

# Spectral calibration of exponential Lévy models

Denis Belomestny•Markus Reiß

August 17, 2006

**Abstract.** We investigate the problem of calibrating an exponential Lévy model based on market prices of vanilla options. We show that this inverse problem is in general severely ill-posed and we derive exact minimax rates of convergence. The estimation procedure we propose is based on the explicit inversion of the option price formula in the spectral domain and a cut-off scheme for high frequencies as regularisation.

---

**Key words.** European option, jump diffusion, minimax rates, severely ill-posed, nonlinear inverse problem, spectral cut-off

**Mathematics Subject Classification (2000):** 60G51; 62G20; 91B28

**JEL Subject Classification:** G13; C14

## 1. Introduction

The work on calibration methods for financial models based on Lévy processes has mainly focused on certain parametrisations of the underlying Lévy process with the notable exception of Cont and Tankov (2004b). Since the characteristic triplet of a Lévy process is a priori an infinite-dimensional object, the parametric approach is always exposed to the problem of misspecification, in particular when there is no inherent economic foundation of the parameters and they are only used to generate different shapes of possible jump distributions. The goal of this paper is to investigate mathematically the problem of nonparametric inference for the Lévy triplet when the asset price  $(S_t)$  follows an exponential Lévy model

$$S_t = S e^{rt+X_t} \text{ with a Lévy process } X_t \text{ for } t \geq 0. \quad (1)$$

We suppose that at time  $t = 0$  we dispose of prices for vanilla European call and put options on this asset with different strike prices and possibly different maturities. By basing our estimation on option data we draw inference on the underlying risk neutral price process, which in general cannot be determined from historical price data due to the incompleteness of the Lévy market.

---

Weierstraß Institute for Applied Analysis and Stochastics (WIAS), Mohrenstraße 39, 10117 Berlin, Germany, e-mail: [belomest@wias-berlin.de](mailto:belomest@wias-berlin.de)

Institute of Applied Mathematics, University of Heidelberg, Im Neuenheimer Feld 294, 69120 Heidelberg, Germany, e-mail: [reiss@statlab.uni-heidelberg.de](mailto:reiss@statlab.uni-heidelberg.de)

The observed option prices will be slightly unprecise due to bid-ask spreads or other market frictions. In the ideal case of precise observations for all possible strike prices the state price density and hence the Lévy triplet can be uniquely identified using the formula by Breeden and Litzenberger (1978). Under the realistic model of finitely many noisy observations we cannot hope to determine the triplet correctly, we should rather try to provide an estimator which is as good as possible for the given accuracy of the data. This optimality property is usually assessed by the minimax paradigm, which quantifies the error in a worst case scenario and measures the inherent complexity of the statistical problem class. One of the main results of the present paper is a lower bound, showing that already in the simple exponential Lévy model the estimation problem is in general severely ill-posed, that is, the estimation error for any part of the Lévy triplet as a function of the accuracy of the observations will only converge with a logarithmic rate for any conceivable estimation procedure.

On the other hand, we propose an explicit construction of an estimator that attains this optimal minimax rate. The procedure is based on the inversion of the explicit pricing formula via Fourier transforms by Carr and Madan (1999) and a regularisation in the spectral domain. Using the Fast Fourier Transformation, the procedure is easy to implement and yields good results in simulations in view of the severe ill-posedness. Below, we discuss the features of our method in detail and compare it with the penalized least squares approach by Cont and Tankov (2004b). In comparison with standard statistical ill-posed problems, the main challenges are the nonlinearity involved and the complex interplay between the jump measure as nonparametric part and the drift and diffusion coefficient as parametric parts.

The exponential Lévy model reflects the assumption that the log returns of the asset evolve independently and with identical distribution for the same time steps, which is plausible for liquid markets and not too long time horizons. This basic model class, first introduced by Merton (1976), has been considered recently for a variety of pricing and optimisation problems in finance, cf. the recent works by Kallsen (2000), Mordecki (2002), Emmer and Klüppelberg (2004), Cont and Voltchkova (2005) and the references therein.

When no model for the price process is specified, calibration from option data can be used to estimate the state price density, see Aït-Sahalia and Duarte (2003). This density yields the distribution of the asset price at the times of maturity, but does not provide any information on the evolution of the price in time. A structural assumption on the price process allows to find prices for path-dependent options or to perform a dynamic risk management. In financial engineering information about the expected time evolution is obtained by smoothing implied Black-Scholes volatility surfaces, e.g. Fengler (2005). For the generalised Black-Scholes model Dupire's formula permits the calibration from option prices, see e.g. Jackson, Süli, and Howison (1999) for a numerical approach and Crépey (2003) for a theoretical study. The calibration of parametric exponential Lévy models has

been studied for example by Eberlein, Keller, and Prause (1998) and Carr, Geman, Madan, and Yor (2002).

After introducing the financial and statistical model in Section 2, the estimation method is developed in Section 3. The main theoretical results are formulated in Section 4. We conclude in Section 5. The proofs of the upper and lower bounds are deferred to Sections 6 and 7, respectively.

## 2. The model

### 2.1. The exponential Lévy model and option prices

Since we base our calibration on option prices, we place ourselves immediately in a risk neutral world, modeled by a filtered probability space  $(\Omega, \mathcal{F}, \mathbb{Q}, (\mathcal{F}_t))$ , on which the price process  $(S_t, t \geq 0)$  of an asset after discounting forms a martingale. As is standard in the calibration literature, the martingale measure  $\mathbb{Q}$  is assumed to be settled by the market and to be identical for all options under consideration.

We suppose that under  $\mathbb{Q}$  the process  $S_t$  follows the Lévy model (1), where  $S > 0$  is the present value of the asset and  $r \geq 0$  is the riskless interest rate, which is assumed to be known and constant. An excellent reference for this model in finance is the monograph by Cont and Tankov (2004a). In this paper we shall only consider Lévy processes  $X_t$  with a jump component of finite intensity and absolutely continuous jump distribution. Extensions to infinite intensity cases can be found in Belomestny and Reiß (2005). The characteristic function of  $X_T$  is then given by the Lévy-Khintchine representation

$$\varphi_T(u) := \mathbb{E}[e^{iuX_T}] = \exp\left(T\left(-\frac{\sigma^2 u^2}{2} + i\gamma u + \int_{-\infty}^{\infty} (e^{iux} - 1)\nu(x) dx\right)\right). \quad (2)$$

$\sigma \geq 0$  is called volatility,  $\gamma \in \mathbb{R}$  drift and the non-negative function  $\nu \in L^1(\mathbb{R})$  is the jump density with intensity  $\lambda := \|\nu\|_{L^1(\mathbb{R})}$ .

A risk neutral price at time  $t = 0$  for a European call option with strike  $K$  and maturity  $T$  is given by

$$C(K, T) = e^{-rT} \mathbb{E}[(S_T - K)^+], \quad (3)$$

where  $(A)^+ := \max(A, 0)$ . By the independence of increments in  $X_t$  the martingale condition on  $e^{-rt}S_t$  may be equivalently characterized by

$$\forall t \geq 0 : \mathbb{E}[e^{X_t}] = 1 \iff \frac{\sigma^2}{2} + \gamma + \int_{-\infty}^{\infty} (e^x - 1)\nu(x) dx = 0. \quad (4)$$

Observe that we have imposed implicitly the exponential moment condition  $\int_0^\infty (e^x - 1)\nu(x) dx < \infty$  to ensure the existence of  $\mathbb{E}[S_t]$ . Another consequence is that the characteristic function  $\varphi_T$  is defined on the whole strip  $\{z \in \mathbb{C} \mid \text{Im}(z) \in [-1, 0]\}$  in the complex plane, which will be important

later. We reduce the number of parameters by introducing the negative log-forward moneyness

$$x := \log(K/S) - rT,$$

such that the call price in terms of  $x$  is given by

$$\mathcal{C}(x, T) = S \mathbb{E}[(e^{X_T} - e^x)^+].$$

The analogous formula for the price of a put option is  $\mathcal{P}(x, T) = S \mathbb{E}[(e^x - e^{X_T})^+]$  and the well-known put-call parity is easily established:

$$\mathcal{C}(x, T) - \mathcal{P}(x, T) = S \mathbb{E}[e^{X_T} - e^x] = S(1 - e^x). \quad (5)$$

## 2.2. The observations

We focus on the calibration from options with a fixed maturity  $T > 0$  and mention the straight-forward extension to several maturities in Section 3.1. The prices of  $N$  call options (or by the put-call parity (5) alternatively put options) are observed at different strikes  $K_j$ ,  $j = 1, \dots, N$ , corrupted by noise:

$$Y_j = C(K_j, T) + \sigma_j \varepsilon_j, \quad j = 1, \dots, N. \quad (6)$$

We assume the observational noise  $(\varepsilon_j)$  to consist of independent centred random variables with  $\mathbb{E}[\varepsilon_j^2] = 1$  and  $\sup_j \mathbb{E}[\varepsilon_j^4] < \infty$ . The noise levels  $(\sigma_j)$  are assumed to be positive and known.

For observational noise with a known and smooth correlation structure the calibration problem becomes more stable. As long as no empirically validated model for the observational noise exists, we work under the assumption of independent perturbations which is canonical and least favourable.

As we need to employ Fourier techniques, we introduce the function

$$\mathcal{O}(x) := \begin{cases} S^{-1}\mathcal{C}(x, T), & x \geq 0, \\ S^{-1}\mathcal{P}(x, T), & x < 0 \end{cases} \quad (7)$$

in the spirit of Carr and Madan (1999).  $\mathcal{O}$  records normalised call prices for  $x \geq 0$  and normalised put prices for  $x \leq 0$ . The following important properties of  $\mathcal{O}$  are easily obtained from the put-call parity (5) and the martingale condition (4), see Belomestny and Reiß (2005) for the exact derivation.

### Proposition 1.

- (a) We have  $\mathcal{O}(x) = S^{-1}\mathcal{C}(x, T) - (1 - e^x)^+$  for all  $x \in \mathbb{R}$ .
- (b)  $\mathcal{O}(x) \in [0, 1 \wedge e^x]$  holds for all  $x \in \mathbb{R}$ .
- (c) If  $C_\alpha := \mathbb{E}[e^{\alpha X_T}]$  is finite for some  $\alpha \geq 1$ , then  $\mathcal{O}(x) \leq C_\alpha e^{(1-\alpha)x}$  holds for all  $x \geq 0$ .

- (d) At any  $x \in \mathbb{R} \setminus \{0\}$ , respectively  $x \in \mathbb{R} \setminus \{0, \gamma T\}$  in the case  $\sigma = 0$ , the function  $\mathcal{O}$  is twice differentiable with  $\|\mathcal{O}''\|_{L^1(\mathbb{R})} \leq 3$ . The first derivative  $\mathcal{O}'$  has a jump of height  $-1$  at zero and, in the case  $\sigma = 0$ , a jump of height  $+e^{T(\gamma-\lambda)}$  at  $\gamma T$ .
- (e) The Fourier transform of  $\mathcal{O}$  satisfies

$$\mathcal{FO}(v) := \int_{-\infty}^{\infty} \mathcal{O}(x)e^{ivx} dx = \frac{1 - \varphi_T(v-i)}{v(v-i)}, \quad v \in \mathbb{R}. \quad (8)$$

This identity extends to all complex values  $v$  with  $\text{Im}(v) \in [0, 1]$ .

