

SFB 649 Discussion Paper 2011-073

# Calibration of self-decomposable Lévy models

Mathias Trabs\*



\* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>  
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin  
Spandauer Straße 1, D-10178 Berlin



SFB 649 ECONOMIC RISK BERLIN

# Calibration of self-decomposable Lévy models

Mathias Trabs\*

Institute for Mathematics, Humboldt-Universität zu Berlin, Germany, [trabs@math.hu-berlin.de](mailto:trabs@math.hu-berlin.de).

We study the nonparametric calibration of exponential, self-decomposable Lévy models whose jump density can be characterized by the  $k$ -function, which is typically nonsmooth at zero. On the one hand the estimation of the drift, the activity measure  $\alpha := k(0+) + k(0-)$  and analog parameters for the derivatives are considered and on the other hand we estimate the  $k$ -function outside of a neighborhood of zero. Minimax convergence rates are derived, which depend on  $\alpha$ . Therefore, we construct estimators adapting to this unknown parameter. Our estimation method is based on spectral representations of the observed option prices and on regularization by cutting off high frequencies. Finally, the procedure is applied to simulations and real data.

*Keywords:* adaptation, European option, infinite activity jump process, minimax rates, non linear inverse problem, self-decomposability.

*AMS Classification (2010):* 60G51, 62G20, 91B25.

*JEL Classification:* C14, G13.

## 1. Introduction

Since Merton [20] introduced his discontinuous asset price model, stock returns were frequently described by exponentials of Lévy processes. A review of recent pricing and hedging results for these models is given by Tankov [25]. The calibration of the underlying model, that is in the case of Lévy models the estimation of the characteristic triplet  $(\sigma, \gamma, \nu)$ , from historical asset prices is mostly studied in parametric models only. Remarkable exceptions are the nonparametric penalized least squares method of Cont and Tankov [10] and the spectral calibration procedure of Belomestny and Reiß [3]. Both articles concentrate on models of finite jump activity. Our goal is to extend their results to infinite intensity models. More precisely, we study pure-jump self-decomposable Lévy processes. For instance, this class was considered in the hyperbolic model (Eberlein, Keller and Prause [11]) or the variance gamma model (Madan and Seneta [19]). Moreover, self-decomposable distributions are discussed in the financial investigation using Sato processes (Carr et al. [8], Eberlein and Madan [12]). Our results can be applied in this context, too. The nonparametric calibration of Lévy models is not only relevant for stock prices, for instance, it can be used for the Libor market as well (see Belomestny and Schoenmakers [4]). In the context of Ornstein-Uhlenbeck processes, the nonparamet-

---

\*The author thanks Markus Reiß and Jakob Söhl for providing many helpful ideas and comments. The research was supported by the Collaborative Research Center 649 “Economic Risk” of the German Research Foundation (Deutsche Forschungsgemeinschaft).

ric inference of self-decomposable Lévy processes was considered by Jongbloed, van der Meulen and van der Vaart [15].

The jump density of self-decomposable processes can be characterized by

$$\nu(x) = \frac{k(x)}{|x|}, \quad x \in \mathbb{R} \setminus \{0\}, \quad (1.1)$$

with a non-negative so-called k-function  $k : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}_+$  which increases on  $(-\infty, 0)$  and decreases on  $(0, \infty)$ . While the Blumenthal-Gettoor index, which was estimated by Belomestny [1], is zero in our model, the infinite activity can be described on a finer scale by the parameter

$$\alpha := k(0+) + k(0-).$$

Since  $k$  is typically nonsmooth at zero, we face two estimation problems: Firstly, to give a proper description of  $k$  at zero, we propose estimators for  $\alpha$  and its analogs for the derivatives  $k^{(j)}(0+) + k^{(j)}(0-)$ , with  $j \geq 1$ , as well as for the drift  $\gamma$ , which can be estimated similarly. We prove convergence rates for their mean squared error which turn out to be optimal in minimax sense up to a logarithmic factor that depends on the precise setup. Secondly, we estimate the shape of the k-function outside of a neighborhood of zero. To this end, we construct an explicit estimator of  $k$  whose mean integrated squared error on the set  $\mathbb{R} \setminus [-\tau, \tau]$ , for any  $\tau > 0$ , converges with nearly optimal rates.

Owing to bid-ask spreads and other market frictions, we observe only noisy option prices. The definition of the estimators is based on the relation between these prices and the characteristic function of the driving process established by Carr and Madan [6] and on different spectral representations of the characteristic exponent. Smoothing is done by cutting off all frequencies higher than a critical value depending on a maximal permitted parameter  $\alpha$ . The whole estimation procedure is computationally efficient and achieves good results in simulations and in real data examples.

All estimators converge with a polynomial rate, where the maximal  $\alpha$  determines the ill-posedness of the problem. Assuming sub-Gaussian error distributions, we provide an estimator with  $\alpha$ -adaptive rates. The main tool for this result is a concentration inequality for our estimator  $\hat{\alpha}$  which might be of independent interest.

This work is organized as follows: In Section 2 we describe the setting of our estimation procedure and give some details about self-decomposable processes. Subsequently, we derive the necessary representations of the characteristic exponent in Section 3. The estimation procedure is described in Section 4, where we also determine the convergence rates of our estimators. The construction of the  $\alpha$ -adaptive estimator of  $\alpha$  is contained in Section 5. In view of simulations we discuss our theoretical results and the implementation of the procedure in Section 6. Applying the proposed calibration to real data, we compare our method with the spectral calibration of Belomestny and Reiß [3]. All proofs are given in Section 7.

## 2. The model

### 2.1. Self-decomposable Lévy processes

A real valued random variable  $X$  has a *self-decomposable* law if for any  $b > 0$  there is an independent random variable  $Z_b$  such that  $X \stackrel{d}{=} bX + Z_b$ . Since each self-decomposable distribution  $\mu$  is infinitely divisible (Sato [22, Prop. 15.5]), we can define the corresponding *self-decomposable Lévy process* as the Lévy process whose law at unit time equals  $\mu$ .

Self-decomposable laws can be understood as the class of limit distributions of converging scaled sums of independent random variables [22, Thm. 15.3]. This characterization is of economical interest. If we understand the price of an asset as an aggregate of small independent influences and release from the  $\sqrt{n}$  scaling, which leads to diffusion models, we automatically end up in a self-decomposable price process. Owing to the infinite activity, the features of market prices can be reproduced even without a diffusion part (cf. Carr et al. [7]). Examples of pure-jump and self-decomposable models for option pricing are the variance gamma model, studied by Madan and Seneta [19] and Madan, Carr and Chang [18], and the hyperbolic model introduced by Eberlein, Keller and Prause [11].

Sato [22, Cor. 15.11] shows that the jump measure of a self-decomposable distribution is always absolutely continuous with respect to the Lebesgue measure and its density can be characterized through equation (1.1). Note that self-decomposability does not affect the volatility  $\sigma$  nor the drift  $\gamma$  of the Lévy process.

Assuming  $\alpha$  to be finite and  $\sigma = 0$ , the process  $X_t$  has finite variation and the characteristic function of  $X_T$  is given by the Lévy-Khintchine representation:

$$\varphi_T(u) := \mathbb{E}[e^{iuX_T}] = \exp\left(T\left(i\gamma u + \int_{-\infty}^{\infty} |e^{iux} - 1| \frac{k(x)}{|x|} dx\right)\right). \quad (2.1)$$

Motivated by a martingale argument, we will suppose the exponential moment condition  $\mathbb{E}[e^{X_t}] = 1$  for all  $t \geq 0$ , which yields

$$0 = \gamma + \int_{-\infty}^{\infty} (e^x - 1) \frac{k(x)}{|x|} dx. \quad (2.2)$$

In particular, we will impose  $\int_{-\infty}^{\infty} (e^x - 1) \frac{k(x)}{|x|} dx < \infty$ . In this case  $\varphi_T$  is defined on the strip  $\{z \in \mathbb{C} \mid \text{Im } z \in [-1, 0]\}$ .

Besides Lévy processes there is another class that is closely related to self-decomposability. Dropping the condition of stationary increments while retaining the other properties of Lévy processes, we obtain so-called additive processes. An additive process  $(Y_t)$  which is additionally self-similar, that means for all  $a > 0$  it satisfies  $(Y_{at}) \stackrel{d}{=} (a^H Y_t)$ , for some exponent  $H > 0$ , is called *Sato process*. Sato [21] showed that self-decomposable distributions can be characterized as the laws at unit time of self-similar additive processes. From the self-similarity and self-decomposability follows for  $T > 0$

$$\varphi_{Y_T}(u) = \mathbb{E}[e^{iuY_T}] = \mathbb{E}[e^{iT^H u Y_1}] = \exp\left(iT^H \gamma u + \int_{-\infty}^{\infty} (e^{iux} - 1) \frac{k(T^{-H}x)}{|x|} dx\right).$$

Since our estimation procedure only depends through equation (2.1) on the distributional structure of the underlying process, we can apply the estimators directly to Sato processes using

$$T_s = 1, \quad \gamma_s = T^H \gamma, \quad \text{and} \quad k_s(\cdot) = k(T^{-H} \bullet)$$

instead of  $T$ ,  $\gamma$  and  $k$  in the case of Lévy processes.

Going back to a Lévy process  $(X_t)$ , the parameter  $\alpha$  captures many of its properties such as the smoothness of the densities of the marginal distributions [22, Thm. 28.4] and the tail behavior of the characteristic function of a self-decomposable distribution. Since the stochastic error in our model is driven by  $|\varphi_T(u-i)|^{-1}$ , we prove the following lemma in the appendix.

