative of the discretized operator. The linear equation in the Gauss-Newton step was solved by a CG-method. For that only matrix-vector operations are used so we do not have to compute the full derivative matrix, but directional derivatives are sufficient, which helps reducing the computational effort.

In order to verify the predicted convergence rates we performed some numerical results using simulated data. At first, we consider the rates in Theorem 4, i.e., the case when a source condition (45) holds. For this we specified the exact solution

$$a^\dagger(K,T) = \exp(-2(K-1)^2)\sin(K)(1+0.2T),$$

computed the forward operator $F(a^\dagger)$ and added some random noise to this to get data with noise level δ. For this simulation the setup of the problem parameters was as follows: We used a 50×50 discretization for the domain $[0,K_0] \times [0,T^*]$ with $T^* = 1$. The volatility σ and η were chosen to be constant 1, the interest rate was $r = 0.05$. As observation set we used the discretization points in the interval $[0.6,4] \times [0.1,1]$. The double exponential tail for the Lévy measure was chosen as in [6]

$$\psi(z) = \begin{cases} \frac{\beta\exp(-(G+1)|z|)}{G^2+G} & z > 0 \\ \frac{\exp(-(M-1)z)}{M^2-M} & z \geq 0 \end{cases} \qquad (\beta,G,M) = (\frac{1}{2},1,2), \tag{47}$$

which corresponds to the Kou model (see [23, 24] for details on this model).
Since in this case $a^\dagger$ is known we constructed the initial guess a^* such that (45)

holds. With the a priori regularization parameter choice $\alpha = \delta$ we computed the error $\|a_{\alpha,\delta} - a^\dagger\|$ for different choices of δ. The result is shown in Figure 1. The predicted order of convergence rate $\|a_{\alpha,\delta} - a^\dagger\| \sim \sqrt{\delta}$ is indicated by the

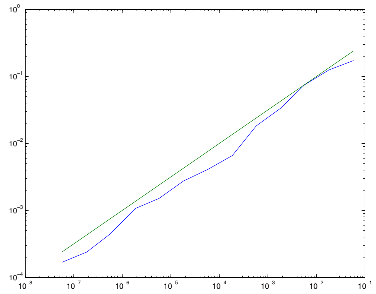

Figure 1: Error versus noise level with source condition

straight line in the picture. It can be seen that the computed convergence rate follows the theoretically predicted one.
Knowing the exact solution for this simulated example allowed an initial guess for which the source condition is satisfied. In practical applications such an initial guess might not be available and a source condition of type (45) might not hold. It is possible that a weaker source condition is fullfilled (e.g. of Hölder or logarithmic type [22]), in which case weaker convergence rates can be proven. In many applications the type of source condition and the rates thus depend on the smoothness of the exact solution $a^\dagger$. In the next example we tested how convergence rates depend on the smoothness of the exact solution.
Note that now the source condition is not known, so the parameter choice $\alpha \sim \delta$ cannot be used here (it might not give the optimal rate). We therefore rely on an a posteriori parameter choice rule [16], which has the advantage that the exact dependence of α on the noise level does not have to be known, but the regularization parameter is defined implicitly. For the computations we used the discrepancy principle, which can be described as follows: First, a geometrically decaying sequence of regularization parameters α_k is generated. For each α_k the corresponding Tikhonov regularized solution is computed and the first parameter for which the residuum is of order of the noise level is chosen as regularization parameter:

$$\alpha := \max\{\alpha_k : \|F(a_{\alpha_k,\delta}) - y_\delta\| \le \tau\delta\},$$

where $\tau > 1$ is a fixed parameter. A convergence proof of this rule for nonlinear Tikhonov regularization can be found in [25, 37, 33].
In Figure 2 we show the error $\|a_{\alpha,\delta} - a^\dagger\|$ over the noise level δ for two different solutions $a^\dagger$: The first one is a smooth Gaussian function,

$$a_1^\dagger(K,T) = \exp\left(-\frac{(K-2.5)^2}{2}\right)\exp\left(-\frac{(T-0.5)^2}{0.1}\right)$$

while the second one is chosen as the tensor product of two piece-wise constant functions

$$a_2^\dagger(K,T) = \text{pwconst}_1(K)\text{pwconst}_2(T),$$

where $\text{pwconst}_1(K)$ is nonzero in the interval $[1,4]$, and constant 1 in the interval $[2,3]$ with continuous linear interpolation in between. $\text{pwconst}_2(T)$ is of similar shape, nonzero in $[0.2,0.8]$ and constant 1 in $[0.4,0.6]$. Both functions $a_1^\dagger$ and $a_2^\dagger$ have a similar support but the first one is smooth while the second one is not even continuously differentiable (although it is in the tensor product Hilbert space $H^1[0,K_0] \otimes H^1[0,T^*]$). The dashed line in Figure 2 corresponds to the error for the non-smooth solution $a_2^\dagger$ and the solid line to the smooth one $a_1^\dagger$. It can be seen that the convergence in the non-smooth case is slower than for the smooth case. This indicates that amongst other factors the smoothness of the solution influences the convergence rate.