Remark that an interesting way to estimate  $\gamma$  and  $\lambda$  (but not  $\nu$ ) is suggested by Proposition 1(d): a change point detection algorithm for jumps in the derivative of  $\mathcal{O}$ , as proposed by Goldenshluger, Tsybakov, and Zeevi (2006), yields an estimate of  $\gamma$  and a subsequent estimate of the jump size an estimate of  $\lambda$ .

We transform our observations  $(Y_j)$  and predictors  $(K_j)$  to

$$O_j := Y_j/S - (1 - K_j e^{-rT}/S)^+ = \mathcal{O}(x_j) + \delta_j \varepsilon_j, \quad (9)$$

$$x_j := \log(K_j/S) - rT, \quad (10)$$

where  $\delta_j = S^{-1}\sigma_j$ . In practice, the design  $(x_j)$  will be rather dense around  $x = 0$  and sparse for options further out of the money or in the money, cf. Fengler (2005) for a study on the German DAX index.

In order to facilitate the subsequent analysis we make a mild moment assumption on the price process, which guarantees by Proposition 1(b,c) the exponential decay of  $\mathcal{O}$ .

**Assumption 1** We assume that  $C_2 := \mathbb{E}[e^{2X_T}]$  is finite. This is equivalent to postulating for the asset price a finite second moment:  $\mathbb{E}[S_T^2] < \infty$ .

### 3. The method of estimation

#### 3.1. Outline of the method

Since our asset follows an exponential Lévy model, the jumps in the Lévy process appear exponentially transformed in the asset prices and it is intuitive that inference on the exponentially weighted jump measure

$$\mu(x) := e^x \nu(x), \quad x \in \mathbb{R},$$

will lead to spatially more homogeneous properties of the estimator than for  $\nu$  itself. Our calibration procedure relies essentially upon the formula

$$\begin{aligned} \psi(v) &:= \frac{1}{T} \log \left( 1 + iv(1 + iv)\mathcal{FO}(v) \right) = \frac{1}{T} \log(\varphi_T(v-i)) \\ &= -\frac{\sigma^2 v^2}{2} + i(\sigma^2 + \gamma)v + (\sigma^2/2 + \gamma - \lambda) + \mathcal{F}\mu(v), \end{aligned} \quad (11)$$

which is a simple consequence of the formulae (2) and (8). Note that the function  $\psi$  is up to a shift in the argument the cumulant-generating function of the Lévy process and a continuous version of the logarithm must be taken such that  $\psi(0) = 0$ , which is implied by the martingale condition.

Formula (11) shows that the Lévy triplet is uniquely identifiable given the observation of the whole option price function  $\mathcal{O}$  without noise:  $\mathcal{F}\mu(v)$  tends to zero as  $|v| \rightarrow \infty$  due to the Riemann-Lebesgue Lemma such that  $\psi$  is the sum of a quadratic polynomial and a function vanishing at infinity. Then  $\sigma^2$ ,  $\gamma$ ,  $\lambda$  are identifiable as coefficients in the polynomial for arguments tending to infinity. The function  $\mathcal{F}\mu(v)$  is obtained as the difference between  $\psi$  and the polynomial.

This identification procedure, however, is not stable such that the problem becomes ill-posed. Still, a properly refined application of this approach combined with a spectral regularisation method will equip us with estimators for the whole triplet  $\mathcal{T} = (\sigma^2, \gamma, \mu)$  (we parametrize Lévy triplets equivalently with  $\mu$  or  $\nu$ ).

The model (11) has a structure similar to the well-known partial linear models, but in fact there is one substantial difference: the function  $\mathcal{F}\mu$  is not supposed to be smooth, but instead it is decaying for high frequencies because we work in the spectral domain. This is also why we shall regularize the problem by discarding frequencies  $|v|$  higher than a certain cut-off value  $U$ , which depends on the noise level and the smoothness assumptions on the unknown jump density. Let us present the basic estimation procedure. Further details are specified in Section 6.1, while a more elaborate numerical implementation is presented in Belomestny and Reiß (2006).

- (a) We approximate the function  $\mathcal{O}$  by building a function  $\tilde{\mathcal{O}} \in L^1$ , approximating the true function  $\mathcal{O}$ , based on the observations  $(O_j)$ . It suffices to interpolate the data points  $(O_j)$  linearly, but in simulations it turns out that some smoothing procedure is preferable, compare the discussion in Section 3.2.
- (b) For  $\kappa(v) \in (0, 1)$ , specified later in (27), we calculate

$$\tilde{\psi}(v) := \frac{1}{T} \log_{\geq \kappa(v)} \left( 1 + iv(1 + iv)\mathcal{F}\tilde{\mathcal{O}}(v) \right), \quad v \in \mathbb{R}, \quad (12)$$

where the trimmed log-function  $\log_{\geq \kappa} : \mathbb{C} \setminus \{0\} \rightarrow \mathbb{C}$  is given by

$$\log_{\geq \kappa}(z) := \begin{cases} \log(z), & |z| \geq \kappa \\ \log(\kappa z/|z|), & |z| < \kappa \end{cases} \quad (13)$$

and  $\log(\bullet)$  is taken in such a way that  $\tilde{\psi}(v)$  is continuous with  $\tilde{\psi}(0) = 0$  (almost surely the argument of the logarithm in (12) does not vanish). If we observe option prices for different maturities  $T_k$ , we perform the steps (a) and (b) for each  $T_k$  separately and aggregate at this point the different estimators for  $\psi$  to obtain one estimator with less variance, e.g. by taking a weighted average. Similarly, estimators obtained on different days can be aggregated at this stage.

- (c) With an estimate  $\tilde{\psi}$  of  $\psi$  at hand, we obtain estimators for the parametric part  $(\sigma^2, \gamma, \lambda)$  by an averaging procedure taking into account the polynomial structure in (11). Upon fixing the spectral cut-off value  $U > 0$ , we set

$$\hat{\sigma}^2 := \int_{-U}^U \operatorname{Re}(\tilde{\psi}(u)) w_{\sigma}^U(u) du, \quad (14)$$

$$\hat{\gamma} := -\hat{\sigma}^2 + \int_{-U}^U \operatorname{Im}(\tilde{\psi}(u)) w_{\gamma}^U(u) du, \quad (15)$$

$$\hat{\lambda} := \frac{\hat{\sigma}^2}{2} + \hat{\gamma} - \int_{-U}^U \operatorname{Re}(\tilde{\psi}(u)) w_{\lambda}^U(u) du, \quad (16)$$

where the weight functions  $w_{\sigma}^U$ ,  $w_{\gamma}^U$  and  $w_{\lambda}^U$  satisfy

$$\int_{-U}^U \frac{-u^2}{2} w_{\sigma}^U(u) du = 1, \quad \int_{-U}^U u w_{\gamma}^U(u) du = 1, \quad \int_{-U}^U w_{\lambda}^U(u) du = 1; \quad (17)$$

$$\int_{-U}^U w_{\sigma}^U(u) du = 0, \quad \int_{-U}^U u^2 w_{\lambda}^U(u) du = 0. \quad (18)$$

The estimate of the coefficients can be understood as an orthogonal projection estimate with respect to a weighted  $L^2$ -scalar product.

- (d) Finally, we define the estimator for  $\mu$  as the inverse Fourier transform of the remainder:

$$\hat{\mu}(x) := \mathcal{F}^{-1} \left[ \left( \tilde{\psi}(\bullet) + \frac{\hat{\sigma}^2}{2}(\bullet - i)^2 - i\hat{\gamma}(\bullet - i) + \hat{\lambda} \right) \mathbf{1}_{[-U, U]}(\bullet) \right](x). \quad (19)$$

Then the identity  $\mathcal{F}\hat{\mu}(0) = -\frac{\hat{\sigma}^2}{2} - \hat{\gamma} + \hat{\lambda}$  shows that the estimated triplet still satisfies the martingale condition (4).

### 3.2. Discussion of the method

First note that the computational complexity of the estimation procedure is very low. Step (a) is a standard interpolation or regression estimation procedure, which is well established and fast. The only time consuming steps are the three integrations in step (c) and the (fast) Fourier transforms in steps (a) and (d).

In step (a) a reasonable approximation of  $\mathcal{FO}$  based on discrete data must be found. Asymptotically, it suffices to use simple linear interpolation because all regularisation takes place later in the spectral domain by damping high frequencies. Depending on the observation design  $(x_j)$  and the noise levels  $(\delta_j)$ , it may nevertheless pay off to invest more in obtaining a good approximating function  $\tilde{\mathcal{O}}$ . When the distance  $\Delta_j = x_j - x_{j-1}$  between the transformed prices is rather large compared to the noise level  $\delta_j$ , the numerical approximation error prevails and higher order interpolation schemes might significantly reduce the total error  $\|\tilde{\mathcal{O}} - \mathcal{O}\|_{L^1}$  if  $\mathcal{O}$  is smooth. In Proposition 2 below, we only take advantage of the fact that  $\mathcal{O}$  is almost

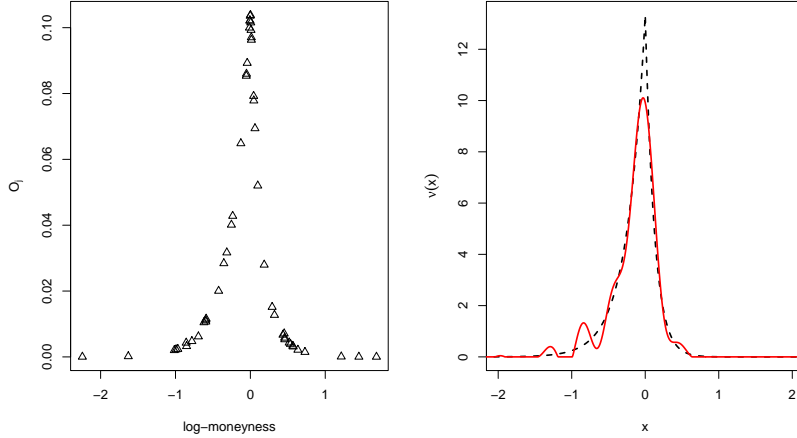
everywhere twice differentiable, but higher regularity will persist for regular jump densities. Note that for a financial Lévy model it is quite reasonable to assume that the Lévy measure is absolutely continuous and has even a smooth density (at least off the origin). Prices are conceived by a large number of agents on the market who in addition all share some uncertainty about possible jump sizes, which smears out possible point masses.