**Lemma 2.1.** *Let  $(X_t)$  be a self-decomposable Lévy process with  $\sigma = 0$  and  $k$ -function  $k$  such that the martingale condition (2.2) is valid.*

- i) *If  $\|e^x k(x)\|_{L^1} < \infty$  then there exists a constant  $C_\varphi = C_\varphi(T, \|e^x k(x)\|_{L^1}, \alpha) > 0$  such that for all  $u \in \mathbb{R}$  with  $|u| \geq 1$  we obtain the bound*

$$|\varphi_T(u-i)| \geq C_\varphi |u|^{-T\alpha}.$$

- ii) *Let  $\bar{\alpha}, R > 0$  then the constant  $C_\varphi(T, R, \bar{\alpha})$  holds uniformly for all functions  $k$  with  $\alpha \leq \bar{\alpha}$  and  $\|e^x k(x)\|_{L^1} \leq R$ .*

## 2.2. Asset prices and Vanilla options

Let  $r \geq 0$  be the riskless interest rate in the market and  $S_0 > 0$  denote the initial value of the asset. In an exponential Lévy model the price process is given by

$$S_t = S_0 e^{rt + X_t},$$

where  $X_t$  is a Lévy process described by the characteristic triplet  $(\sigma, \gamma, \nu)$ . Throughout these notes, we assume  $X_t$  to be self-decomposable with  $\sigma = 0$  and  $\alpha < \infty$ . On the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  with pricing (or martingale) measure  $\mathbb{P}$  the discounted process  $(e^{-rt} S_t)$  is a martingale with respect to its natural filtration  $(\mathcal{F}_t)$ . This property is equivalent to  $\mathbb{E}[e^{X_t}] = 1$  for all  $t \geq 0$  and thus, the martingale condition (2.2) holds.

At time  $t = 0$  the risk neutral price of an European call option with underlying  $S$ , time to maturity  $T$  and strike price  $K$  is given by

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

where  $A_+ := \max\{0, A\}$ . Similarly, an European put has the price  $P(K, T) = e^{-rT} \mathbb{E}[(K - S_T)_+]$ . In terms of the negative log-forward moneyness  $x := \log(K/S_0) - rT$  the prices can be expressed as

$$\mathcal{C}(x, T) = S_0 \mathbb{E}[(e^{X_T} - e^x)_+] \quad \text{and} \quad \mathcal{P}(x, T) = S_0 \mathbb{E}[(e^x - e^{X_T})_+].$$

Carr and Madan [6] introduced the option function

$$\mathcal{O}(x) := \begin{cases} S_0^{-1}\mathcal{C}(x, T), & x \geq 0, \\ S_0^{-1}\mathcal{P}(x, T), & x < 0 \end{cases}$$

and set the Fourier transform  $\mathcal{FO}(u) := \int_{-\infty}^{\infty} e^{iux} \mathcal{O}(x) dx$  in relation to the characteristic function  $\varphi_T$  through the pricing formula

$$\mathcal{FO}(u) = \frac{1 - \varphi_T(u - i)}{u(u - i)}, \quad u \in \mathbb{R} \setminus \{0\}. \quad (2.3)$$

The properties of  $\mathcal{O}$  were studied further by Belomestny and Reiß [3, Prop. 2.1]: At any  $x \in \mathbb{R} \setminus \{0\}$  the function  $\mathcal{O}$  is twice differentiable with  $\int_{\mathbb{R}} |\mathcal{O}''(x)| dx \leq 3$  and the first derivative  $\mathcal{O}'$  has a jump of height -1 at zero. Additionally, they showed that Assumption 1 ensures an exponential decay of the option function, i.e.  $|\mathcal{O}(x)| \lesssim e^{-|x|}$  holds for  $x \in \mathbb{R}$ .

**Assumption 1.** *We assume that  $C_2 := \mathbb{E}[e^{2X_T}]$  is finite, which is equivalent to the moment condition  $\mathbb{E}[S_t^2] < \infty$ .*

Our observations are given by

$$O_j = \mathcal{O}(x_j) + \delta_j \varepsilon_j, \quad j = 1, \dots, N, \quad (2.4)$$

where the noise  $(\varepsilon_j)$  consists of independent, centered random variables with  $\mathbb{E}[\varepsilon_j^2] = 1$  and  $\sup_j \mathbb{E}[\varepsilon_j^4] < \infty$ . The noise levels  $\delta_j$  are assumed to be positive and known. In practice, the uncertainty is due to market frictions such as bid-ask spreads.

### 3. Representation of the characteristic exponent

Using (2.1) and (2.3), the shifted characteristic exponent is given by

$$\psi(u) := \frac{1}{T} \log(1 + iu(1 + iu)\mathcal{FO}(u)) = \frac{1}{T} \log(\varphi_T(u - i)) \quad (3.1)$$

$$= i\gamma u + \gamma + \int_{-\infty}^{\infty} (e^{i(u-i)x} - 1) \frac{k(x)}{|x|} dx \quad (3.2)$$

for  $u \in \mathbb{R}$ . Note that the last line equals zero for  $u = 0$  because of the martingale condition (2.2). Throughout, we choose a distinguished logarithm, that is a version of the complex logarithm such that  $\psi$  is continuous with  $\psi(0) = 0$ . On the assumption  $\int_{-\infty}^{\infty} (1 \vee e^x)k(x) dx < \infty^1$  we can apply Fubini's theorem to obtain

$$\psi(u) = i\gamma u + \gamma + \int_0^1 i(u - i)\mathcal{F}(\text{sgn}(x)k(x))((u - i)t) dt, \quad (3.3)$$

---

<sup>1</sup>We denote  $A \wedge B := \min\{A, B\}$  and  $A \vee B := \max\{A, B\}$  for  $A, B \in \mathbb{R}$ .

where the Fourier transform  $\mathcal{F}(\operatorname{sgn} \cdot k)$  is well-defined on  $\{z \in \mathbb{C} \mid \operatorname{Im} z \in [-1, 0]\}$ .

Typically, the  $k$ -function and its derivatives are not continuous at zero. Moreover, for all non-zero  $k$  the function  $x \mapsto \operatorname{sgn}(x)k(x)$  has a jump at zero. Therefore, the Fourier transform decreases very slowly. Let  $k$  be smooth on  $\mathbb{R} \setminus \{0\}$  and fulfill an integrability condition which will be important later:

**Assumption 2.** *Assume  $k \in C^s(\mathbb{R} \setminus \{0\})$  with all derivatives having a finite right- and left-hand limit at zero and  $(1 \vee e^x)k(x), \dots, (1 \vee e^x)k^{(s)} \in L^1(\mathbb{R})$ .*

Our idea is to compensate those discontinuities by adding a linear combination of the functions

$$h_j(x) := x^j e^{-x} \mathbf{1}_{[0, \infty)}(x), \quad x \in \mathbb{R}, j \in \mathbb{N} \cup \{0\}.$$

For  $j \geq 1$  it holds  $h_j \in C^{j-1}(\mathbb{R})$  and all  $h_j$  are contained in  $C^\infty(\mathbb{R} \setminus \{0\})$ . Hence, we can find  $\alpha_j, j = 0, \dots, s-2$ , such that

$$g(x) := \operatorname{sgn}(x)k(x) - \sum_{j=0}^{s-2} \alpha_j h_j(x) \in C^{s-2}(\mathbb{R}) \cap C^s(\mathbb{R} \setminus \{0\}).$$

These coefficients are given recursively by the following formula, which can be proved by straight forward calculations. We omit the details.

**Lemma 3.1.** *Grant Assumption 2. The factors  $\alpha_j, j = 0, \dots, s-2$ , satisfying  $g \in C^{s-2}(\mathbb{R}) \cap C^s(\mathbb{R} \setminus \{0\})$ , can be calculated via*

$$\alpha_j = \frac{1}{j!} (k^{(j)}(0+) + k^{(j)}(0-)) - \sum_{m=1}^j \frac{(-1)^m}{m!} \alpha_{j-m},$$

especially  $\alpha_0 = \alpha$  holds.

Hence, the Fourier transform in (3.3) can be written as  $\mathcal{F}(\operatorname{sgn} \cdot k)(z) = \mathcal{F}g(z) + \sum_{j=0}^{s-2} \alpha_j \mathcal{F}h_j(z)$ , where integration by parts yields

$$\mathcal{F}h_j(v - ti) = \int_0^\infty e^{i(v+i(1-t))x} x^j dx = \frac{j!}{(1-t-iv)^{j+1}}, \quad v \in \mathbb{R}, t \in [0, 1),$$

and  $|\mathcal{F}g(u)|$  decreases as  $|u|^{-s}$  because of the smoothness of  $g$ . From these preparations we derive a representation of  $\psi$  which allows us to estimate  $\gamma$  and  $\alpha_0, \dots, \alpha_{s-2}$ . A plug-in approach yields estimators for  $k^{(j)}(0+) + k^{(j)}(0-), j = 0, \dots, s-2$ , using Lemma 3.1. Since we only apply this representation when  $\psi$  is multiplied with weight functions having roots of degree  $s-1$  at zero, the poles that appear in (3.4) do no harm.

**Proposition 3.2.** *Let  $s \geq 2$ . On Assumption 2 there exist functions  $D : \{-1, 1\} \rightarrow \mathbb{C}$  and  $\rho : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{C}$  such that  $|u^{s-1}\rho(u)|$  is bounded in  $u$  and it holds*

$$\psi(u) = D(\operatorname{sgn}(u)) + i\gamma u - \alpha_0 \log(|u|) + \sum_{j=1}^{s-2} \frac{i^j (j-1)! \alpha_j}{u^j} + \rho(u), \quad u \neq 0, \quad (3.4)$$

From the proof in Section 7.1 we deduce the form of the mapping  $D : \{-1, 1\} \rightarrow \mathbb{C}$ :

$$D(\pm 1) = \gamma \mp i \frac{\pi}{2} - \sum_{j=1}^{s-2} (j-1)! \alpha_j + \int_{-\infty}^{\infty} g(x) \frac{e^{x/2} - 1}{x} dx \pm i \int_0^{\infty} \mathcal{F}(e^{x/2} g(x))(\pm v) dv.$$

Proposition 3.2 covers the case  $s \geq 2$ . For  $s = 1$  we conclude from (3.2) and the martingale condition (2.2)

$$\psi(u) = i\gamma u + \int_{-\infty}^{\infty} (e^{iux} - 1) e^x \frac{k(x)}{|x|} dx = i\gamma u + i \int_0^u \mathcal{F}(\operatorname{sgn}(x) e^x k(x))(v) dv, \quad (3.5)$$

where the last equation follows from Fubini's theorem on the condition  $\int_{-\infty}^{\infty} e^x k(x) dx < \infty$ , which is implied by Assumption 2. Hence,  $\psi$  is a sum of a constant from the integration, the linear drift  $i\gamma u$  and a remainder of order  $\log |u|$ , which follows from the decay of the Fourier transform as  $|u|^{-1}$  (cf. Lemma 7.1). One can even show Corollary B.6 that there exists no  $L^2$ -consistent estimator of  $\alpha$  for  $s = 1$ . Therefore, we concentrate on the case  $s \geq 2$  in the sequel.