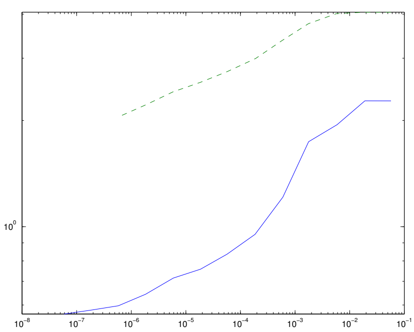

Figure 2: Rates for smooth and non-smooth speed functions

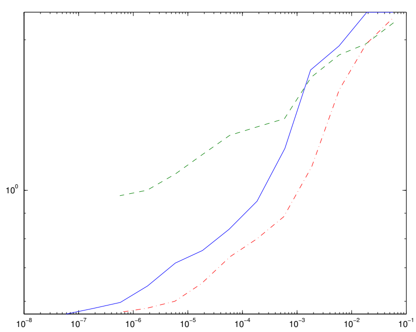

Figure 3: Rates for different kernels

In a similar experiment we tested how convergence rates depend on the smoothness of the double exponential kernel. As indicated in the discussion in Section 5, a source condition of type (45) is more difficult to fulfill for a given solution if the double exponential kernel is smoother. Thus, it is to be expected that if a solution $a^\dagger$ is kept fixed but the identification problem is computed using different kernels, the convergence rate for the case of smooth kernels should be slower. We choose three different kernels: The first one was as in the previous cases the one-times weakly differentiable kernel (47), the second one was a smooth Gaussian centered at 0 with variance $\sigma = \sqrt{5}$, the third one was a narrow Gaussian with variance $\sigma \sim 0.15$. Although the last kernel is smooth, in a numerical sense it approximates a Dirac-Delta distribution and hence can be considered a highly non-smooth kernel for the discretized problem. The results are shown in Figure 3: The dashed line corresponds to the smooth Gaussian, the dashed-dotted line to the delta-like kernel and the solid line to (47). The results confirm the prediction (see discussion at the end of Section 5) that the rates are faster for non-smooth kernels - the delta-like kernel yields the fastest rates, while the Gaussian with large variance result in the slowest ones.
It should be mentioned that other factors determine the convergence rates as well, for instance we observed that it is harder to identify features of the speed function which are located close to the boundary K_0, where the call price has little variation. This is also not surprising in view of the discussion of the source condition in Section 5.
Finally we tested the calibration procedure on real data. For that purpose we used the data given in [2] as well as the calibrated Merton jump model of that paper. The jumps in that model are log-normally distributed, i.e.

$\nu(dx, dt) = \phi(x)dxdt$ with ϕ denoting the density function of the normal distribution with mean μ and variance γ^2. The given data in [2] are bid and ask prices of European calls having maturities form 0.08 to 10 years and strikes between 0.5 and 2 (with $S_0 = 1$). For the solution to the PIDEs we used the scheme outlined in the beginning of this section and a 401×201 discretization. As "observed price" we used the arithmetic average between bid and ask price. For the initial guess for the local speed function we employed the constant jump intensity of the fitted model of [2], which is equal to 0.089. The remaining fitted parameters are: $\sigma = 0.1765$, $\mu = -0.8898$, $\gamma = 0.4505$, $r = 0.059$ and $\eta = 0.0114$. While Andersen and Andreasen in [2] fitted the local volatility leaving the jump process unchanged over time by optimizing the least-squares problem with the constraint for the model prices to fall into the bid-ask spread, we calibrate the local speed function, which only affects the jump term of the process, by the developed regularization procedure. The bid-ask spread is interpreted as the noise level δ and set to the root of the average squared difference between the midpoints of the bid and ask prices and the bid prices, which resulted in $\delta = 0.0022$. The fitted local speed function is plotted in Figure 4 in the relevant region, where data were available (the local speed function was set to 0 elsewhere). The procedure terminated after 2 main Gauss-Newton iterations and took 4 minutes on a 2.4 GHz Pentium 4 with 512 MB RAM.
It is quite obvious, that there is a sharp decrease in the local speed function around time 1.5 and strike 1.5. This is due to the fact that in this region the volatility $\sigma = 0.1765$ in the model suffices to imply call prices as high as in the data, and hence no jump term is needed to explain the option prices there.
The root of the mean squared error of the fit to the data was 0.0014, when the algorithm terminated, while it was 0.0245 for the initial guess and hence the introduction of the local speed function has considerably increased the fitting quality. Note that the error is smaller than the noise level, which is in some sense the best one can expect from a fitting procedure.