As usually, the estimation procedure is specified by certain parameters. The stabilisation of the logarithm by the function  $\kappa(v)$  is enforced mainly for theoretical reasons to prevent explosions in the logarithm due to large deviations, its practical importance is minor. For the weights  $w_\sigma^U$ ,  $w_\gamma^U$ ,  $w_\lambda^U$  it suffices to use weight functions satisfying (29) below for some  $s_{max}$ , which should be chosen larger than the expected regularity of the jump density, see Belomestny and Reiß (2006) for an example. The value  $s_{max}$  has the same meaning as the order of the kernel in nonparametric estimation. As for classical kernel estimators, the exact choice of the weight functions is not very critical. We are thus left with one important tuning parameter, the spectral cut-off frequency  $U$ . In Theorem 1 an asymptotically optimal choice is given, while Belomestny and Reiß (2006) discuss some methods to determine  $U$  directly from the data. Note, however, that a proper mathematical analysis for these completely data-driven (i.e., unsupervised) choices of  $U$  seems challenging due to the underlying nonlinear 'change point detection'-structure, for which a data-driven algorithm even in the idealized linear setting of Goldenshluger, Tsybakov, and Zeevi (2006) is not yet available.

While the spectral calibration method is here only applied to the nonparametric estimation of the Lévy triplet in an exponential Lévy model, it is more generally applicable. Suppose we prescribe a finite-dimensional parametric model for the Lévy measure. Then we can follow steps (a) through (c) and fit the remainder term in step (d) to the parameters by a least-squares criterion. In comparison with the classical least-squares approach this has the advantage of yielding faster algorithms, which are also more robust due to the variance reduction caused by the spectral cut-off. Moreover, many more financial models have been propagated where the option price and the model parameters are linked by a relationship in the spectral domain, cf. Duffie, Filipovic, and Schachermayer (2003) and the references therein. Although each model needs to be analyzed in detail, the general principles of the spectral calibration method will apply.

Let us finally make a comparison with the nonlinear penalized least-squares (PLS) approach by Cont and Tankov (2004b) for the same calibration problem. There an exponential Lévy model is selected as a prior and exponential Lévy models are considered that are obtained by a martingale measure equivalent to the prior. For each model the sum of squared distances between observed and model option prices is penalized by the relative entropy with respect to the prior. The estimated triplet is obtained by minimizing this penalized least squares criterion. In practice, the Lévy measure is approximated by a finite-dimensional collection of point measures and the minimizer is found by an iterative descent algorithm.





**Fig. 1.** Kou model. Left: Simulated option price data ( $O_j$ ) as function of  $x_j$ . Right: Calibrated Lévy density  $\hat{\nu}$  and true  $\nu$  (dashed).

Both methods display certain distinctive features. In the PLS method the penalisation plays a double role, it uses prior knowledge for the unknown Lévy triplet and it stabilizes the inverse problem. Because of the relative entropy penalisation, the volatility  $\sigma$  is fixed in advance by the prior and cannot be calibrated. In comparison, the spectral calibration (SC) method profits from the regularity of the jump density for regularizing the inverse problem. In particular, the volatility can be calibrated. The SC estimators depend only on one critical tuning parameter, the cut-off frequency  $U$ , and given  $U$  they are always uniquely defined.

In simulations the experience is that both methods yield roughly comparable results. The PLS method fits better the option price function, while the SC method performs better in estimating the Lévy triplet. The PLS method has difficulties in estimating the jump density near zero, because there are no smoothness constraints, while the SC method can yield negative values for jump densities, which can be corrected a posteriori (Belomestny and Reiß 2006). Because it avoids a numerical minimisation procedure, the SC method is considerably faster than the PLS method.

### 3.3. A numerical example

We consider the double exponential jump diffusion model proposed by Kou (2002), where the Lévy triplet is specified by the jump density

$$\nu(x) = \lambda \left( p \lambda_+ e^{-\lambda_+ x} \mathbf{1}_{[0, \infty)}(x) + (1 - p) \lambda_- e^{\lambda_- x} \mathbf{1}_{(-\infty, 0)}(x) \right), \quad x \in \mathbb{R}, \quad (20)$$

and the parameters  $\sigma, \lambda, \lambda_+, \lambda_- \geq 0$  and  $p \in [0, 1]$ , while  $\gamma$  is uniquely determined by the martingale condition. We simulate the Kou model with

parameters  $\sigma = 0.1, \lambda = 5, \lambda_- = 4, \lambda_+ = 8, p = 1/3$  and apply the non-parametric calibration procedure given the noisy observation of  $N = 50$  European options with maturity  $T = 0.25$ , interest rate  $r = 0.06$  and noise levels  $\delta_j = \mathcal{O}(x_j)/10$ . The strike prices giving rise to the design points  $(x_j)$  have been obtained by a random sample which yields more option prices at the money than further in or out of the money.

In this example we use a standard procedure based on cubic smoothing splines with cross validation for step (a) of the algorithm. The spectral cut-off frequency  $U$  is selected in a data-driven way by looking for the values where the estimates stabilize when  $U$  increases. As a postprocessing step the estimated Lévy density is corrected to ensure that it is non-negative. A more precise description of the implementation of the entire procedure is given in Belomestny and Reiß (2006).

In Figure 1 (left) the simulated observations  $(O_j)$  are depicted as a function of the corresponding log-forward moneyness  $(x_j)$  on the horizontal axis. The calibrated Lévy density  $\hat{\nu}$  is shown in Figure 1 (right) together with the true density  $\nu$  from (20). The parameters were estimated as  $\hat{\sigma} = 0.131, \hat{\lambda} = 4.983, \hat{\gamma} = 0.424$  ( $\gamma = 0.423$ ). We observe that the calibration recovers the main features of the Lévy triplet like the magnitude of the volatility and jump intensity or the mode and the skewness of the jump density.

Simulations show that for twice as many data points ( $N = 100$ ) and observations with half as much noise ( $\delta_j = \mathcal{O}(x_j)/20$ ) the calibration results are already very satisfactory. Usually, the quality of the estimators depends slightly more on the noise level ( $\delta_j$ ) than on the number of observations  $N$ . The double-exponential jump density in the Kou model is difficult to estimate because of its nondifferentiability at zero. For smoother jump densities, as in the Merton (1976) model, even better estimates are obtained. In Belomestny and Reiß (2006) further calibration results are presented.

## 4. Risk bounds

### 4.1. The main results

We shall use throughout the notation  $A \lesssim B$  if  $A$  is bounded by a constant multiple of  $B$ , independently of the parameters involved, that is, in the Landau notation  $A = \mathcal{O}(B)$ . Equally  $A \gtrsim B$  means  $B \lesssim A$  and  $A \sim B$  stands for  $A \lesssim B$  and  $A \gtrsim B$  simultaneously.

To assess the quality of the estimators, we quantify their risks under a smoothness condition of order  $s$  on the transformed jump density  $\mu$ .

**Definition 1.** For  $s \in \mathbb{N}$  and  $R, \sigma_{max} > 0$  let  $\mathcal{G}_s(R, \sigma_{max})$  denote the set of all Lévy triplets  $\mathcal{T} = (\sigma^2, \gamma, \mu)$ , satisfying the martingale condition and Assumption 1 with  $C_2 \leq R$ , such that  $\mu$  is  $s$ -times (weakly) differentiable and

$$\sigma \in [0, \sigma_{max}], \quad |\gamma|, \lambda \in [0, R], \quad \max_{0 \leq k \leq s} \|\mu^{(k)}\|_{L^2(\mathbb{R})} \leq R, \quad \|\mu^{(s)}\|_{L^\infty(\mathbb{R})} \leq R.$$

	$\sigma^2$	$\gamma$	$\lambda$	$\mu$
$\sigma_{max} > 0$	$ \log(\varepsilon) ^{-(s+3)/2}$	$ \log(\varepsilon) ^{-(s+2)/2}$	$ \log(\varepsilon) ^{-(s+1)/2}$	$ \log(\varepsilon) ^{-s/2}$
$\sigma_{max} = 0$	0	$\varepsilon^{(2s+4)/(2s+5)}$	$\varepsilon^{(2s+2)/(2s+5)}$	$\varepsilon^{2s/(2s+5)}$

**Table 1.** The minimax rates  $v_{q, \sigma_{max}}$  for the different parameters  $q$ .

Since the underlying Lévy triplet is only identifiable if  $\mathcal{O}(x)$  is known for all  $x \in \mathbb{R}$ , we consider the asymptotics of a growing number of observations with

$$\Delta := \max_{j=2, \dots, N} (x_j - x_{j-1}) \rightarrow 0 \quad \text{and} \quad A := \min(x_N, -x_1) \rightarrow \infty. \quad (21)$$

In contrast to standard regression estimates we shall always track explicitly the dependence on the level  $(\delta_k)$  of the noise in the observations, which is usually rather small for observed option prices. We shall use the norms  $\|\delta\|_{l^\infty} := \sup_k \delta_k$  and  $\|\delta\|_{l^2}^2 := \sum_k \delta_k^2$ . The subsequent analysis can certainly be improved for a concrete design  $(x_k)$  and concrete noise levels  $(\delta_k)$ , but for revealing the main features it is more transparent and concise to state the results in terms of the abstract noise level

$$\varepsilon := \Delta^{3/2} + \Delta^{1/2} \|\delta\|_{l^\infty}, \quad (22)$$

comprising the level of the numerical interpolation error and of the statistical error simultaneously. If the level  $\|\delta\|_{l^\infty}$  of the observational noise remains constant in the asymptotics, then we are in the typical regression setting where the noise level is governed by  $\Delta^{1/2}$ , the square root of the distance of observations. If, however, the observational noise level  $\|\delta\|_{l^\infty}$  tends to zero, then much faster rates can be achieved.

We now state the main results about the risk upper bounds of the estimators obtained by the basic procedure, given the specific choices in Section 6.1, and about the risk lower bounds valid for any estimation procedure whatsoever. The proofs are given in Sections 6 and 7 for the upper and lower bounds, respectively.

**Theorem 1.** Assume  $e^{-A} \lesssim \Delta^2$  and  $\Delta \|\delta\|_{l^2}^2 \lesssim \|\delta\|_{l^\infty}^2$ . For any  $\bar{\sigma} > \sigma_{max}$  we choose

$$U_{\bar{\sigma}} := \bar{\sigma}^{-1} (2 \log(\varepsilon^{-1})/T)^{1/2}, \quad U_0 := \varepsilon^{-2/(2s+5)}, \quad (23)$$

in the cases  $\sigma_{max} > 0$  and  $\sigma_{max} = 0$ , respectively. Then every estimator  $\hat{q} \in \{\hat{\sigma}^2, \hat{\gamma}, \hat{\lambda}, \hat{\mu}\}$  for the corresponding parameter  $q$  satisfies the following asymptotic risk bound:

$$\sup_{T \in \mathcal{G}_s(R, \sigma_{max})} \mathbb{E}_T [\|\hat{q} - q\|^2]^{1/2} \lesssim v_{q, \sigma_{max}},$$

where  $\|\bullet\|$  denotes the absolute value for  $q \in \{\sigma^2, \gamma, \lambda\}$  and the  $L^2(\mathbb{R})$ -norm for  $q = \mu$  and the rate  $v_{q, \sigma_{max}}$  is given in Table 1.

The two assumptions in the theorem are not very severe: because of the exponential decay of  $\mathcal{O}$  the width  $A$  of the design only needs to grow logarithmically and the error levels  $(\delta_k)$  need only be square summable after renormalisation. The latter condition could be further relaxed since this term is caused by a rough bound on the quadratic remainder term.