Equation (3.5) allows another useful observation. Defining the exponentially scaled  $k$ -function

$$k_e(x) := \operatorname{sgn}(x) e^x k(x), \quad x \in \mathbb{R},$$

we obtain by differentiation

$$\psi'(u) = \frac{1}{T} \frac{(i-2u)\mathcal{FO}(u) - (u+iu^2)\mathcal{F}(x\mathcal{O}(x))(u)}{1 + (iu-u^2)\mathcal{FO}(u)} = i\gamma + i\mathcal{F}k_e(u). \quad (3.6)$$

Using this relation, we can define an estimator of  $k_e$ .

## 4. Estimation procedure

### 4.1. Definition of the estimators and weight functions

Given the observations  $\{(x_1, O_1), \dots, (x_N, O_N)\}$ , we fit a function  $\tilde{\mathcal{O}}$  to these data using linear B-splines

$$b_j(x) := \frac{x - x_{j-1}}{x_j - x_{j-1}} \mathbf{1}_{[x_{j-1}, x_j)} + \frac{x_{j+1} - x}{x_{j+1} - x_j} \mathbf{1}_{[x_j, x_{j+1}]}, \quad j = 1, \dots, N,$$

and a function  $\beta_0$  with  $\beta_0'(0+) - \beta_0'(0-) = -1$  to take care of the jump of  $\mathcal{O}'$ :

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



We choose  $\beta_0$  with support  $[x_{j_0-1}, x_{j_0}]$  where  $j_0$  satisfies  $x_{j_0-1} < 0 \leq x_{j_0}$ . Replacing  $\mathcal{O}$  with  $\tilde{\mathcal{O}}$  in the representations (3.1) and (3.6) of  $\psi$  and  $\psi'$ , respectively, allows us to define their empirical versions through

$$\begin{aligned}\tilde{\psi}(u) &:= \frac{1}{T} \log \left( v_{\kappa(u)}(1 + iu(1 + iu)\mathcal{F}\tilde{\mathcal{O}}(u)) \right), \\ \tilde{\psi}'(u) &:= \frac{1}{T} \frac{(i - 2u)\mathcal{F}\tilde{\mathcal{O}}(u) - (u + iu^2)\mathcal{F}(x\tilde{\mathcal{O}}(x))(u)}{v_{\kappa(u)}(1 + iu(1 + iu)\mathcal{F}\tilde{\mathcal{O}}(u))}, \quad u \in \mathbb{R},\end{aligned}$$

where  $\kappa$  is a positive function and we apply a trimming function given by

$$v_{\kappa}(z) : \mathbb{C} \setminus \{0\} \rightarrow \mathbb{C}, \quad z \mapsto \begin{cases} z, & |z| \geq \kappa, \\ \kappa z/|z|, & |z| < \kappa \end{cases}$$

to stabilize for large stochastic errors. A reasonable choice of  $\kappa$  will be derived below. The function  $\tilde{\psi}$  is well-defined on the interval  $[-U, U]$  on the event

$$A := \{\omega \in \Omega : 1 + iu(1 + iu)\mathcal{F}(\tilde{\mathcal{O}}(\omega, \bullet))(u) \neq 0 \forall u \in [-U, U]\} \subseteq \Omega.$$

For  $\omega \in \Omega \setminus A$  we set  $\tilde{\psi}$  arbitrarily, for instance equal to zero. The more  $\tilde{\mathcal{O}}$  concentrates around the true function  $\mathcal{O}$  the greater is the probability of  $A$ . Söhl [23] shows even that in the continuous-time Lévy model with finite jump activity the identity  $\mathbb{P}(A) = 1$  holds.

In the spirit of Belomestny and Reiß [3] we estimate the parameters  $\gamma$  and  $\alpha_j, j = 0, \dots, s-2$ , as coefficients of the different powers of  $u$  in equation (3.4). Using a spectral cut-off value  $U > 0$ , we define

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

and for  $0 \leq j \leq s-2$

$$\hat{\alpha}_j := \begin{cases} \int_{-U}^U \operatorname{Re}(\tilde{\psi}(u)) w_{\alpha_j}^U(u) \, du, & \text{if } j \text{ is even,} \\ \int_{-U}^U \operatorname{Im}(\tilde{\psi}(u)) w_{\alpha_j}^U(u) \, du, & \text{otherwise.} \end{cases}$$

Owing to (3.6), the nonparametric object  $k_e$  can be estimated by

$$\hat{k}_e(x) := \mathcal{F}^{-1} \left( \left( -\hat{\gamma} - i\tilde{\psi}'(u) \right) w_k \left( \frac{u}{U} \right) \right)(x), \quad x \in \mathbb{R}.$$

The weight functions  $w_{\gamma}^U$  and  $w_{\alpha_j}^U$  are chosen such that they filter the coefficients of interest. Moreover,  $w_k$  should decrease fast in the spatial domain and should cut off high frequencies:

**Assumption 3.** *We assume:*

- $w_\gamma^U$  fulfills for all odd  $j \in \{1, \dots, s-2\}$

$$\int_{-U}^U u w_\gamma^U(u) du = 1, \quad \int_{-U}^U u^{-j} w_\gamma^U(u) du = 0 \quad \text{and} \quad \int_0^U w_\gamma^U(\pm u) du = 0.$$

- $w_{\alpha_0}^U$  satisfies for all even  $j \in \{1, \dots, s-2\}$

$$\int_{-U}^U \log(|u|) w_{\alpha_0}^U(u) du = -1, \quad \int_{-U}^U u^{-j} w_{\alpha_0}^U(u) du = 0 \quad \text{and} \quad \int_0^U w_{\alpha_0}^U(\pm u) du = 0.$$

- For  $j = 1, \dots, s-2$  the weight functions  $w_{\alpha_j}^U$  fulfill

$$\int_{-U}^U u^{-j} w_{\alpha_j}^U(u) du = \frac{(-1)^{\lfloor j/2 \rfloor}}{(j-1)!}, \quad \int_{-U}^U u^{-l} w_{\alpha_j}^U(u) du = 0 \quad \text{and} \quad \int_0^U w_{\alpha_j}^U(\pm u) du = 0,$$

where  $1 \leq l \leq s-2$  and  $l$  is even for even  $j$  and odd otherwise. For even  $j$  we impose additionally

$$\int_{-U}^U \log(|u|) w_{\alpha_j}^U(u) du = 0.$$

- $w_k$  is contained in  $C^m(\mathbb{R})$  for some  $m \geq 2s+1$  and satisfies  $\text{supp } w_k \subseteq [-1, 1]$  as well as  $w_k \equiv 1$  on  $(-a_k, a_k)$  for some  $a_k \in (0, 1)$ .

Furthermore, we assume continuity and boundedness of the functions  $u \mapsto u^{-s+1} w_q^1(u)$  for  $q \in \{\gamma, \alpha_0, \dots, \alpha_{s-2}\}$ .

The integral conditions can be provided by rescaling: Let  $w_q^1$  satisfy Assumption 3 for  $q \in \{\gamma, \alpha_0, \dots, \alpha_{s-2}\}$  and  $U = 1$ . Since  $1 = \int_{-1}^1 u w_\gamma^1(u) du = \int_{-U}^U u U^{-2} w_\gamma^1(u/U) du$ , we can choose  $w_\gamma^U(u) := U^{-2} w_\gamma^1(\frac{u}{U})$ . Similarly, a rescaling is possible for  $w_{\alpha_0}^U$ :

$$\begin{aligned} -1 &= \int_{-1}^1 \log(|u|) w_{\alpha_0}^1(u) du = \int_{-U}^U \log(|u|) U^{-1} w_{\alpha_0}^1\left(\frac{u}{U}\right) du - \frac{\log(U)}{U} \int_{-U}^U w_{\alpha_0}^1\left(\frac{u}{U}\right) du \\ &= \int_{-U}^U \log(|u|) U^{-1} w_{\alpha_0}^1\left(\frac{u}{U}\right) du. \end{aligned}$$

Therefore, we define  $w_{\alpha_0}^U(u) := U^{-1} w_{\alpha_0}^1(\frac{u}{U})$  and analogously  $w_{\alpha_j}^U(u) := U^{j-1} w_{\alpha_j}^1(\frac{u}{U})$ . The continuity condition in Assumption 3 is set to take advantage of the decay of the remainder  $\rho$ . In connection with the rescaling it implies

$$|w_\gamma^U(u)| \lesssim U^{-s-1} |u|^{s-1} \quad \text{and} \quad |w_{\alpha_j}^U(u)| \lesssim U^{-s+j} |u|^{s-1}, \quad j = 0, \dots, s-2. \quad (4.1)$$

In the sequel we assume that the weight functions satisfy Assumption 3 and the property (4.1).

## 4.2. Convergence rates

To ensure a well-defined procedure, an exponential decay of  $\mathcal{O}$ , the identity (3.5) and to obtain a lower bound of  $|\varphi_T(u - i)|$ , we consider the class  $\mathcal{G}_0(R, \bar{\alpha})$ . Uniform convergence results for the parameters will be derived in the smoothness class  $\mathcal{G}_s(R, \bar{\alpha})$ .