7 Conclusion and further research

We have investigated a procedure to robustly calibrate the local speed function of a given Lévy model to available European option prices. It turned out that the Tikhonov regularized problem is not only stably solvable, but one can also prove convergence rates and verify them numerically. The numerical results are quite satisfactory also in the real data case. The fact that the numerically reconstructed local speed function is quite small or even zero in some regions suggests further research on model selection in general. Another interesting open problem is the case of a pure jump model without diffusion component. From the viewpoint of calibration this situation is more delicate in both investigating the forward operator and the regularization. Finally, in the present paper we investigated the calibration of the local speed function when both the local volatility and the Lévy measure are known. The suitability of a simultaneous calibration of volatility, speed function and Lévy measure from market data will be investigated in a future study.

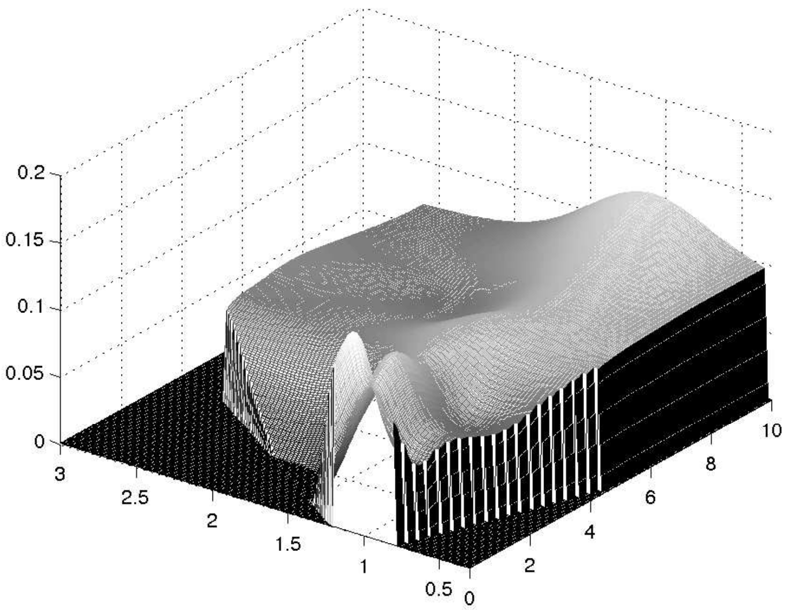

Figure 4: Local speed function for the Merton jump-diffusion model of [2]

Acknowledgement

We thank Herbert Egger for providing MATLAB code for the calibration of the Black-Scholes model.

References

[1] Y. Achdou and O. Pironneau. *Computational methods for option pricing.* SIAM, Philadelphia, PA, 2005.

[2] L. Andersen and J. Andreasen. Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research*, 4:231–262, 2000.

[3] D. Belomestny and M. Reiss. Spectral calibration of exponential levy models: Implementation. SFB 649 Discussion Paper No. 35, Berlin, 2006.

[4] D. Belomstny and M. Reiss. Spectral calibration for exponential Lévy models. *Finance and Stochastics*, 10:449–474, 2006.

[5] P. Carr, H. Geman, D. Madan, and M. Yor. Stochastic volatility for Lévy processes. *Mathematical Finance*, 13:345–382, 2003.

[6] P. Carr, D. Madan, H. Geman, and M. Yor. From local volatility to local Lévy models. *Quantitative Finance*, 4:581–588, 2004.

[7] R. Cont and P. Tankov. *Financial Modelling with Jump Processes.* Chapman and Hall, Boca Raton, FL, 2003.

[8] R. Cont and P. Tankov. Nonparametric calibration of jump-diffusion option pricing models. *Journal of Computational Finance*, 7:1–49, 2004.