For the lower bounds we appeal to the equivalence between the regression and the Gaussian white noise model, as established by Brown and Low (1996), and consider merely the idealized observation model

$$dZ(x) = \mathcal{O}(x) dx + \varepsilon dW(x), \quad x \in \mathbb{R}, \quad (24)$$

with the noise level asymptotics  $\varepsilon \rightarrow 0$ , a two-sided Brownian motion  $W$  and with  $\mathcal{O} = \mathcal{O}_{\mathcal{T}}$  denoting the option price function from (7) for the given triplet  $\mathcal{T}$ . Here, the noise level  $\varepsilon$  corresponds exactly to the statistical regression error  $\Delta^{1/2} \|\delta\|_{l^\infty}$ . Due to Assumption 1 the option price functions  $\mathcal{O}$  decrease exponentially and the results by Brown and Low (1996) remain valid for unbounded intervals. This simplification avoids tedious numerical approximations in the proofs.

**Theorem 2.** *Let  $s \in \mathbb{N}$ ,  $R > 0$  and  $\sigma_{max} \geq 0$  be given. For the observation model (24) and any quantity  $q \in \{\sigma^2, \gamma, \lambda, \mu\}$  the following asymptotic risk lower bounds hold:*

$$\inf_{\hat{q}} \sup_{\mathcal{T} \in \mathcal{G}_s(R, \sigma_{max})} \mathbb{E}_{\mathcal{T}}[\|\hat{q} - q\|^2]^{1/2} \gtrsim v_{q, \sigma_{max}},$$

where  $\|\bullet\|$  denotes the absolute value for  $q \in \{\sigma^2, \gamma, \lambda\}$  and the  $L^2(\mathbb{R})$ -norm for  $q = \mu$ , the infimum is always taken over all estimators, that is all measurable functions of the observation  $Z$ , and the rate  $v_{q, \sigma_{max}}$  is given in Table 1. Consequently, the spectral estimators are rate-optimal.

#### 4.2. Discussion of the results

As we want to identify the Lévy triplet exactly in the limit, we have to assume the asymptotics  $\Delta \rightarrow 0$  and  $A \rightarrow \infty$  in the upper bound result. The numerical interpolation error term  $\Delta^{3/2}$  contained in  $\varepsilon$  can be made smaller by using higher-order schemes, see Section 3.2. On the other hand, the statistical error term  $\Delta^{1/2} \|\delta\|_{l^\infty}$  cannot be avoided as proved by the lower bound. Another way to study the calibration problem is to keep the number  $N$  of observations fixed and just to consider the asymptotics  $\|\delta\|_{l^\infty} \rightarrow 0$ . In this case the original Lévy triplet is not identifiable and the triplet of interest has to be properly defined in the set of triplets giving rise to the uncorrupted option prices, cf. Cont and Tankov (2005) for a minimum relative entropy approach.

We observe that for  $\sigma > 0$  the rate corresponds to a severely ill-posed problem, while for known  $\sigma = 0$  the rates are much better, but still ill-posed compared to those obtained in classical nonparametric regression. The severe ill-posedness in the case  $\sigma > 0$  is due to an underlying deconvolution

problem with the Gaussian kernel of variance  $\sigma^2$ : the law of the diffusion part of  $X_T$  is convolved with that of the compound Poisson part to give the density of  $X_T$ . This type of estimation problem has been studied thoroughly by Butucea and Matias (2005) in an idealized density estimation setup. Note the general order in which the (asymptotic) quality of estimation decreases:  $\sigma^2$ ,  $\gamma$ ,  $\lambda$  and finally  $\mu$ , which is related to the domination property formulated in Aït-Sahalia and Jacod (2006). For small values of  $\sigma$  and finite samples the performance is not so bad, compare the simulations in Section 3.3; it just needs a lot more observations to improve on that.

At first sight the rates for the parametric estimation part are astonishing. They are worse than in usual semi-parametric problems which also indicates that misspecified parametric models will give unreliable estimates for the volatility and jump intensity. In the case  $\sigma = 0$ , however, these rates are easily understood when employing the language of distributions. With  $\delta_0$  denoting the Dirac measure in zero and  $\delta'_0$  its derivative we have

$$\log(\varphi_T(u)) = T\mathcal{F}(\gamma\delta'_0 + \nu - \lambda\delta_0)(u).$$

Estimating the density of  $X_T$  and similarly its characteristic function from the noisy observations of  $\mathcal{O}$  amounts roughly to differentiate the observed function twice, cf. Aït-Sahalia and Duarte (2003) and the remark after equation (34) below. This gives the minimax rate for  $\nu$  and  $\mu$  as that of estimating the second derivative of a regression function of regularity  $s + 2$ . For the parameter  $\lambda$  it suffices to estimate the jump in the antiderivative of  $\mathcal{F}^{-1}(\log(\varphi_T))$ , which corresponds to a pointwise estimation problem in the first derivative of a regression function, while for  $\gamma$  the analogy is the estimation of the regression function itself at zero. This explains also why in the class  $\mathcal{G}_s$  we have measured the regularity not only in  $L^2$ , but also uniformly. In fact, if we only assume an  $L^2$ -Sobolev condition, then the same lower bound techniques will yield slower rates for the parameters, as is typical for pointwise estimation problems.

Observe that the estimation of the jump density at zero is only possible by imposing a certain regularity there, otherwise it is clearly not possible to detect jumps of height zero.

## 5. Conclusion

We have developed an estimation procedure for the nonparametric calibration of exponential Lévy models which is mathematically satisfying because of its minimax properties and which yields a straight-forward algorithm for the implementation. The corresponding lower bound results show that the calibration is in general a hard problem to solve, at least if high accuracy is desired. Nevertheless, the estimation procedure is well suited to gain general insight into the size of the parameters and the structure of the jump density. Even if reasonable parametric models exist that can be better fitted, a goodness-of-fit test based on our nonparametric approach should always be used to check against model misspecification.

Our procedure can be adapted to different models as long as the inverse transformation from the option prices to the characteristic function can be calculated and the unknown quantities can be determined from the structure of the characteristic function. In fact, the inhomogeneity in empirical jump densities across maturities (see Cont and Tankov (2004a) and Belomestny and Reiß (2006)) suggests that the exponential Lévy model should be extended, for example by abandoning homogeneity in time. In conclusion we believe that the question of calibration for models in financial mathematics should be addressed with the same rigour and intensity as other primary questions like pricing, hedging and risk management.

## 6. Proof of the upper bounds

All calculations take place in the setting of Section 4. To facilitate the calculations we introduce the exponentially increasing function

$$\mathcal{E}(x) := \frac{e^x - 1}{x}, \quad x > 0, \text{ and set } \mathcal{E}(0) := 1. \quad (25)$$

### 6.1. Specification of the method

In step (a) we interpolate the data  $(O_j)$  by setting

$$\tilde{\mathcal{O}}(x) = \beta_0(x) + \sum_{j=1}^N O_j b_j(x), \quad x \in \mathbb{R},$$

where  $(b_j)$  are linear splines and the function  $\beta_0$  is added to take care of the jump in the derivative of  $\mathcal{O}$  at zero:  $\beta'_0(0+) - \beta'_0(0-) = -1$ . We choose  $b_k$ ,  $k = 1, \dots, N$ , as the linear B-spline with knots at  $x_{k-1}$ ,  $x_k$ ,  $x_{k+1}$  and  $\beta_0$  as the linear spline with knots at  $x_{j_0-1}$ ,  $0$ ,  $x_{j_0}$  and with  $\beta_0(x_{j-1}) = \beta_0(x_j) = 0$ ,  $\beta_0(0) = x_{j-1}x_j/(x_{j-1} - x_j)$ , where the index  $j_0$  is defined by  $x_{j_0-1} < 0 < x_{j_0}$  (excluding the improbable case  $x_j = 0$ ). To ease the mathematical treatment of the extrapolation error, we assume that all data is contained in the interval  $(-A - \Delta, A + \Delta)$ . Adding the extrapolated design points  $x_0 = -A - \Delta$  and  $x_{N+1} = A + \Delta$ , we set  $\tilde{\mathcal{O}}(x_0) = \tilde{\mathcal{O}}(x_{N+1}) = 0$ . As bias we encounter the following linear interpolation of  $\mathcal{O}$

$$\mathcal{O}_l(x) := \mathbb{E}[\tilde{\mathcal{O}}(x)] = \sum_{j=1}^N \mathcal{O}(x_j) b_j(x) + \beta_0(x), \quad x \in \mathbb{R}. \quad (26)$$

More generally, we merely need to ensure for step (a) that the results of Proposition 2 and estimate (35) are satisfied.

We have enforced  $|\tilde{\psi}_T(v)| \geq \log(\kappa(v))$  in (12) to prevent unboundedness in the case of large stochastic errors. For Lévy triplets in  $\mathcal{G}_s(R, \sigma_{max})$  a reasonable choice for  $\kappa(v)$  can be obtained from the following calculation using the identity  $\frac{\sigma^2}{2} + \gamma + \mathcal{F}\mu(0) = \lambda$  derived from the martingale condition (4):

$$\frac{1}{2}|\varphi_T(v - i)| = \frac{1}{2} \exp\left(-T \frac{\sigma^2}{2} v^2 - T \mathcal{F}\mu(0) + T \operatorname{Re}(\mathcal{F}\mu(v))\right)$$

$$\geq \frac{1}{2} \exp\left(-T \frac{\sigma_{max}^2}{2} v^2 - 4TR\right) =: \kappa(v). \quad (27)$$

The only reason for the factor  $1/2$  is the mathematical tractability giving later the bound of Lemma 1.

Concerning the choice of the weight functions, we take advantage of the smoothness  $s$  of  $\mu$  by taking functions  $w$  such that  $\mathcal{F}w$  has  $s$  vanishing moments. Equivalently expressed in the spectral domain, the weight functions  $w(u)$  grow with frequencies  $|u|$  like  $|u|^s$  to profit from the decay of  $|\mathcal{F}\mu(u)|$ . Hence, we define for all  $U > 0$  families of weight functions by rescaling:

$$w_\sigma^U(u) = U^{-3} w_\sigma^1(u/U), \quad w_\gamma^U(u) = U^{-2} w_\gamma^1(u/U), \quad w_\lambda^U(u) = U^{-1} w_\lambda^1(u/U), \quad (28)$$

where the functions  $w_\sigma^1, w_\gamma^1, w_\lambda^1$  satisfy conditions (17) and (18) as well as

$$\mathcal{F}(w_\sigma^1(u)/u^s), \mathcal{F}(w_\gamma^1(u)/u^s), \mathcal{F}(w_\lambda^1(u)/u^s) \in L^1(\mathbb{R}). \quad (29)$$

In addition the support of the weight functions  $w_\sigma^1, w_\gamma^1, w_\lambda^1$  is assumed to be contained in  $[-1, 1]$ . Note that the property  $\mathcal{F}(w(u)/u^s) \in L^1(\mathbb{R})$  means in particular that  $w(u)/u^s$  is continuous and bounded such that

$$|w_\sigma^U(u)| \lesssim U^{-(s+3)} |u|^s, \quad |w_\gamma^U(u)| \lesssim U^{-(s+2)} |u|^s, \quad |w_\lambda^U(u)| \lesssim U^{-(s+1)} |u|^s. \quad (30)$$

## 6.2. A numerical approximation result

**Proposition 2.** *Under the hypothesis  $e^{-A} \lesssim \Delta^2$  we obtain uniformly over all Lévy triplets satisfying Assumption 1*

$$\sup_{u \in \mathbb{R}} |\mathbb{E}[\mathcal{F}\tilde{\mathcal{O}}(u) - \mathcal{F}\mathcal{O}(u)]| = \sup_{u \in \mathbb{R}} |\mathcal{F}\mathcal{O}_l(u) - \mathcal{F}\mathcal{O}(u)| \lesssim \Delta^2. \quad (31)$$

*Proof.* By standard Fourier estimates the assertion follows once we have proved  $\|\mathcal{O}_l - \mathcal{O}\|_{L^1} \lesssim \Delta^2$ .