**Definition 4.1.** *Let  $s \in \mathbb{N}$  and  $R, \bar{\alpha} > 0$ . We define*

- i)  $\mathcal{G}_0(R, \bar{\alpha})$  as the set of all pairs  $\mathcal{P} = (\gamma, k)$  where  $k$  is a  $k$ -function and the corresponding Lévy process  $X$  given by the triplet  $(0, \gamma, k(x)/|x|)$  satisfies Assumption 1 with  $C_2 \leq R$ , martingale condition (2.2) as well as

$$\alpha \in [0, \bar{\alpha}] \quad \text{and} \quad \|k_e\| \leq R,$$

- ii)  $\mathcal{G}_s(R, \bar{\alpha})$  as the set of all pairs  $\mathcal{P} = (\gamma, k) \in \mathcal{G}_0(R, \bar{\alpha})$  satisfying additionally Assumption 2 with

$$\begin{aligned} |k^{(l)}(0+) + k^{(l)}(0-)| &\leq R, \quad \text{for } l = 1, \dots, s-1, \\ \|(1 \vee e^x)k^{(l)}(x)\|_{L^1} &\leq R, \quad \text{for } l = 0, \dots, s. \end{aligned}$$

In the class  $\mathcal{G}_0(R, \bar{\alpha})$  Lemma 2.1 ii) provides a common lower bound of  $|\varphi_T(u - i)|$  for  $|u| \geq 1$ . Using  $\max_{x \in \mathbb{R}} \frac{1 - \cos(x)}{x} \in (0, 1]$ , we estimate roughly for  $u \in (-1, 1) \setminus \{0\}$ :

$$\begin{aligned} |\varphi_T(u - i)| &= \exp\left(T \int_{-\infty}^{\infty} (\cos(ux) - 1) \frac{e^x k(x)}{x} dx\right) \\ &\geq \exp\left(-T \int_{-\infty}^{\infty} e^{x/|u|} k\left(\frac{x}{|u|}\right) dx\right) \geq \exp(-TR). \end{aligned}$$

Hence, the choice

$$\kappa(u) := \kappa_{\bar{\alpha}}(u) := \begin{cases} \frac{1}{3} e^{-TR}, & |u| < 1, \\ \frac{1}{3} C_{\varphi}(T, R, \bar{\alpha}) |u|^{-T\bar{\alpha}}, & |u| \geq 1, \end{cases}$$

satisfies

$$\frac{1}{3} |\varphi_T(u - i)| \geq \kappa(u), \quad u \in \mathbb{R}, \quad (4.2)$$

where the factor  $1/3$  is used for technical reasons. As discussed above, we can restrict our investigation to the case  $s \geq 2$ .

Since the Lévy process is only identifiable if  $\mathcal{O}$  is known on the whole real line, we consider asymptotics of a growing number of observations with

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

Taking into account the numerical interpolation error and the stochastic error, we analyze the risk of the estimators in terms of the abstract noise level

$$\varepsilon := \Delta^{3/2} + \Delta^{1/2} \|\delta\|_{l^\infty}.$$

**Theorem 4.2.** *Let  $s \geq 2, R, \bar{\alpha} > 0$  and assume  $e^{-A} \lesssim \Delta^2$  and  $\Delta \|\delta\|_{l^2}^2 \lesssim \|\delta\|_{l^\infty}^2$ . We choose the cut-off value  $U_{\bar{\alpha}} := \varepsilon^{-2/(2s+2T\bar{\alpha}+1)}$  to obtain the uniform convergence rates*

$$\begin{aligned} \sup_{\mathcal{P}=(\gamma,k) \in \mathcal{G}_s(R,\bar{\alpha})} \mathbb{E}_{\mathcal{P}}[|\hat{\gamma} - \gamma|^2]^{1/2} &\lesssim \varepsilon^{2s/(2s+2T\bar{\alpha}+1)} \quad \text{and} \\ \sup_{\mathcal{P}=(\gamma,k) \in \mathcal{G}_s(R,\bar{\alpha})} \mathbb{E}_{\mathcal{P}}[|\hat{\alpha}_j - \alpha_j|^2]^{1/2} &\lesssim \varepsilon^{2(s-1-j)/(2s+2T\bar{\alpha}+1)}, \quad j = 0, \dots, s-2. \end{aligned}$$

As one may expect the rates for  $\alpha_j, j = 0, \dots, s-2$ , become slower as  $j$  gets closer to its maximal value because the profit from the smoothness of  $k$  decreases. Note that the cut-off for all estimators is the same.

**Remark 4.3.** The proof in Section 7.2 reveals that the condition  $\Delta \|\delta\|_{l^2}^2 \lesssim \|\delta\|_{l^\infty}^2$  is only used to estimate the remainder term. In the case  $s \geq 3$  our bound is not strict and we can replace the constraint by the weaker one

$$\Delta^r \|\delta\|_{l^2}^2 \lesssim \|\delta\|_{l^\infty}^{4-2r} \quad \text{for some } r \in \left(1, \frac{3s+2T\bar{\alpha}-1}{2s+2T\bar{\alpha}+1}\right).$$

In this setting  $\delta_j$  can be bounded away from 0 if  $A$  increases slowly enough whereas for  $r = 1$  the noise  $\delta_j$  must tend to 0 for  $x_j \rightarrow \pm\infty$ . Otherwise  $\Delta \|\delta\|_{l^2}^2$  could not be bounded because of  $\Delta N \geq \frac{2A}{N} N \rightarrow \infty$ .

For  $\tau \in (0, \frac{1}{2})$  we study the loss of the exponentially scaled k-function  $k_e$  in the norm

$$\|k_e\|_{L^2, \tau} := \left( \int_{\mathbb{R} \setminus [-\tau, \tau]} |k_e(x)|^2 dx \right)^{1/2}.$$

In contrast to  $\mathcal{G}_s(R, \bar{\alpha})$  we assume Sobolev conditions on  $k_e$  in the class  $\mathcal{H}_s(R, \bar{\alpha})$  in order to apply  $L^2$ -Fourier analysis.

**Definition 4.4.** *Let  $s \in \mathbb{N}$  and  $R, \bar{\alpha} > 0$ . We define  $\mathcal{H}_s(R, \bar{\alpha})$  as the set of all pairs  $\mathcal{P} = (\gamma, k) \in \mathcal{G}_0(R, \bar{\alpha})$  satisfying additionally  $k \in C^s(\mathcal{R} \setminus \{0\})$ ,  $\mathbb{E}_{\mathcal{P}}[|X_T e^{X_T}|] \leq R$  for corresponding Lévy process  $X$  as well as*

$$|\gamma| \leq R, \quad \text{and} \quad \|k_e^{(l)}\|_{L^2} \leq R, \quad \text{for } l = 0, \dots, s.$$

In the next theorem the conditions on  $A$  and  $\delta$  are stronger than for the upper bounds of the parameters which is due to the necessity to estimate also the derivative of  $\psi$ . However, the estimation of  $\psi'$  does not lead to a loss in the rate.

**Theorem 4.5.** *Let  $s \geq 1, R, \bar{\alpha} > 0, \tau \in (0, \frac{1}{2})$  and assume  $Ae^{-A} \lesssim \Delta^2$  and  $\Delta(\|\delta_j\|_{l^2}^2 + \Delta^2 \|(x_j \delta_j)_j\|_{l^2}^2) \lesssim \|\delta\|_{l^\infty}^2$ . We choose the cut-off value  $U_{\bar{\alpha}} := \varepsilon^{-2/(2s+2T\bar{\alpha}+5)}$ . Then we obtain for the risk of  $\hat{k}_e$  the uniform convergence rate*

$$\sup_{\mathcal{P}=(\gamma,k) \in \mathcal{H}_s(R,\bar{\alpha})} \mathbb{E}_{\mathcal{P}}[\|\hat{k}_e - k_e\|_{L^2, \tau}^2]^{1/2} \lesssim \varepsilon^{2s/(2s+2T\bar{\alpha}+5)}.$$

**Remark 4.6.** The convergence rates in the Theorems 4.2 and 4.5 are minimax optimal up to a logarithmic factor, which is shown in Appendix B and C.

## 5. Adaptation

The convergence rate of our estimation procedure depends on the bound  $\bar{\alpha}$  of the true but unknown  $\alpha \in \mathbb{R}_+$ . Therefore, we construct an  $\alpha$ -adaptive estimator. For simplicity we concentrate on the estimation of  $\alpha$  itself whereas the results can be easily extended to  $\gamma$ ,  $\alpha_j$ ,  $j = 1, \dots, s-2$ , and  $k_e$ . In this section we will require the following

**Assumption 4.** *Let  $R > 0$ ,  $s \geq 2$  and  $\alpha \in [0, \bar{\alpha}]$  for some maximal  $\bar{\alpha} > 0$ . Furthermore, we suppose  $e^{-A} \lesssim \Delta^2$  and  $\Delta \|\delta\|_{l^2}^2 \lesssim \|\delta\|_{l^\infty}^2$ .*

These conditions only recall the setting in which the convergence rates of our parameter estimators were proven. Given a consistent preestimator  $\hat{\alpha}_{pre}$  of  $\alpha$ , let  $\tilde{\alpha}_0$  be the estimator using the data-driven cut-off value and the trimming parameter

$$\tilde{U} := U_{\hat{\alpha}_{pre}} := \varepsilon^{-2/(2s+2T\hat{\alpha}_{pre}+1)} \quad \text{and} \quad (5.1)$$

$$\tilde{\kappa}(u) := \kappa_{\tilde{\alpha}_{pre}}(u) := \begin{cases} \frac{1}{2}e^{-TR}, & |u| < 1, \\ \frac{1}{2}C_{\tilde{\alpha}_{pre}}|u|^{-T\tilde{\alpha}_{pre}}, & |u| \geq 1, \end{cases} \quad (5.2)$$

respectively, with  $\tilde{\alpha}_{pre} := \hat{\alpha}_{pre} + |\log \varepsilon|^{-1}$ . If  $\hat{\alpha}_{pre}$  is sufficiently concentrated around the true value, the adaptation does not lead to losses in the rate as the following proposition shows. Note that the condition  $\tilde{\alpha}_0 \in [0, \bar{\alpha}]$  is not restrictive since any estimator  $\hat{\alpha}$  of  $\alpha \in [0, \bar{\alpha}]$  can be improved by using  $(0 \vee \hat{\alpha}) \wedge \bar{\alpha}$  instead.

**Proposition 5.1.** *On Assumption 4 let  $\hat{\alpha}_{pre}$  be a consistent estimator which is independent of the data  $O_j$ ,  $j = 1, \dots, N$ , and fulfills for  $\varepsilon \rightarrow 0$  the inequality*

$$\mathbb{P}(|\hat{\alpha}_{pre} - \alpha| \geq |\log \varepsilon|^{-1}) \leq d\varepsilon^2 \quad (5.3)$$

with a constant  $d \in (0, \infty)$ . Furthermore, we suppose  $\tilde{\alpha}_0 \in [0, \bar{\alpha}]$  almost surely. Then  $\tilde{\alpha}_0$  satisfies the asymptotic risk bound

$$\sup_{\mathcal{P} \in \mathcal{G}_s(R, \alpha)} \mathbb{E}_{\mathcal{P}, \hat{\alpha}_{pre}} [|\tilde{\alpha}_0 - \alpha|^2]^{1/2} \lesssim \varepsilon^{2(s-1)/(2s+2T\alpha+1)}$$

where the expectation is taken with respect to the common distribution  $\mathbb{P}_{\mathcal{P}, \hat{\alpha}_{pre}}$  of the observations  $O_1, \dots, O_N$  and the preestimator  $\hat{\alpha}_{pre}$ .