[9] R. Cont and P. Tankov. Recovering exponential Lévy models from option prices: regularization of an ill-posed inverse problem. *SIAM Journal on Control and Optimization*, 43:1–25, 2006.

[10] R. Cont and E. Voltchkova. Integro-differential equations for option prices in exponential Lévy models. *Finance and Stochastics*, 9:299–325, 2005.

[11] S. Crepey. Calibration of the local volatility in a generalized Black-Scholes model using Tikhonov regularization. *SIAM Journal on Mathematical Analysis*, 34:1183–1206, 2003.

[12] E. Derman and I. Kani. Riding on a smile. *Risk*, 7:32–39, 1994.

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

[14] H. Egger and H. W. Engl. Tikhonov regularization applied to the inverse problem of option pricing: convergence analysis and rates. *Inverse Problems*, 21(3):1027–1045, 2005.

[15] H. W. Engl. *Integralgleichungen.* Springer Wien, 1997.

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

[17] H. W. Engl, K. Kunisch, and A. Neubauer. Convergence rates for Tikhonov regularization of nonlinear ill-posed problems. *Inverse Problems*, 5:523–540, 1989.

[18] L. C. Evans. *Partial differential equations.* Graduate Studies in Mathematics. 19. Providence, RI: American Mathematical Society (AMS), 1998.

[19] J.-P. Fouque, G. Papanicolaou, and K. R. Sircar. *Derivatives in Financial Markets with Stochastic Volatility.* Cambridge University Press, 2000.

[20] C. Groetsch. *The theory of Tikhonov regularization for Fredholm equations of the first kind.* Pitman, Boston, 1984.

[21] I. Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an Ito differential. *Probability Theory and Related Fields*, 71:501–516, 1986.

[22] T. Hohage. Regularization of exponentially ill-posed problems. *Numerical Functional Analysis and Optimization*, 21:439–464, 2000.

[23] S. G. Kou. A jump diffusion model for option pricing. *Management Science*, 48:1088–1101, 2002.

[24] S. G. Kou and H. Wang. First passage time for a jump diffusion process. *Advances in Applied Probability*, 35:504–531, 2003.

[25] C. Kravaris and J. H. Seinfeld. Identification of parameters in distributed parameter systems by regularization. *SIAM Journal on Control and Optimization*, 23:217–241, 1985.

[26] R. Kress. *Linear integral equations.* Springer New York, 1999.

[27] O. Ladyzhenskaya, V. Solonnikov, and N. Ural'tseva. *Linear and quasilinear equations of parabolic type. Translated from the Russian by S. Smith.* Translations of Mathematical Monographs. 23. Providence, RI: American Mathematical Society (AMS). XI, 648 p. , 1968.

[28] F. Liu and M. Z. Nashed. Tikhonov regularization of nonlinear ill-posed problems with closed operators in Hilbert scales. *Journal of Inverse and Ill-Posed Problems*, 5(4):363–376, 1997.

[29] A.-M. Matache, T. von Petersdorff, and C. Schwab. Fast deterministic pricing of options on Lévy driven assets. *Mathematical Modelling and Numerical Analysis*, 38:37–71, 2004.

[30] M. Nashed. Approximate regularized solutions to improperly posed linear integral operator equations. In D. Colton, R.G. Gilbert, editor, *Constructive and computational methods for differential and integral equations*, Lecture Notes in Mathematics, Vol. 430, pages 289–332. Springer, Berlin-Heidelberg-NewYork 1974.

[31] A. Neubauer. Tikhonov regularization of nonlinear ill-posed problems in Hilbert scales. *Applicable analysis*, 46(1-2):59–72, 1992.

[32] P. Protter. *Stochastic Integration and Differential Equations (2nd Edition).* Springer, 2004.

[33] R. Ramlau. Morozov's discrepancy principle for Tikhonov regularization of nonlinear operators. *Numerical Functional Analysis and Optimization*, 23(1-2):147–172, 2002.

[34] M. Renardy and R. C. Rogers. *An Introduction to Partial Differential Equations.* Springer, 1993.

[35] K.-I. Sato. *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, 1999.

[36] W. Schoutens, *Lévy Processes in Finance: Pricing Financial Derivatives.* Wiley, 2003.

[37] O. Scherzer, The use of Morozov's discrepancy principle for Tikhonov regularization for solving nonlinear ill-posed problems. *Computing*, 51: 45–60, 1993.