Note that  $\mathcal{O} - \beta_0$  is twice differentiable except at the points  $x_{j_0-1}, 0, x_{j_0}$  and possibly  $\gamma T$  by Proposition 1(d). Moreover,  $\mathcal{O} - \beta_0$  has a derivative near zero which is uniformly bounded by a constant  $C_0$ , which follows from the  $L^1$ -estimate for  $\mathcal{O}''$  in Proposition 1(d).

Starting with the case  $\sigma > 0$ , we obtain the classical quadrature estimate for the trapezoidal rule using the mean value theorem:

$$\int_{x_1}^{x_N} |\mathcal{O}_l(x) - \mathcal{O}(x)| dx \leq \|\mathcal{O}''\|_{L^1} \Delta^2 + 2C_0 \Delta^2.$$

By Assumption 1 and Proposition 1(b,c) the extrapolation error is bounded by

$$\int_{[x_0, x_1] \cup [x_N, x_{N+1}]} |\mathcal{O}_l(x) - \mathcal{O}(x)| dx \leq 4C_2 \Delta e^{-(A-\Delta)}.$$

An application of Proposition 1(d) therefore shows for  $\sigma > 0$

$$\int_{-\infty}^{\infty} |\mathcal{O}_l(x) - \mathcal{O}(x)| dx \leq e^{-A} + 3\Delta^2 + 2C_0\Delta^2 + 4C_2\Delta e^{-(A-\Delta)} \lesssim \Delta^2.$$

In the case  $\sigma = 0$  we consider the index  $j_*$  with  $x_{j_*-1} \leq \gamma T < x_{j_*}$  and face an additional error estimated by

$$\begin{aligned} \int_{x_{j_*-1}}^{x_{j_*}} |\mathcal{O}_l(x) - \mathcal{O}(x)| dx &\leq \int_{x_{j_*-1}}^{x_{j_*}} \|(\mathcal{O} - \beta_0)'\|_{L^\infty} \left| \frac{2(x - x_{j_*-1})(x_{j_*} - x)}{x_{j_*} - x_{j_*-1}} \right| dx \\ &\leq \|(\mathcal{O} - \beta_0)'\|_{L^\infty} (x_{j_*} - x_{j_*-1})^2 \end{aligned}$$

We infer that this error term is also of order  $\Delta^2$  and thus does not enlarge the convergence rate.

### 6.3. Upper bound for $\sigma^2$

The rate for  $\sigma^2$  follows once the general risk estimate

$$\mathbb{E}[|\hat{\sigma}^2 - \sigma^2|^2] \lesssim U^{-2(s+3)} + \mathcal{E}(T\sigma^2 U^2)U^{-1}\varepsilon^2 + \mathcal{E}(T\sigma_{\max}^2 U^2)^2 U^4 \varepsilon^4 \quad (32)$$

has been shown for  $U \lesssim \Delta^{-1}$  uniformly over  $\mathcal{G}_s(R, \sigma_{\max})$ , since the explicit choice of  $U$  renders the second and third term asymptotically negligible.

Consider in the definition (12) of  $\tilde{\psi}$  separately the linearisation  $\mathcal{L}$ , neglecting the stabilisation by  $\kappa$ , and the remainder term  $\mathcal{R}$ :

$$\mathcal{L}(u) := T^{-1}\varphi_T(u-i)^{-1}(u-i)u\mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O})(u), \quad (33)$$

$$\mathcal{R}(u) := \tilde{\psi}(u) - \psi(u) - \mathcal{L}(u). \quad (34)$$

When neglecting the remainder term, we may view  $\tilde{\psi}(u)$  as observation of  $\psi(u)$  in additive noise, whose intensity grows like  $|\varphi_T(u-i)|^{-1}|(u-i)u| \sim u^2 e^{T\sigma^2 u^2}$  for  $|u| \rightarrow \infty$ . This heteroskedasticity reflects the degree of ill-posedness of the estimation problem.

**Lemma 1.** *For all  $u \in \mathbb{R}$  the remainder term satisfies*

$$|\mathcal{R}(u)| \leq T^{-1}\kappa(u)^{-2}(u^4 + u^2)|\mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O})(u)|^2.$$

*Proof.* Let us set  $\tilde{\varphi}_T(u-i) := 1 - u(u-i)\mathcal{F}\tilde{\mathcal{O}}(u)$  which equals  $e^{T\tilde{\psi}(u)}$  if  $|\tilde{\varphi}_T(u-i)| \geq \kappa(u)$ . Using  $|e^{T\tilde{\psi}(u)}| \geq \kappa(u)$ ,  $u \in \mathbb{R}$ , we obtain by a second-order expansion of the logarithm

$$\begin{aligned} |T\tilde{\psi}(u) - \log(\varphi_T(u-i)) - \varphi_T(u-i)^{-1}(e^{T\tilde{\psi}(u)} - \varphi_T(u-i))| \\ \leq \frac{1}{2}\kappa(u)^{-2}|e^{T\tilde{\psi}(u)} - \varphi_T(u-i)|^2. \end{aligned}$$

This gives the result whenever  $|\tilde{\varphi}_T(u-i)| \geq \kappa(u)$ . For the other values  $u$  the inequalities  $|\tilde{\varphi}_T(u-i)| < \kappa(u) \leq |\varphi_T(u-i)|/2$  imply  $1 \leq |\tilde{\varphi}_T(u-i) - \varphi_T(u-i)|\kappa(u)^{-1}$  and hence

$$|\varphi_T(u-i)^{-1}(e^{T\tilde{\psi}(u)} - \tilde{\varphi}_T(u-i))| \leq \frac{1}{2}\kappa(u)^{-2}|\tilde{\varphi}_T(u-i) - \varphi_T(u-i)|^2$$



$$= \frac{1}{2}\kappa(u)^{-2}(u^4 + u^2)|\mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O})(u)|^2.$$

Together with the previous result this gives for all  $u \in \mathbb{R}$  the assertion of the lemma.  $\square$

We shall frequently use the following norm bounds for the B-splines  $(b_k)$ , which follow from  $\|b_k\|_\infty = 1$  and  $|x_{k+1} - x_{k-1}| \leq 2\Delta$ :

$$\|\mathcal{F}b_k\|_{L^2} = \sqrt{2\pi}\|b_k\|_{L^2} \leq (4\pi\Delta)^{1/2}, \quad \|\mathcal{F}b_k\|_\infty \leq \|b_k\|_{L^1} \leq 2\Delta. \quad (35)$$

We decompose  $\hat{\sigma}^2$  in terms of  $\mathcal{L}$  and  $\mathcal{R}$  from (33) and (34):

$$\begin{aligned} \hat{\sigma}^2 &= \int_{-U}^U \left( -\frac{\sigma^2}{2}(u^2 - 1) + \gamma + \operatorname{Re}(\mathcal{F}\mu(u)) - \lambda + \operatorname{Re}(\mathcal{L}(u) + \mathcal{R}(u)) \right) w_\sigma^U(u) \, du \\ &= \sigma^2 + \int_{-U}^U \operatorname{Re}(\mathcal{F}\mu(u) + \mathcal{L}(u) + \mathcal{R}(u)) w_\sigma^U(u) \, du, \end{aligned} \quad (36)$$

which yields

$$\begin{aligned} \mathbb{E}[|\hat{\sigma}^2 - \sigma^2|^2] &\leq 3 \left| \int_{-U}^U \mathcal{F}\mu(u) w_\sigma^U(u) \, du \right|^2 + 3 \mathbb{E} \left[ \left| \int_{-U}^U \mathcal{L}(u) w_\sigma^U(u) \, du \right|^2 \right] \\ &\quad + 3 \mathbb{E} \left[ \left| \int_{-U}^U \mathcal{R}(u) w_\sigma^U(u) \, du \right|^2 \right]. \end{aligned}$$

Let us consider the three terms in the sum separately. The nuisance of  $\mathcal{F}\mu$  causes a deterministic error which can be bounded using  $(iu)^s \mathcal{F}\mu(u) = \mathcal{F}\mu^{(s)}(u)$  and the Plancherel isometry by

$$\begin{aligned} \left| \int_{-U}^U \mathcal{F}\mu(u) w_\sigma^U(u) \, du \right| &= 2\pi \left| \int_{-\infty}^\infty \mu^{(s)}(x) \overline{\mathcal{F}^{-1}(w_\sigma^U(u)/(iu)^s)(x)} \, dx \right| \\ &\leq U^{-(s+3)} \|\mu^{(s)}\|_\infty \|\mathcal{F}(w_\sigma^1(u)/u^s)\|_{L^1}. \end{aligned} \quad (37)$$

The linear error term can be split into a bias and a variance part ( $\operatorname{Var}[Z] := \mathbb{E}[|Z - \mathbb{E}[Z]|^2]$ ):

$$\begin{aligned} \mathbb{E} \left[ \left| \int_{-U}^U \mathcal{L}(u) w_\sigma^U(u) \, du \right|^2 \right] &= \left| \int_{-U}^U \frac{u(u-i) \mathbb{E}[\mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O})(u)] w_\sigma^U(u)}{\varphi_T(u-i)} \, du \right|^2 \\ &\quad + \operatorname{Var} \left[ \int_{-U}^U \frac{u(u-i) \mathcal{F}\tilde{\mathcal{O}}(u) w_\sigma^U(u)}{\varphi_T(u-i)} \, du \right] =: \mathcal{L}_b^2 + \mathcal{L}_v. \end{aligned}$$

The bias term is easily bounded by Proposition 2, using the uniform bound on  $U^{s+3} w_\sigma^U(u)/u^s$ :

$$\begin{aligned} |\mathcal{L}_b| &\leq \|\mathcal{F}(\mathcal{O}_l - \mathcal{O})\|_\infty \int_{-U}^U |\varphi_T(u-i)|^{-1} (u^4 + u^2)^{1/2} |w_\sigma^U(u)| \, du \\ &\lesssim \Delta^2 U^{-(s+3)} \int_{-U}^U e^{T \frac{\sigma^2}{2} u^2 + 2T \|\mu\|_{L^1}} |u|^{s+2} \, du. \end{aligned}$$