To use  $\hat{\alpha}_0$  on an independent sample as preestimator, we establish a concentration result for the proposed procedure. Therefor, we require  $(\varepsilon_j)$  to be uniformly sub-Gaussian (see e.g. van de Geer [27]). That means there are constants  $C_1, C_2 \in (0, \infty)$  such that the following concentration inequality holds for all  $t, N > 0$  and  $a_1, \dots, a_N \in \mathbb{R}$

$$\mathbb{P}\left(\left|\sum_{j=1}^N a_j \varepsilon_j\right| \geq t\right) \leq C_1 \exp\left(-C_2 \frac{t^2}{\sum_{j=1}^N a_j^2}\right). \quad (5.4)$$

**Proposition 5.2.** *Additionally to Assumption 4 let  $(\varepsilon_j)$  be uniformly sub-Gaussian fulfilling (5.4). Then there is a constant  $c > 0$  and for all  $\kappa > 0$  there is an  $\varepsilon_0 \sim \kappa^{(2s+2T\bar{\alpha}+1)/(2s-2)}$ , such that for all  $\varepsilon < \varepsilon_0 \wedge 1$  the estimator  $\hat{\alpha}_0$  satisfies*

$$\mathbb{P}(|\hat{\alpha}_0 - \alpha| \geq \kappa) \leq ((7N + 1)C_1 + 2) \exp(-c(\kappa^2 \wedge \kappa^{1/2})\varepsilon^{-(s-1)/(2s+2T\bar{\alpha}+1)}). \quad (5.5)$$

Concentration (5.5) is stronger than in Proposition 5.1 needed. To apply the proposed estimation procedure, let  $S_{pre}$  and  $S$  be two independent samples with noise levels  $\varepsilon_{pre}$  and  $\varepsilon$  as well as sample sizes  $N_{pre}$  and  $N$ , respectively. Using  $S_{pre}$  for the estimator  $\hat{\alpha}_{pre}$ , we construct adaptively  $\tilde{\alpha}_0$  on  $S$ . We suppose  $N_{pre}$  grows at most polynomial in  $\varepsilon_{pre}$ , that is  $N_{pre} \lesssim \varepsilon_{pre}^{-p}$  holds for some  $p > 0$ . This is fulfilled for polynomially strike distributions with a logarithmically growing domain as considered in Appendix B. To satisfy (5.3), it is sufficient if there exists a power  $q > 0$ , which can be arbitrary small, such that  $\varepsilon_{pre} \sim \varepsilon^q$  owing to the exponential inequality (5.5). Using  $\varepsilon^2 \gtrsim A_N/N \geq 1/N$ , we estimate

$$\frac{N_{pre}}{N} \lesssim \varepsilon_{pre}^{-p} \varepsilon^2 \sim \varepsilon^{2-pq} \rightarrow 0$$

for  $q < 2/p$ . Thus, relatively to all available data the necessary number of observations for the preestimator tends to zero.

## 6. Discussion and application

### 6.1. Numerical example

We apply the proposed estimation procedure to the variance gamma model (see [18]). In view of the empirical study of Madan, Carr and Chang [18] we choose the parameters  $\nu \in \{0.05, 0.1, 0.2, 0.5\}$ ,  $\sigma = 1.2$  and  $\theta = -0.15$ . The value of  $\gamma$  is then given by the martingale condition (2.2):

$$\gamma = \frac{1}{\nu} \log(1 - \theta\nu - \sigma^2\nu/2).$$

According to the different choices of  $\nu$ , we set  $\bar{\alpha} = 40$  as maximal value of  $\alpha$ .

The deterministic design of the sample  $\{x_1, \dots, x_N\}$  is distributed normally with mean zero and variance  $1/3$ . The observations  $O_j$  are computed from the characteristic function  $\varphi_T$  using the fast Fourier transform method of Carr and Madan [6]. The additive noise consists of normal centered random variables with variance  $|\delta\mathcal{O}(x_j)|^2$  for some  $\delta > 0$ .

We estimate  $q \in \{\gamma, \alpha_0, \alpha_1, \alpha_2, k_e\}$ . Hence, we need  $s \geq 4$  (see Corollary B.6). We used maturity  $T = 0.25$ , interest  $r = 0.06$ , smoothness  $s = 6$ , sample size  $N = 100$  and noise level  $\delta = 0.01$ , which generates values of  $\varepsilon$  on average 0.168. The results of our Monte Carlo simulations are summarized in Tables 1 and 2.

In order to apply the estimation procedure, we need to choose the tuning parameters. Owing to the typically unknown smoothness  $s$ , let the weight functions satisfy Assumption 3 for some large value  $s_{max}$ . The weights for the parameters can be chosen

$\alpha$	$\mathbb{E}[ \hat{\alpha}_0 - \alpha ^2]^{1/2}$	$\mathbb{E}[ \tilde{\alpha}_0 - \alpha ^2]^{1/2}$
40	20.7998	23.3589
20	5.8362	7.7724
10	1.0505	2.4534
4	0.1729	1.1158

**Table 1.** 1000 Monte Carlo simulation of the variance gamma model with  $N = 100$ ,  $\delta = 0.01$  and  $\nu \in \{0.05, 0.1, 0.2, 0.5\}$ .

	$q$	$\mathbb{E}[ \hat{q} - q ^2]^{1/2}$	$\mathbb{E}[ \tilde{q} - q ^2]^{1/2}$
$\gamma$	0.1408	0.0065	0.0126
$\alpha_0$	10.0000	1.0505	2.4534
$\alpha_1$	-94.1667	34.5066	85.9605
$\alpha_2$	4458.1250	1203.1827	2741.7031
	$\ q\ _{L^2}^{1/2}$	$\mathbb{E}[\ \hat{K} - q\ _{L^2}^2]^{1/2}$	$\mathbb{E}[\ \tilde{q} - q\ _{L^2}^2]^{1/2}$
$k_e$	0.9556	0.3289	0.3368

**Table 2.** 1000 Monte Carlo simulations of the variance gamma model with  $N = 100$ ,  $\delta = 0.01$  and  $\nu = 0.2$ .

polynomial whereas a flat-top kernel function can be used as  $w_k$ , as done by Belomestny [2]. The trimming parameter  $\kappa$  is included mainly for theoretical reasons and is not important to the implementation. The most crucial point is the choice of the cut-off value  $U$ . For  $\hat{q}$  we implement the oracle method  $U = \operatorname{argmin}_{V \geq 0} |\hat{q}(V) - q|$  and an adaptive estimator  $\tilde{q}$  based on the construction of Section 5. The sample size for the preestimator is chosen  $N_{pre} = 25$ . This adaptation to  $\alpha$  is a first step to a data-driven procedure and should be developed further.

## 6.2. Discussion

The rates show that the studied estimation problem is (mildly) ill-posed compared with classical nonparametric regression models. In order to understand the convergence rate of the estimators for  $\gamma$  and  $\alpha_j$  better, we rewrite equation (3.6) in the distributional sense, denoting the Dirac distribution at zero by  $\delta_0$ , and differentiate representation (3.4)

$$\psi'(u) = \mathcal{F}(i\gamma\delta_0 + ik_e)(u) = i\gamma - \sum_{j=0}^{s-2} i^j j! \alpha_j u^{-j-1} + \rho'(u), \quad u \in \mathbb{R} \setminus \{0\}.$$

Hence,  $\psi'$  can be seen as Fourier transform of an  $s$ -times weakly differentiable function and estimating  $\gamma$  from noisy observations of  $\psi'$  corresponds to a nonparametric regression with regularity  $s$ . Since dividing by  $u$  on the right-hand side of the above equation corresponds to taking the derivative in the spatial domain, the estimation of  $\alpha_j$  is similar to the estimation of the  $(j + 1)$ th derivative in a regression model. The convergence rate of  $k_e$  is in line with the results of Belomestny and Reiß [3] for  $\sigma = 0$ . Outside a

neighborhood of zero estimating the k-function amounts to estimating the jump density itself so that their rates equals ours in the case  $\alpha = 0$ .

For  $\hat{k}_e(x)$  with  $x$  different zero the degree of ill-posedness is given by  $T\alpha + 2$ . This can be seen analytically by observing that the noise is governed by  $u^2|\varphi_T(u - i)|^{-1}$ , which grows with rate  $T\alpha + 2$ . From a statistical point of view a higher value of  $\alpha$  leads to a more active Lévy process and hence, it is harder to distinguish the small jumps of the process from the additive noise. The influence of the time to maturity  $T$  on the convergence rates is an interesting deviation from the analysis of Belomestny and Reiß [3]. The simulation shown in Table 1 demonstrates the improvement of the estimation for small the values of  $\alpha$ .

The proposed estimator  $\hat{k}_e$  does not take into account the shape restrictions of the k-function. Therefore, it can be understood as estimator of the function  $|x|\nu(x)$  for arbitrary absolutely continuous Lévy measures. Thus, the estimation procedure can be applied to exponential Lévy models with Blumenthal-Gettoor index larger than zero, for example tempered stable processes. However, the behavior of the Lévy density at zero needs different methods in these cases and should be studied further. For instance, Belomestny [1] discusses the estimation of the fractional order for regular Lévy models of exponential type.

In the self-decomposable framework we reduce the loss of  $\hat{k}_e$  by truncating positive values on  $\mathbb{R}_-$  and negative ones on  $\mathbb{R}_+$ . The monotonicity can be generated by a rearrangement of the function. Chernozhukov, Fernández-Val and Galichon [9] show that the rearrangement reduces weakly the error for increasing target functions on compact subsets. This result carries over to our estimation problem, where  $k_e$  is decreasing and we restrict its support to a possibly large interval.