Making use of  $\int_0^U 2ue^{cu^2} du = \frac{e^{cU^2}-1}{c} = \mathcal{E}(cU^2)U^2$  for any  $c \geq 0$ , we estimate the last integral by

$$\int_{-U}^U e^{T\frac{\sigma^2}{2}u^2+2T\|\mu\|_{L^1}|u|^{s+2}} du \leq e^{2T\|\mu\|_{L^1}} U^{s+3} \mathcal{E}(T\frac{\sigma^2}{2}U^2)$$

and derive from  $\|\mu\|_{L^1} = \mathcal{F}\mu(0) \leq 2R$  for the bias part in the linear term

$$|\mathcal{L}_b| \lesssim \Delta^2 \mathcal{E}(T\frac{\sigma^2}{2}U^2). \quad (38)$$

For the variance part of the linear error term we use the support properties  $\text{supp}(w_\sigma^U) \in [-U, U]$  and  $\text{supp}(b_k) = [x_{k-1}, x_{k+1}]$ . Several applications of the Plancherel identity, the Cauchy-Schwarz inequality and estimate (35) then yield

$$\begin{aligned} \mathcal{L}_v &= \int_{-U}^U \int_{-U}^U \text{Cov}\left(\frac{(u-i)u\mathcal{F}\tilde{\mathcal{O}}(u)}{\varphi_T(u-i)}, \frac{(v-i)v\mathcal{F}\tilde{\mathcal{O}}(v)}{\varphi_T(v-i)}\right) w_\sigma^U(u) w_\sigma^U(v) du dv \\ &= \sum_{k=1}^N \delta_k^2 \left| \int_{-U}^U \varphi_T(u-i)^{-1} (u-i) u \mathcal{F}b_k(u) w_\sigma^U(u) du \right|^2 \\ &= 2\pi \sum_{k=1}^N \delta_k^2 \left| \int_{-\infty}^{\infty} \mathcal{F}^{-1}\left(\varphi_T(u-i)^{-1} (u-i) u w_\sigma^U(u)\right)(x) b_k(-x) dx \right|^2 \\ &\leq 2\pi \sum_{k=1}^N \delta_k^2 \int_{x_{k-1}}^{x_{k+1}} \left| \mathcal{F}^{-1}\left(\varphi_T(u-i)^{-1} (u-i) u w_\sigma^U(u)\right)(-x) \right|^2 dx \|b_k\|_{L^2}^2 \\ &\lesssim \Delta \|\delta\|_{l^\infty}^2 \int_{-\infty}^{\infty} \left| \mathcal{F}^{-1}\left(\varphi_T(u-i)^{-1} (u-i) u w_\sigma^U(u)\right)(-x) \right|^2 dx \\ &\sim \Delta \|\delta\|_{l^\infty}^2 \int_{-U}^U |\varphi_T(u-i)|^{-2} (u^4 + u^2) w_\sigma^U(u)^2 du \\ &\lesssim \Delta U^{-1} \mathcal{E}(T\sigma^2 U^2) \|\delta\|_{l^\infty}^2. \end{aligned}$$

Altogether we obtain for the linear error term

$$\mathbb{E}\left[\left|\int_{-U}^U \mathcal{L}(u) w_\sigma^U(u) du\right|^2\right] \lesssim \mathcal{E}(T\sigma^2 U^2) \left(\Delta^4 + U^{-1} \Delta \|\delta\|_{l^\infty}^2\right). \quad (39)$$

It remains to estimate the quadratic remainder term. Due to Lemma 1 and Proposition 2 we have

$$\begin{aligned} &\mathbb{E}\left[\left|\int_{-U}^U \mathcal{R}(u) w_\sigma^U(u) du\right|^2\right] \\ &\lesssim \int_{-U}^U \int_{-U}^U \mathbb{E}\left[\left|\mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O})(u) \mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O})(v)\right|^2\right] \frac{u^4 w_\sigma^U(u) v^4 w_\sigma^U(v)}{\kappa(u)^2 \kappa(v)^2} du dv. \end{aligned} \quad (40)$$

The independence of  $(\varepsilon_k)$  and the finiteness of their fourth order moments entail the inequality

$$\mathbb{E} \left[ \left| \mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O})(u) \mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O})(v) \right|^2 \right] \leq \|\mathcal{F}(\mathcal{O}_l - \mathcal{O})\|_\infty^4 + \mathbb{E}[\|\mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O}_l)(u) \mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O}_l)(v)\|^2],$$

which together with estimates (30), (35) yields that (40) is bounded in order by

$$\begin{aligned} & \int_{-U}^U \int_{-U}^U \left( \Delta^8 + \mathbb{E} \left[ \left| \sum_{k,l=1}^N \delta_k \delta_l \varepsilon_k \varepsilon_l \mathcal{F}b_k(u) \mathcal{F}b_l(v) \right|^2 \right] \right) \frac{u^4 w_\sigma^U(u) v^4 w_\sigma^U(v)}{\kappa(u)^2 \kappa(v)^2} du dv \\ & \lesssim \int_{-U}^U \int_{-U}^U \left( \Delta^8 + \sum_{k,l=1}^N \delta_k^2 \delta_l^2 |\mathcal{F}b_k(u)|^2 |\mathcal{F}b_l(v)|^2 \right) \frac{u^4 w_\sigma^U(u) v^4 w_\sigma^U(v)}{\kappa(u)^2 \kappa(v)^2} du \\ & = \left( \Delta^4 \int_{-U}^U \frac{u^4 w_\sigma^U(u)}{\kappa(u)^2} du \right)^2 + \left( \int_{-U}^U \sum_{k=1}^N \delta_k^2 |\mathcal{F}b_k(u)|^2 \frac{u^4 w_\sigma^U(u)}{\kappa(u)^2} du \right)^2 \\ & \lesssim (\Delta^8 U^4 + \Delta^4 U^4 \|\delta\|_{l^2}^2) \mathcal{E}(T \sigma_{max}^2 U^2)^2. \end{aligned}$$

Putting all estimates together and using  $U \lesssim \Delta^{-1}$  as well as  $\Delta \|\delta\|_{l^2}^2 \lesssim \|\delta\|_{l^\infty}^2$  we obtain (32) and consequently the rate for  $\sigma^2$ .

#### 6.4. Upper bound for $\gamma$ and $\lambda$

Since the claimed risk bound for  $\hat{\gamma}$  is larger than for  $\hat{\sigma}^2$ , we only need to estimate the risk of  $\hat{\gamma} + \frac{\hat{\sigma}^2}{2}$  instead of that for  $\hat{\gamma}$ . Equally, we can restrict to  $\hat{\lambda} - \frac{\hat{\sigma}^2}{2} - \hat{\gamma}$  instead of  $\hat{\lambda}$ . Then the proof follows exactly the lines of the proof for  $\hat{\sigma}^2$ , the only difference being the different norming in estimate (30) giving rise to a factor  $U$  for  $\gamma$  and a factor  $U^2$  for  $\lambda$ . It remains to note that we obtain the bounds in the compound Poisson case by setting  $\sigma = \sigma_{max} = 0$  and considering the continuous extension of the bounds for that case. For  $\hat{\gamma}$  we obtain as bias

$$\left| \int_{-U}^U \mathcal{F}\mu(u) w_\gamma^U(u) du \right| \lesssim U^{-(s+2)}. \quad (41)$$

The linear error term is estimated by

$$\mathbb{E} \left[ \left( \int_{-U}^U \mathcal{L}(u) w_\gamma^U(u) du \right)^2 \right] \lesssim \mathcal{E}(T \sigma^2 U^2) (U^2 \Delta^4 + U \Delta \|\delta\|_{l^\infty}^2) \quad (42)$$

and the remainder satisfies

$$\mathbb{E} \left[ \left| \int_{-U}^U \mathcal{R}(u) w_\gamma^U(u) du \right|^2 \right] \lesssim \mathcal{E}(T \sigma_{max}^2 U^2)^2 (\Delta^8 U^6 + \Delta^4 U^6 \|\delta\|_{l^2}^2). \quad (43)$$

For  $\hat{\lambda}$  we obtain the same asymptotic error bounds as for  $\hat{\gamma}$ , but multiplied by  $U$  when regarding the root mean square error. With the rate-optimal choice (23) of  $U$  this gives the asserted risk bounds for  $\hat{\gamma}$  and  $\hat{\lambda}$ .

### 6.5. Upper bound for $\mu$

The assertion follows as soon as the following  $\mathcal{G}_s(R, \sigma_{max})$ -uniform risk bound for general  $U$  holds:

$$\mathbb{E} \left[ \int_{-\infty}^{\infty} |\hat{\mu}(x) - \mu(x)|^2 dx \right] \lesssim U^{-2s} + \mathcal{E}(T\sigma^2 U^2) U^5 \varepsilon^2 + \mathcal{E}(2T\sigma_{max}^2 U^2) U^9 \varepsilon^4. \quad (44)$$

The bias in estimating  $\mu$  due to the cutoff at  $U$  can be estimated by

$$\int_{-\infty}^{\infty} |\mathcal{F}\mu(u)(1 - \mathbf{1}_{[-U, U]})|^2 du \leq U^{-2s} \int_{-\infty}^{\infty} |u|^{2s} |\mathcal{F}\mu(u)|^2 du = U^{-2s} \|\mu^{(s)}\|_{L^2}^2. \quad (45)$$

The variance term can be split up according to the different risk contributions. For  $u \in [-U, U]$  we obtain

$$\begin{aligned} & \mathbb{E}[|\mathcal{F}(\hat{\mu} - \mu)(u)|^2] \\ & \leq 4\mathbb{E}[|\tilde{\psi}(u) - \psi(u)|^2] + 4(u^2 + 1)^2 \mathbb{E}[|\hat{\sigma}^2 - \sigma^2|^2] \\ & \quad + 4(u^2 + 1) \mathbb{E}[|\hat{\gamma} - \gamma|^2] + 4\mathbb{E}[|\hat{\lambda} - \lambda|^2] \\ & \lesssim \mathbb{E}[|\mathcal{L}(u)|^2] + \mathbb{E}[|\mathcal{R}(u)|^2] + U^4 \mathbb{E}[|\hat{\sigma}^2 - \sigma^2|^2] + U^2 \mathbb{E}[|\hat{\gamma} - \gamma|^2] + \mathbb{E}[|\hat{\lambda} - \lambda|^2] \\ & \lesssim \mathbb{E}[|\mathcal{L}(u)|^2] + \mathbb{E}[|\mathcal{R}(u)|^2] + U^{-2(s+1)} + \mathcal{E}(T\sigma^2 U^2) U^3 \varepsilon^2 + \mathcal{E}(T\sigma_{max}^2 U^2) U^8 \varepsilon^4. \end{aligned}$$

In analogy to the previous estimates for  $\hat{\sigma}^2$  we find

$$\begin{aligned} \mathbb{E}[|\mathcal{L}(u)|^2] & \leq |\varphi_T(u - i)|^{-2} (u^4 + u^2) (\|\mathcal{F}(\mathcal{O} - \mathcal{O}_l)\|_{\infty}^2 + \text{Var}[\mathcal{F}\tilde{\mathcal{O}}(u)]) \\ & \lesssim e^{T\sigma^2 u^2} u^4 (\Delta^4 + \Delta^2 \|\delta\|_{l^2}^2). \end{aligned}$$