To calibrate the self-decomposable model completely, we combine the estimator  $\hat{k}_e$ , which works away from zero asymptotically optimal, and the estimators  $\hat{\alpha}_j, j \geq 0$ , which provide a proper description of the true k-function at zero. Using only  $\hat{\alpha}_0$  and  $\hat{\alpha}_1$ , this can be done as follows: Choosing some  $\tau > 0$ , we take the estimation of  $\hat{k}_e(x)$  for  $|x| \geq \tau$  and extend it continuously with linear functions on  $(-\tau, \tau)$  such that the result fits to  $\hat{\alpha}_j, j = 1, 2$ . We define the combined estimator as

$$\hat{K}(x) := \begin{cases} m_-(x + \tau) + \hat{k}_e(-\tau), & -\tau < x < 0, \\ m_+(x - \tau) + \hat{k}_e(\tau), & 0 \leq x < \tau, \\ \hat{k}_e(x), & |x| \geq \tau \end{cases}$$

where  $m_{\pm}$  are uniquely given by the conditions

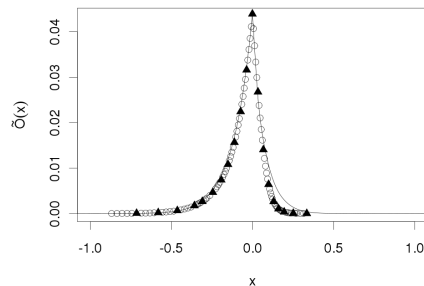
$$\hat{\alpha}_0 = \hat{K}(0+) - \hat{K}(0-) \quad \text{and} \quad 2\hat{\alpha}_0 + \hat{\alpha}_1 = \hat{K}'(0+) - \hat{K}'(0-).$$

Since  $k$  is monoton, we force  $m_{\pm} \leq 0$ , which might lead to a violation of the second equation for large stochastic errors. Table 2 contains simulation results for the estimators  $\hat{q}$  and  $\tilde{q}$ ,  $q \in \{\gamma, \alpha_0, \alpha_1, \alpha_2, k_e\}$ , corresponding to oracle and  $\alpha$ -adaptive cut-off values, respectively. The optimal combination of estimators  $\hat{\alpha}_j$  and  $\hat{k}_e$  should be developed further, for instance an exponential Taylor expansion could be used. However, taking  $\hat{\alpha}_j$



$T$	0.314	0.567
$r$	0.045	0.044
$N_{pre}$	20	21
$N$	81	85
$\tilde{\gamma}$	0.109	0.506
$\tilde{\alpha}_0$	24.850	29.846
$\tilde{\alpha}_1$	-59.595	256.049
$\tilde{\alpha}_2$	9319.844	7570.380

**Table 3.** Estimation based on ODAX from 29 May 2008.



**Figure 1.** Given ODAX data points from 29 May 2008 with  $T = 0.314$  and the option function generated from the estimated model.

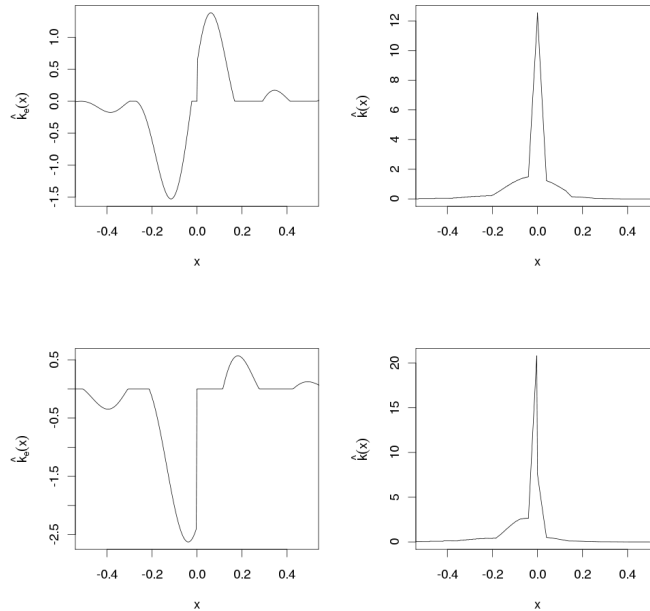
for higher  $j$  into account leads to a loss in the convergence rate and it is actually not clear how to decompose the estimators into the left and right limits of the derivatives of  $k$ . Assuming finite right- and left-hand limits of  $k$  and its derivatives at zero, one-sided kernels might estimate the  $k$ -function even in the neighborhood of zero optimally.

Even if the practitioner prefers specific parametric models that might achieve smaller errors and faster rates, the nonparametric method should be used as a goodness-of-fit test against model misspecification. This issue makes progress through study of confidence sets in the framework of Lévy processes with finite activity done by Söhl [24] and it would be interesting to derive confidence intervals for  $\alpha$ .

### 6.3. Real data example

We apply our estimation method to a data set from the Deutsche Börse database Eurex<sup>2</sup>. It consists of settlement prices of put and call options on the DAX index with three and six months to maturity from 29 May 2008. The sample sizes are 101 and 106, respectively. The interest rate is chosen such that the put-call parity holds as best as possible for all pairs of put and call options with the same strike and maturity. The subsample for the preestimator consists of every fifth strike while the main estimation is done from the remaining data points. By a rule of thumb the bid-ask spread is chosen as 1% of the option prices. Therefore, we get noise levels  $\varepsilon$  with values 0.0138 and 0.069 for the two maturities, respectively. Table 3 shows the result of the proposed method. The estimations of  $k_e$  are presented in Figure 2, which show  $\hat{k}_e$  without rearrangement as well as the estimated  $k$ -function which results from  $\hat{K}$ . In Figure 1 the calibrated model is used to generate the option function in the case of three months to maturity, where the data points used for the preestimator are marked with triangles in the figure.

<sup>2</sup>provided through the Collaborative Research Center 649 “Economic Risk”



**Figure 2.** Estimation of ODAX data from 29 May 2008 with three (*top*) and six (*bottom*) months maturity. *Left:* Estimated function  $\hat{k}_e$ . *Right:* Estimation of the k-function using  $\hat{K}$ .

Finally, we compare the outcome of our estimation procedure with the spectral calibration of Belomestny and Reiß [3], where the cut-off value is chosen by the penalized least squares criterion. The estimation results of the latter method applied to the same data set are presented in Table 4. We obtain that the higher  $\alpha$  in the selfdecomposable model corresponds to a higher  $\sigma$  in the Lévy model with finite jump activity. The parameter  $\lambda$  is even smaller for  $T = 0.567$ .

## 7. Proofs

### 7.1. Proof of Proposition 3.2

Standard Fourier analysis yields the decay of  $|\mathcal{F}g(u)|$ :

**Lemma 7.1.** *Let  $f \in C^{s-2}(\mathbb{R}) \cap C^s(\mathbb{R} \setminus \{0\})$  for  $s \geq 2$  and  $f \in C^1(\mathbb{R} \setminus \{0\})$  in case of  $s = 1$ , respectively. Furthermore, we assume finite left- and right-hand limits of  $f^{(s-1)}$  and  $f^{(s)}$  at zero and  $f^{(0)}, \dots, f^{(s)} \in L^1(\mathbb{R})$ . Then we obtain*

$$|\mathcal{F}f(u)| \lesssim |u|^{-s} \quad \text{for } |u| \rightarrow \infty.$$

$T$	0.314	0.567
$r$	0.045	0.044
$N$	101	106
$\tilde{\sigma}$	0.112	0.127
$\tilde{\gamma}$	0.160	0.100
$\tilde{\lambda}$	1.381	0.546

**Table 4.** Estimation based on ODAX from 29 May 2008.

*Epecially, on Assumption 2 there is a constant  $C_g > 0$  independent from  $t$  and  $u$  such that*

$$|\mathcal{F}g(v - it)| \leq C_g |v|^{-s} \quad \text{for } t \in [0, 1) \text{ and } |v| \geq \frac{1}{2}.$$

**Proof.** *Part 1:* Since  $f \in C^{s-2}(\mathbb{R})$  has a piecewise continuous  $(s-1)$ th derivative and all derivatives are in  $L^1(\mathbb{R})$ , standard Fourier analysis yields

$$\mathcal{F}(f^{(s-1)})(u) = (-iu)^{s-1} \mathcal{F}f(u).$$

Therefore, it is enough to show for  $f \in C^1(\mathbb{R} \setminus \{0\})$  with  $f, f' \in L^1(\mathbb{R})$  that  $|\mathcal{F}f(u)| \leq C|u|^{-1}$ ,  $|u| \geq 1$ , where  $C > 0$  does not depend on  $u$ . The integrability of  $f'$  ensures the existence of the limits of  $f$  for  $x \rightarrow \pm\infty$ . Since  $f$  itself is absolutely integrable, those limits equal 0. Integration by parts applied to the piecewise  $C^1$ -function verifies for  $u \neq 0$ :

$$\begin{aligned} |\mathcal{F}f(u)| &= \left| \int_{-\infty}^0 e^{iux} f(x) dx + \int_0^{\infty} e^{iux} f(x) dx \right| = \frac{1}{|u|} |f(0-) - f(0+) - \mathcal{F}(f')(u)| \\ &\leq \frac{1}{|u|} (|f(0-) - f(0+)| + \|f'\|_{L^1}). \end{aligned}$$

*Part 2:* From Part 1 and the Leibniz rule follow for  $t \in [0, 1)$  and  $|v| \geq \frac{1}{2}$

$$\begin{aligned} &|\mathcal{F}g(v - it)| \\ &= \frac{1}{|v|^s} \left| \mathcal{F} \left( \frac{\partial^{s-1}}{\partial x^{s-1}} e^{tx} g(x) \right) (v) \right| \left| g^{(s-1)}(0-) - g^{(s-1)}(0+) - \mathcal{F} \left( \frac{\partial^s}{\partial x^s} e^{tx} g(x) \right) (v) \right| \\ &\leq \frac{1}{|v|^s} \left( \sum_{l=0}^{s-1} \left| \mathcal{F} \left( e^{tx} g^{(l)}(x) \right) (v) \right| \right) \left( \left| g^{(s-1)}(0-) - g^{(s-1)}(0+) \right| + \sum_{l=0}^s \left| \mathcal{F} \left( e^{tx} g^{(l)}(x) \right) (v) \right| \right). \end{aligned}$$

Hence, it remains to bound  $|\mathcal{F}(e^{tx} g^{(l)}(x))(v)|$  uniformly over  $t \in [0, 1)$  and  $|v| \geq \frac{1}{2}$ , where  $l = 0, \dots, s$ . For each  $j = 0, \dots, s-2$  and  $l = 0, \dots, s$  there is a linear combination  $h_j^{(l)}(x) = \sum_{m=0}^j \beta_m^{(j,l)} h_m(x)$  with  $\beta_m^{(j,l)} \in \mathbb{R}$ ,  $m = 0, \dots, j$ . Thus, we can find  $\beta_j^{(l)} \in \mathbb{R}$ ,  $j = 0, \dots, s-2$ , such that the derivatives of  $g$  are given by

$$g^{(l)}(x) = \operatorname{sgn}(x) k^{(l)}(x) + \sum_{j=0}^{s-2} \alpha_j \beta_j^{(l)} h_j(x), \quad x \in \mathbb{R} \setminus \{0\}, l = 0, \dots, s.$$

Therefore, we obtain for all  $t \in [0, 1)$ ,  $|v| \geq \frac{1}{2}$  and  $l = 0, \dots, s$

$$\begin{aligned} \left| \mathcal{F} \left( e^{tx} g^{(l)}(x) \right) (v) \right| &= \left| \mathcal{F} \left( \operatorname{sgn}(x) e^{tx} k^{(l)}(x) \right) (v) + \sum_{j=0}^{s-2} \alpha_j \beta_j^{(l)} \mathcal{F} \left( e^{tx} h_j(x) \right) (v) \right| \\ &\leq \| (1 \vee e^x) k^{(l)}(x) \|_{L^1} + \sum_{j=0}^{s-2} \frac{j! |\alpha_j \beta_j^{(l)}|}{|1 - t - iv|^{j+1}} \\ &\leq \| (1 \vee e^x) k^{(l)}(x) \|_{L^1} + \sum_{j=0}^{s-2} 2^{(j+1)} j! |\alpha_j \beta_j^{(l)}|. \quad \square \end{aligned}$$

With this lemma at hand the representation (3.4) can be proved as follows: Owing to the symmetry  $\psi(-u) = \overline{\psi(u)}$ ,  $u \in \mathbb{R}$ , it is sufficient to consider the case  $u > 0$ . We recall representation (3.3) of  $\psi$ :

$$\psi(u) = i\gamma u + \gamma + \int_0^1 i(u-i) \mathcal{F}(\operatorname{sgn} \cdot k)((u-i)t) dt.$$

To develop this integral further we consider for  $\tau \in (0, \frac{1}{2})$

$$\begin{aligned} \xi_\tau(u) &:= \int_0^{1-\tau} i(u-i) \mathcal{F}(\operatorname{sgn} \cdot k)((u-i)t) dt \\ &= \sum_{j=0}^{s-2} \int_0^{1-\tau} i(u-i) \alpha_j \mathcal{F} h_j((u-i)t) dt + \int_0^{1-\tau} i(u-i) \mathcal{F} g((u-i)t) dt \\ &= - \sum_{j=1}^{s-2} (j-1)! \alpha_j - \alpha_0 \log(\tau - iu(1-\tau)) + \sum_{j=1}^{s-2} \frac{(j-1)! \alpha_j}{(\tau - iu(1-\tau))^j} \\ &\quad + \int_0^{1-\tau} i(u-i) \mathcal{F} g((u-i)t) dt \end{aligned}$$

To calculate the last integral we split its domain in  $[0, \frac{1}{2}]$  and  $(\frac{1}{2}, 1-\tau]$ . By assumption and choice of  $h_j$  we obtain  $|e^{i(u-i)tx} g(x)| \leq |(1 \vee e^{x/2}) g(x)| \in L^1$ , for  $0 \leq t \leq \frac{1}{2}$ , and thus, we can apply Fubini's theorem to the first part:

$$\int_0^{1/2} i(u-i) \mathcal{F} g((u-i)t) dt = \int_{-\infty}^{\infty} g(x) \int_0^{1/2} i(u-i) e^{i(u-i)tx} dt dx.$$

Since  $z \mapsto e^{izx}$  is holomorphic, Cauchy's integral theorem yields

$$\int_0^{1/2} i(u-i) e^{i(u-i)tx} dt = \int_0^{(u-i)/2} i e^{izx} dz = \int_0^{-i/2} i e^{izx} dz + \int_{-i/2}^{(u-i)/2} i e^{izx} dz.$$

Hence,

$$\int_0^{1/2} i(u-i)\mathcal{F}g((u-i)t) dt = \int_{-\infty}^{\infty} g(x) \left( \int_0^{1/2} e^{tx} dt + \int_0^{u/2} ie^{ivx+x/2} dv \right) dx.$$

Another application of Fubini's theorem to the second term shows

$$\begin{aligned} & \int_0^{1/2} i(u-i)\mathcal{F}g((u-i)t) dt \\ &= \int_{-\infty}^{\infty} g(x) \frac{e^{x/2} - 1}{x} dx + i \int_0^{\infty} \mathcal{F}(e^{x/2}g(x))(v) dv - i \int_{u/2}^{\infty} \mathcal{F}(e^{x/2}g(x))(v) dv. \end{aligned} \quad (7.1)$$

The first two summands are independent from  $u$  whereas we can use Lemma 7.1 to estimate the last integral for  $u \geq 1$ :

$$\left| \int_{u/2}^{\infty} \mathcal{F}(e^{x/2}g(x))(v) dv \right| \leq C_g \int_{u/2}^{\infty} |v|^{-s} dv = \frac{2^{s-1}C_g}{s-1} |u|^{-s+1}. \quad (7.2)$$

Also the integral over  $(\frac{1}{2}, 1-\tau]$  can be estimated using Lemma 7.1. For all  $\tau \in (0, \frac{1}{2})$  and for all  $u \geq 1$  we obtain uniformly:

$$\left| \int_{1/2}^{1-\tau} i(u-i)\mathcal{F}g((u-i)t) dt \right| \leq C_g |(u-i)u^{-s}| \int_{1/2}^1 t^{-s} dt \sim |u|^{-s+1}. \quad (7.3)$$

Thus, (7.1) yields

$$\begin{aligned} \xi_{\tau}(u) &= - \sum_{j=1}^{s-2} (j-1)! \alpha_j + \int_{-\infty}^{\infty} g(x) \frac{e^{x/2} - 1}{x} dx + i \int_0^{\infty} \mathcal{F}(e^{x/2}g(x))(v) dv \\ &\quad - \alpha_0 \log(\tau - iu(1-\tau)) + \sum_{j=1}^{s-2} \frac{(j-1)! \alpha_j}{(\tau - iu(1-\tau))^j} + \rho_{\tau}(u) \end{aligned} \quad (7.4)$$

with

$$\begin{aligned} \rho_{\tau}(u) &:= -i \int_{u/2}^{\infty} \mathcal{F}(e^{x/2}g(x))(v) dv + \int_{1/2}^{1-\tau} i(u-i)\mathcal{F}g((u-i)t) dt \\ &= -i \int_{u/2}^{\infty} \mathcal{F}(e^{x/2}g(x))(v) dv + \int_{1/2}^{1-\tau} i(u-i) \left( \mathcal{F}(\operatorname{sgn} \cdot k)((u-i)t) \right. \\ &\quad \left. - \sum_{j=0}^{s-2} \frac{j! \alpha_j}{(1 - i(u-i)t)^{j+1}} \right) dt. \end{aligned} \quad (7.5)$$

Plugging the estimates (7.2) and (7.3) into equation (7.5), we obtain  $|\rho_{\tau}(u)| \lesssim |u|^{-s+1}$  uniformly over  $\tau > 0$  and  $u \geq 1$ .

For  $u > 0$  there exists  $\rho(u) := \lim_{\tau \rightarrow 0} \rho_\tau(u)$  because  $\mathcal{F}(\text{sgn} \cdot k)$  is defined on  $\{z \in \mathbb{C} \mid \text{Im}(z) \in [-1, 0]\}$  and is continuous on its domain whereas the integral over the sum can be computed explicitly. Then the bound  $|u|^{-s+1}$  holds for  $\rho(u)$ ,  $|u| \geq 1$ , too. Also for small  $u \in (0, 1)$  the term  $|u^{s-1}\rho(u)|$  remains bounded since  $\rho$  has a pole at 0 of maximal order  $s - 2$ . Since all terms in (7.4) are continuous in  $\tau$  at 0 this equation is true for  $\tau = 0$ . Finally, we notice  $\log(-iu) = \log(|-iu|) + i \arg(-iu) = \log(|u|) - i\pi/2$  and insert (7.4) in (3.3).