With a look at Lemma 1 we estimate the remainder by

$$\begin{aligned} \mathbb{E}[|\mathcal{R}(u)|^2] & \leq 16\kappa(u)^{-4} (u^4 + u^2)^2 \mathbb{E}[|\mathcal{F}(\mathcal{O}_l - \mathcal{O})(u)|^4 + |\mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O}_l)(u)|^4] \\ & \lesssim e^{2T\sigma_{max}^2 u^2} u^8 (\Delta^8 + \Delta^4 \|\delta\|_{l^2}^4). \end{aligned}$$

The Plancherel identity and these estimates yield together (44) via

$$\begin{aligned} \int_{-\infty}^{\infty} \mathbb{E}[|\hat{\mu}(x) - \mu(x)|^2] dx & \lesssim U^{-2s} + \mathcal{E}(T\sigma^2 U^2) U^5 \varepsilon^2 + \mathcal{E}(2T\sigma_{max}^2 U^2) U^9 \varepsilon^4 \\ & \quad + \mathcal{E}(T\sigma^2 U^2) U^4 \varepsilon^2 + \mathcal{E}(T\sigma_{max}^2 U^2)^2 U^9 \varepsilon^4 \\ & \sim U^{-2s} + \mathcal{E}(\sigma^2 U^2) U^5 \varepsilon^2 + \mathcal{E}(2T\sigma_{max}^2 U^2) U^9 \varepsilon^4. \end{aligned}$$

## 7. Proof of the lower bounds

We follow the usual Bayes prior technique, see e.g. Korostelev and Tsybakov (1993), and perturb a fixed Lévy triplet  $\mathcal{T}_0 = (0, \gamma_0, \nu_0)$  in the interior of  $\mathcal{G}_s(R, \sigma_{max})$  such that the perturbations remain in  $\mathcal{G}_s(R, \sigma_{max})$ .

### 7.1. Lower bound for $\mu$ in the case $\sigma = 0$

Fix a positive integer  $j$ . Let  $\psi^{(j)} \in C^\infty(\mathbb{R})$  be some function with support in  $[0, 1]$  satisfying  $\|\psi^{(j)}\|_{L^2} = 1$ ,  $\int \psi^{(j)}(x)e^{-2^{-j}x} dx = 0$  and  $\int |\mathcal{F}\psi^{(j)}(u)u^{-2}|^2 du < \infty$ . Certainly, there are infinitely many functions  $\psi^{(j)}$  fulfilling these requirements; the last property follows for instance if  $\psi$  is the second derivative of an  $L^2$ -function. Introduce the wavelet-like notation

$$\psi_{jk}(x) := 2^{j/2}\psi^{(j)}(2^j x - k), \quad j \geq 0, k = 0, \dots, 2^j - 1.$$

Consider for any  $r = (r_k) \in \{-1, +1\}^{2^j}$  and some  $\beta > 0$  the perturbed Lévy triplets  $\mathcal{T}_r = (0, \gamma_0, \mu_r)$  with

$$\mu_r(x) = \mu_0(x) + \beta 2^{-j(s+1/2)} \sum_{k=1}^{2^j} r_k \psi_{jk}(x), \quad x \in \mathbb{R}.$$

We note that due to  $\mathcal{F}\psi_{jk}(0) = 0$  and  $\int e^{-x}\psi_{jk}(x) dx = 0$  the triplet  $\mathcal{T}_r$  satisfies the martingale condition such that  $\mathcal{T}_r \in \mathcal{G}_s(R, 0)$  holds for a sufficiently small choice of the constant  $\beta > 0$ .

The Gaussian likelihood ratio of the observations under the probabilities corresponding to  $\mathcal{T}_{r'}$  and  $\mathcal{T}_r$  under the law of  $\mathcal{T}_r$  for some  $r, r'$  with  $r_k = r'_k$  for all  $k$  except one  $k_0$  is given by

$$\Lambda(r', r) = \exp\left(\int_{-\infty}^{\infty} (\mathcal{O}_{r'} - \mathcal{O}_r)(x) \varepsilon^{-1} dW(x) - \frac{1}{2} \int_{-\infty}^{\infty} |\mathcal{O}_{r'} - \mathcal{O}_r(x)|^2 \varepsilon^{-2} dx\right).$$

Hence, the Kullback-Leibler divergence (relative entropy) between the two observation models equals

$$KL(\mathcal{T}_{r'}|\mathcal{T}_r) = \frac{1}{2} \int_{-\infty}^{\infty} |(\mathcal{O}_{r'} - \mathcal{O}_r)(x)|^2 \varepsilon^{-2} dx.$$

The standard Assouad Lemma (Korostelev and Tsybakov 1993, Thm. 2.6.4) now yields the lower bound for the risk of any estimator  $\hat{\mu}$  of  $\mu$

$$\inf_{\hat{\mu}} \sup_{\mathcal{T}=(0, \gamma, \mu) \in \mathcal{G}_s(R, 0)} \mathbb{E}_{\mathcal{T}} \left[ \int |\hat{\mu}(x) - \mu(x)|^2 dx \right] \gtrsim 2^j \|\mu_r - \mu_{r'}\|_{L^2}^2 \sim 2^{-2js},$$

provided the Kullback-Leibler divergence  $KL(\mathcal{T}_{r'}|\mathcal{T}_r)$  stays uniformly bounded by a small constant. It remains to determine a minimal rate for  $2^j \rightarrow \infty$  such that this holds when the noise level tends to zero.

Arguing in the spectral domain and using the general estimate  $|e^z - 1| \leq 2|z|$ , for  $|z| \leq \delta$  and some small  $\delta > 0$ , together with  $\|\varphi_{T, r'}/\varphi_{T, r}\|_{\infty} \rightarrow 1$  for  $2^j \rightarrow \infty$ , we obtain for all sufficiently large  $j$

$$\begin{aligned} KL(\mathcal{T}_{r'}|\mathcal{T}_r) &= \frac{1}{4\pi\varepsilon^2} \int_{-\infty}^{\infty} |\mathcal{F}(\mathcal{O}_{r'} - \mathcal{O}_r)(u)|^2 du \\ &\leq \varepsilon^{-2} \int_{-\infty}^{\infty} \left| \frac{\varphi_{T, r}(u - i) - \varphi_{T, r'}(u - i)}{u(u - i)} \right|^2 du \end{aligned}$$

$$\begin{aligned}
&\leq 4\varepsilon^{-2} \int_{-\infty}^{\infty} |\varphi_{T,r}(u-i)|^2 T^2 |\mathcal{F}(\mu_r - \mu_{r'})(u)|^2 (u^4 + u^2)^{-1} du \\
&\lesssim \varepsilon^{-2} 2^{-j(2s+1)} \int_{-\infty}^{\infty} |\mathcal{F}\psi_{j_{k_0}}(u)|^2 u^{-4} du \\
&= \varepsilon^{-2} 2^{-j(2s+5)} \int_{-\infty}^{\infty} |\mathcal{F}\psi^{(j)}(v)|^2 v^{-4} dv.
\end{aligned}$$

Hence, for  $2^{j(2s+5)} \sim \varepsilon^2$  with a sufficiently large constant the Kullback-Leibler divergence remains bounded and the asymptotic lower bound for  $\mu$  follows.

## 7.2. Lower bound for $\gamma$ and $\lambda$ in the case $\sigma = 0$

Let us start with the lower bound for  $\gamma$ . We proceed as before by perturbing a triplet  $\mathcal{T}_0 = (0, \gamma_0, \mu_0)$  from the interior of  $\mathcal{G}_s(R, 0)$ , but this time we only consider one alternative  $\mathcal{T}_1 = (0, \gamma_1, \mu_1)$  and choose the perturbation in such a way that the characteristic function  $\varphi_T(u-i)$  does not change for small values of  $|u|$ . For any  $\delta > 0$  and  $U > 0$  put

$$\gamma_1 := \gamma_0 + \delta, \quad \mathcal{F}\mu_1(u) := \mathcal{F}\mu_0(u) - \delta i(u-i)e^{-u^2/U^2}, \quad u \in \mathbb{R}.$$

Then the function  $\mu_1$  is real-valued. Moreover, the martingale condition (4) is satisfied:

$$\gamma_1 + \mathcal{F}\mu_1(0) - \mathcal{F}\mu_1(i) = \gamma_0 + \delta + \mathcal{F}\mu_0(0) - \delta - \mathcal{F}\mu_0(i) + 0 = 0.$$

Because of

$$\|\mu_1^{(s)} - \mu_0^{(s)}\|_{\infty} \leq 2\pi \int_{-\infty}^{\infty} |u|^s |\mathcal{F}(\mu_1 - \mu_0)(u)| du \lesssim \delta \int_{-\infty}^{\infty} |u|^{s+1} e^{-u^2/U^2} du$$

we get  $\|\mu_1^{(s)} - \mu_0^{(s)}\|_{\infty} \lesssim \delta U^{s+2}$  and even better bounds for  $\|\mu_1^{(k)} - \mu_0^{(k)}\|_{L^2}$ ,  $k = 0, \dots, s$ . It suffices to choose  $U \sim \delta^{-1/(s+2)}$  small enough to ensure that  $\mathcal{T}_1$  still lies in our nonparametric class  $\mathcal{G}_s(R, 0)$ . The basic lower bound result (Korostelev and Tsybakov 1993, Prop. 2.2.2) then yields

$$\inf_{\hat{\gamma}} \sup_{(0, \gamma, \mu) \in \mathcal{G}_s(R, 0)} \mathbb{E}_{\gamma, \mu} [|\hat{\gamma} - \gamma|^2] \gtrsim \delta^2,$$

provided the Kullback-Leibler divergence between  $\mathcal{T}_1$  and  $\mathcal{T}_0$  remains asymptotically bounded. As in the lower bound proof for  $\mu$ , in particular using  $\mathcal{F}(\mu_1 - \mu_0)(i) = 0$ , we obtain

$$\begin{aligned}
&KL(\mathcal{T}_1|\mathcal{T}_0) \\
&\leq \frac{4}{\varepsilon^2} \int_{-\infty}^{\infty} \frac{|\varphi_{0,T}(u-i)|^2 T^2 |i(\gamma_1 - \gamma_0)(u-i) + \mathcal{F}(\mu_1 - \mu_0)(u)|^2}{(u^4 + u^2)} du \\
&\lesssim \varepsilon^{-2} \delta^2 \int_{-\infty}^{\infty} |i(u-i)(1 - e^{-u^2/U^2})|^2 (u^4 + u^2)^{-1} du
\end{aligned}$$

$$\begin{aligned}
&= \varepsilon^{-2} \delta^2 \int_{-\infty}^{\infty} (1 - e^{-v^2})^2 U^{-2} v^{-2} U \, dv \\
&\lesssim \varepsilon^{-2} \delta^2 U^{-1} \sim \varepsilon^{-2} \delta^{(2s+5)/(s+2)}.
\end{aligned}$$

Thus, the Kullback-Leibler divergence remains small for  $\delta \sim \varepsilon^{(2s+4)/(2s+5)}$  with a small constant, which gives the asymptotic lower bound for  $\gamma$ .

For the lower bound of  $\lambda$  we perturb the triplet  $\mathcal{T}_0$  leaving  $\gamma_0$  and  $\sigma_0 = 0$  fixed and putting

$$\mathcal{F}\mu_1(u) := \mathcal{F}\mu_0(u) + \delta e^{-u(u-i)/U^2}.$$

By similar estimates as for  $\gamma$ , when choosing  $U \sim \delta^{-1/(s+1)}$  with a sufficiently small constant, the perturbation  $\mathcal{T}_1$  lies in  $\mathcal{G}_s(R, 0)$  and the Kullback-Leibler divergence is asymptotically bounded by

$$KL(\mathcal{T}_1 | \mathcal{T}_0) \lesssim \varepsilon^{-2} \delta^2 U^{-3} \sim \varepsilon^{-2} \delta^{(2s+5)/(s+1)}.$$

The basic lower bound results yields the asserted lower bound for  $\lambda$ .

### 7.3. Lower bound for $\mu$ in the case $\sigma > 0$

The interesting deviation from standard proofs of lower bounds (see e.g. Butucea and Matias (2005)) for severely ill-posed problems is that we face the restriction that  $\mathcal{F}\mu$  is analytic in a strip parallel to the real line and is uniquely identifiable from its values on any open set. So, let  $\mathcal{T}_0 = (\sigma_0^2, \gamma_0, \mu_0)$  with  $\sigma_0 > 0$  be a Lévy triplet from the interior of  $\mathcal{G}_s(R, \sigma_{max})$ . Consider the perturbation  $\mathcal{T}_1 = (\sigma_0^2, \gamma_0, \mu_1)$  with

$$\mathcal{F}\mu_1(u) := \mathcal{F}\mu_0(u) + \delta m^{1/4} e^{-(T\sigma_0^2 u^2/m)^m/2} (T\sigma_0^2/m)^m u^m (u-i)^m, \quad u \in \mathbb{R}.$$

for  $m \in \mathbb{N}$ ,  $\delta > 0$ . Then we have uniformly for  $m \rightarrow \infty$  and  $\delta \rightarrow 0$

$$\|\mu_1 - \mu_0\|_{L^2}^2 = \frac{2\pi\delta^2}{\sqrt{T\sigma_0^2}} \int_0^\infty e^{-v} v^{(1+2m)/2m} (1 + m^{-1}v^{-1/m})^m \, dv \sim \delta^2.$$

Similarly, for  $k = 1, \dots, s$  we derive uniformly in  $m$  and  $\delta$

$$\begin{aligned}
\|\mu_1^{(k)} - \mu_0^{(k)}\|_{L^2} &= \sqrt{2\pi} \|u^k \mathcal{F}(\mu_1 - \mu_0)(u)\|_{L^2} \sim \delta m^{k/2}, \\
\|\mu_1^{(s)} - \mu_0^{(s)}\|_\infty &\leq \|u^s \mathcal{F}(\mu_1 - \mu_0)(u)\|_{L^1} \leq \delta m^{s/2-1/4}.
\end{aligned}$$

Therefore choosing  $\delta \sim m^{-s/2}$  with a small constant yields  $\mathcal{T}_1 \in \mathcal{G}_s(R, \sigma_{max})$  because we then also have that  $\mu_1$  is real-valued and  $\mathcal{T}_1$  satisfies the martingale condition and Assumption 1.

By the same arguments as before and by Stirling's formula to estimate the Gamma function, the Kullback-Leibler divergence between the observations under  $\mathcal{T}_0$  and under  $\mathcal{T}_1$  is asymptotically bounded by

$$4\varepsilon^{-2} \int_{-\infty}^{\infty} |\varphi_{0,T}(u-i)|^2 T^2 |\mathcal{F}(\mu_1 - \mu_0)(u)|^2 (u^4 + u^2)^{-1} \, du$$

$$\begin{aligned}
&\lesssim \varepsilon^{-2} \delta^2 \int_{-\infty}^{\infty} e^{-T\sigma_0^2 u^2} m^{1/2} e^{-(T\sigma_0^2 u^2/m)^m} (T\sigma_0^2/m)^{2m} u^{2m-2} |u-i|^{2m-2} du \\
&= \varepsilon^{-2} \delta^2 m^{-7/2} (T\sigma_0^2 m)^{-1/2} \int_0^{\infty} e^{-mv^{1/m}} e^{-v} v^{\frac{2m-1}{2m}} (1+m^{-1}v^{-1/m})^{m-1} dv \\
&\lesssim \varepsilon^{-2} \delta^2 m^{-4} \int_0^{\infty} e^{-mv^{1/m}} dv \\
&= \varepsilon^{-2} \delta^2 m^{-4} \int_0^{\infty} e^{-z} z^{m-1} m^{1-m} dz \\
&= \varepsilon^{-2} \delta^2 m^{-m-3} \Gamma(m) \lesssim \varepsilon^{-2} \delta^2 m^{-m-3} (m-1)^{m-1/2} e^{1-m} \sim \varepsilon^{-2} m^{-3-s} e^{-m}.
\end{aligned}$$

Consequently, the Kullback-Leibler divergence remains small when choosing  $m \geq 2 \log(\varepsilon^{-1})$ , but  $m \lesssim \log(\varepsilon^{-1})$ , which gives  $\delta \sim \log(\varepsilon^{-1})^{-s/2}$ . From the basic general lower bound result we therefore obtain the asymptotic lower bound for  $\mu$ .

#### 7.4. Lower bound for $\sigma^2$ , $\gamma$ and $\lambda$ in the case $\sigma > 0$

Since the proof is very similar to the preceding calculations, we only give the perturbations of the basic triplet  $\mathcal{T}_0 = (\sigma_0, \gamma_0, \mu_0)$  with  $\sigma_0 > 0$  which are least favourable. More details can be found in Belomestny and Reiß (2005).

For  $\gamma$  we leave  $\sigma_0$  fixed and use a perturbation of the form

$$\gamma_1 := \gamma_0 + \delta, \quad \mathcal{F}\mu_1(u) := \mathcal{F}\mu_0(u) - \delta i(u-i) e^{-u^{2m}/U^{2m}}.$$

For  $\lambda$  we keep  $\sigma_0$  and  $\gamma_0$  fixed and consider

$$\mathcal{F}\mu_1(u) := \mathcal{F}\mu_0(u) + \delta e^{-u^m(u-i)^m/U^{2m}}.$$

For  $\sigma^2$  we leave  $\gamma_0$  invariant and consider the perturbation

$$\sigma_1^2 := \sigma_0^2 + 2\delta, \quad \mathcal{F}\mu_1(u) := \mathcal{F}\mu_0(u) + \delta(u-i)^2 e^{-u^{2m}/U^{2m}}.$$

Each time  $m$  is chosen to be of order  $\log(\varepsilon^{-1})$  and the value of  $\delta > 0$  results from the smoothness class considered.

#### Acknowledgements

We thank Peter Tankov for intense discussions on different calibration approaches and for sharing his experience in assessing their practical performances. We are grateful for the comments and stimulating questions by two anonymous referees which have led to considerable improvements. Support by the Collaborative Research Center 649 *Economic Risk* is gratefully acknowledged.



## References

- AÏT-SAHALIA, Y., AND J. DUARTE (2003): “Nonparametric option pricing under shape restrictions,” *J. Econom.*, 116(1-2), 9–47.
- AÏT-SAHALIA, Y., AND J. JACOD (2006): “Volatility estimators for discretely sampled Lévy processes,” *Ann. Stat.*, to appear.
- BELOMESTNY, D., AND M. REISS (2005): “Optimal calibration of exponential Lévy models,” Preprint 1017, Weierstraß Institute Berlin, [www.wias-berlin.de](http://www.wias-berlin.de).
- (2006): “Spectral calibration of exponential Lévy models [2],” Discussion Paper 35, Collaborative Research Center 649 *Economic Risk*, Berlin, <http://sfb649.wiwi.hu-berlin.de>.
- BREEDEN, D., AND R. LITZENBERGER (1978): “Prices of State-Contingent Claims Implicit in Options Prices,” *J. Business*, 51(4), 621–651.
- BROWN, L. D., AND M. G. LOW (1996): “Asymptotic equivalence of nonparametric regression and white noise,” *Ann. Stat.*, 24(6), 2384–2398.
- BUTUCEA, C., AND C. MATIAS (2005): “Minimax estimation of the noise level and of the deconvolution density in a semiparametric convolution model,” *Bernoulli*, 11(2), 309–340.
- CARR, P., H. GEMAN, D. B. MADAN, AND M. YOR (2002): “The Fine Structure of Asset Returns: An Empirical Investigation,” *J. Business*, 75(2), 305–332.
- CARR, P., AND D. MADAN (1999): “Option valuation using the fast Fourier transform,” *J. Comput. Finance*, 2, 61–73.
- CONT, R., AND P. TANKOV (2004a): *Financial modelling with jump processes*, Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton.
- (2004b): “Nonparametric calibration of jump-diffusion option pricing models,” *Journal of Computational Finance*, 7(3), 1–49.
- (2005): “Retrieving Lévy processes from option prices: regularization of an ill-posed inverse problem,” *SIAM J. Num. Opt. Control*, to appear.
- CONT, R., AND E. VOLCHKOVA (2005): “Integro-differential equations for option prices in exponential Lévy models,” *Finance Stoch.*, 9(3), 299–325.
- CRÉPEY, S. (2003): “Calibration of the local volatility in a generalized Black–Scholes model using Tikhonov regularization,” *SIAM J. Math. Anal.*, 34(5), 1183–1206.
- DUFFIE, D., D. FILIPOVIC, AND W. SCHACHERMAYER (2003): “Affine processes and applications in finance,” *Ann. Appl. Probab.*, 13(3), 984–1053.
- EBERLEIN, E., U. KELLER, AND K. PRAUSE (1998): “New insights into smile, mispricing, and value at risk: the hyperbolic model,” *Journal of Business*, 71(3), 371–405.
- EMMER, S., AND C. KLÜPPELBERG (2004): “Optimal portfolios when stock prices follow an exponential Lévy process,” *Finance Stoch.*, 8(1), 17–44.
- FENGLER, M. (2005): *Semiparametric Modeling of Implied Volatility*. Springer Finance Series.
- GOLDENSHLUGER, A., A. TSYBAKOV, AND A. ZEEVI (2006): “Optimal change-point estimation from indirect observations,” *Ann. Stat.*, 34(1), 350–372.
- JACKSON, N., E. SÜLI, AND S. HOWISON (1999): “Computation of deterministic volatility surfaces,” *Journal Comp. Finance*, 2(2), 5–32.
- KALLSEN, J. (2000): “Optimal portfolios for exponential Lévy processes,” *Math. Meth. Operations Res.*, 51(3), 357–374.
- KOROSTELEV, A., AND A. TSYBAKOV (1993): *Minimax theory of image reconstruction*. Lecture Notes in Statistics (Springer). 82. New York: Springer-Verlag.
- KOU, S. (2002): “A jump diffusion model for option pricing,” *Management Science*, 48(4), 1086–1101.
- MERTON, R. (1976): “Option Pricing When Underlying Stock Returns Are Discontinuous,” *J. Financial Economics*, 3(1), 125–144.
- MORDECKI, E. (2002): “Optimal stopping and perpetual options for Lévy processes,” *Finance Stoch.*, 6(4), 473–493.