## 7.2. Proof of the upper bounds

Let us recall some results of Belomestny and Reiß [3]: Because of the B-spline interpolation we obtain  $\mathcal{O}_l(x) := \mathbb{E}[\tilde{\mathcal{O}}(x)] = \sum_{j=1}^N \mathcal{O}(x_j) b_j(x) + \beta_0(x)$ ,  $x \in \mathbb{R}$ . Furthermore, the decomposition of the stochastic error  $\tilde{\psi} - \psi$  in a linearization  $\mathcal{L}$  and a remainder  $\mathcal{R}$ ,

$$\begin{aligned} \mathcal{L}(u) &:= T^{-1} \varphi_T(u - i)^{-1} (i - u) u \mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O})(u), \\ \mathcal{R}(u) &:= \tilde{\psi}(u) - \psi(u) - \mathcal{L}(u), \end{aligned}$$

$u \in \mathbb{R}$ , has the following properties:

**Proposition 7.2.** *i) Under the hypothesis  $e^{-A} \lesssim \Delta^2$  we obtain uniformly over all Lévy triplets satisfying Assumption 1*

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

*ii) If the function  $\kappa : \mathbb{R} \rightarrow \mathbb{R}_+$  satisfies (4.2) then for all  $u \in \mathbb{R}$  the remainder is bounded by*

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

*Upper bound for  $\gamma$  and  $\alpha_j$  (Theorem 4.2):*

Since Theorem 4.2 can be proven analogously to Theorem 4.2 of Belomestny and Reiß [3], we only sketch the main steps. Note that in  $\mathcal{G}_s(R, \bar{\alpha})$  we can bound uniformly the constant  $C_g$  from Lemma 7.1. Let us consider  $\gamma$  first. The definition of  $\hat{\gamma}$  and  $w_\gamma^U$ , the decomposition of  $\tilde{\psi}$  and representation (3.4) yield

$$\hat{\gamma} = \int_{-U}^U \text{Im}(\tilde{\psi}(u)) w_\gamma^U(u) \, du = \gamma + \int_{-U}^U \text{Im}(\rho(u) + \mathcal{L}(u) + \mathcal{R}(u)) w_\gamma^U(u) \, du.$$

Hence, we obtain

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

where all three summands can be estimated separately. The first one is a deterministic error term. It can be estimated using the decay of  $\rho(u)$  and the weight function property (4.1):

$$\left| \int_{-U}^U \rho(u) w_\gamma^U(u) du \right| \lesssim \int_{-U}^U U^{-(s+1)} |\rho(u) u^{s-1}| du \lesssim U^{-s}.$$

A bias-variance decomposition, with the definition  $\text{Var}(Z) := \mathbb{E}[|Z - \mathbb{E}[Z]|^2]$ , of the linear error term yields

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

Using the approximation result in Proposition 7.2, the bound of  $|\varphi_T(u-i)|^{-1}$  given by  $\kappa^{-1}$  and property (4.1), we infer the estimate of the bias term:

$$|\mathcal{L}_b| \lesssim \Delta^2 U^{-(s+1)} \int_{-U}^U |\varphi_T(u-i)|^{-1} |u|^{s+1} du \lesssim \Delta^2 U^{T\bar{\alpha}+1}.$$

For the variance part we make use of the properties of the the linear spline functions  $b_k$  as well as  $\text{supp}(w_\gamma^U) \subseteq [-U, U]$  and the independence of  $(\varepsilon_k)$ . We estimate  $(\text{Cov}(Y, Z) := \mathbb{E}[(Y - \mathbb{E}[Y])(Z - \mathbb{E}[Z])])$  as in [3]:

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

To estimate the remaining term  $\mathcal{R}$ , we use Proposition 7.2, the property (4.1) of  $w_\gamma^U$  and the choice of  $\kappa$ . In addition the independence of  $(\varepsilon_k)$  and the uniform bound of their fourth moments comes into play.

$$\begin{aligned} &\mathbb{E} \left[ \left| \int_{-U}^U \mathcal{R}(u) w_\gamma^U(u) du \right|^2 \right] \\ &\lesssim \int_{-U}^U \int_{-U}^U \left( \|\mathcal{F}(\mathcal{O}_l - \mathcal{O})\|_\infty^4 + \mathbb{E} \left[ \left| \mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O}_l)(u) \mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O}_l)(v) \right|^2 \right] \right) \frac{u^4 w_\gamma^U(u) v^4 w_\gamma^U(v)}{\kappa(u)^2 \kappa(v)^2} du dv \\ &\lesssim \left( \Delta^4 \int_{-U}^U \frac{u^4 w_\gamma^U(u)}{\kappa(u)^2} du \right)^2 + \left( \int_{-U}^U \sum_{k=1}^N \delta_k^2 |\mathcal{F}b_k(u)|^2 \frac{u^4 w_\gamma^U(u)}{\kappa(u)^2} du \right)^2 \\ &\lesssim \left( \Delta^4 U^{-(s+1)} \int_{-U}^U \kappa(u)^{-2} |u|^{s+3} du \right)^2 + \left( \Delta^2 \|\delta\|_{l^2}^2 U^{-(s+1)} \int_{-U}^U \kappa(u)^{-2} |u|^{s+3} du \right)^2 \\ &\lesssim U^{4T\bar{\alpha}+6} (\Delta^8 + \Delta^4 \|\delta\|_{l^2}^4). \end{aligned}$$

Therefore, the total risk of  $\hat{\gamma}$  is of order

$$\mathbb{E}[|\hat{\gamma} - \gamma|^2] \lesssim U^{-2s} + U^{2T\bar{\alpha}+1}(\Delta^4 U + \Delta \|\delta\|_{l^\infty}^2) + U^{4T\bar{\alpha}+6}(\Delta^8 + \Delta^4 \|\delta\|_{l^2}^4)$$

uniformly over  $\mathcal{G}_s(R, \bar{\alpha})$ . Since the explicit choice of  $U = U_{\bar{\alpha}} = \varepsilon^{-2/(2s+2T\bar{\alpha}+1)}$  fulfills  $U \lesssim \Delta^{-1}$  and  $\Delta \|\delta\|_{l^2}^2 \lesssim \|\delta\|_{l^\infty}^2$  holds by assumption, this bound simplifies to

$$\mathbb{E}[|\hat{\gamma} - \gamma|^2] \lesssim U^{-2s} + U^{2T\bar{\alpha}+1}\varepsilon^2 + U^{4T\bar{\alpha}+6}\varepsilon^4.$$

Here  $U_{\bar{\alpha}}$  balances the trade-off between the first and the second term whereby the third summand is asymptotically negligible. We obtain the claimed rate.

For  $\alpha_j, j = 0, \dots, s-2$ , the only difference to the analysis for  $\hat{\gamma}$  is the rescaling factor of  $w_{\alpha_j}^U$  in (4.1). Since its square appears in front of every summand, we verify

$$\begin{aligned} \mathbb{E}[|\hat{\alpha}_j - \alpha_j|^2] &\lesssim U^{-2(s-1-j)} + U^{2T\bar{\alpha}+2j+3}(\Delta^4 U + \Delta \|\delta\|_{l^\infty}^2) \\ &\quad + U^{4T\bar{\alpha}+2j+8}(\Delta^8 + \Delta^4 \|\delta\|_{l^2}^4) \\ &\lesssim U^{-2(s-1-j)} + U^{2T\bar{\alpha}+2j+3}\varepsilon^2 + U^{4T\bar{\alpha}+2j+8}\varepsilon^4. \end{aligned}$$

The explicit choice of  $U = U_{\bar{\alpha}}$  implies the result.

*Upper bound for  $k_e$  (Theorem 4.5):*

Similarly to the uniform bound of the bias of  $\mathcal{F}\tilde{\mathcal{O}}$  in Proposition 7.2, we prove the following lemma.

**Lemma 7.3.** *If  $Ae^{-A} \lesssim \Delta^2$  holds, we obtain uniformly over all Lévy triplets satisfying Assumption 1 and  $\mathbb{E}[|X_T e^{X_T}|] \lesssim 1$*

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

**Proof.** We follow the lines of the proof of Proposition 6.1 in [3] with the slightly different estimation:

$$\begin{aligned} &\int_{x_1}^{x_N} |x(\mathcal{O}_l - \mathcal{O})(x)| dx \leq \sum_{j=2}^N (|x_{j-1}| \vee |x_j|) \int_{x_{j-1}}^{x_j} |\mathcal{O}_l(x) - \mathcal{O}(x)| dx \\ &\leq \sum_{j \in \{2, \dots, N\} \setminus \{j_0\}} \int_{x_{j-1}}^{x_j} \int_{x_{j-1}}^x \int_{x_{j-1}}^{x_j} (|x_{j-1}| + \Delta) |\mathcal{O}''(z)| dz dy dx + C_0 (|x_{j_0-1}| \vee |x_{j_0}|) \Delta^2 \\ &\leq \|x\mathcal{O}''(x)\|_{L^1} \Delta^2 + \|\mathcal{O}''\|_{L^1} \Delta^3 + 2C_0 \Delta^3. \end{aligned}$$

Since the extrapolation errors can be bounded by  $4C_2\Delta(A + \Delta)e^{-A-\Delta}$ , we obtain

$$\begin{aligned} &\int_{-\infty}^{\infty} |\mathbb{E}[x(\tilde{\mathcal{O}} - \mathcal{O})(x)]| dx \\ &\leq 2C_2(Ae^{-A} + e^{-A}) + \|x\mathcal{O}''(x)\|_{L^1} \Delta^2 + (\|\mathcal{O}''\|_{L^1} + 2C_0)\Delta^3 + 4C_2\Delta(A + \Delta)e^{-A-\Delta} \end{aligned}$$



It remains to bound  $\|x\mathcal{O}''(x)\|_{L^1}$ . Recall from [3, Prop. 2.1] that  $\mathcal{O}''(x) = e^x(\mathbb{P}(X_T < x) + f_T(x) - \mathbf{1}_{\{x>0\}})$ ,  $x \in \mathbb{R}$ . Integration by parts yields

$$\begin{aligned} \int_0^\infty |x\mathcal{O}''(x)| dx &= \int_0^\infty xe^x |\mathbb{P}(X_T < x) + f_T(x) - 1| dx \\ &\leq \int_0^\infty xe^x (1 - \mathbb{P}(X_T < x)) dx + \mathbb{E}[|X_T|e^{X_T} \mathbf{1}_{\{X_T>0\}}] \\ &= 1 - \mathbb{P}(X_T < 0) + \int_0^\infty (x-1)e^x f_T(x) dx + \mathbb{E}[|X_T|e^{X_T} \mathbf{1}_{\{X_T>0\}}] \\ &= \mathbb{P}(X_T \geq 0) + \mathbb{E}[(2|X_T| + 1)e^{X_T} \mathbf{1}_{\{X_T>0\}}]. \end{aligned}$$

We conclude analogously

$$\int_{-\infty}^0 |x\mathcal{O}''(x)| dx = \mathbb{P}(X_T < 0) + \mathbb{E}[(2|X_T| + 1)e^{X_T} \mathbf{1}_{\{X_T<0\}}].$$

Therefore, it holds  $\|x\mathcal{O}''(x)\|_{L^1} \leq 2 + 2\mathbb{E}[|X_T|e^{X_T}]$ , which is bounded by assumption.  $\square$

As we will see, the estimation of  $\gamma$  in the definition of  $\hat{k}_e$  is asymptotically negligible. We thus set  $\hat{\gamma} \equiv 0$  in this section. To show Theorem 4.5 we define the function  $W_k := \mathcal{F}^{-1}w_k$  which can be understood as a kernel with bandwidth  $U^{-1}$ . By the properties of the weight  $w_k$  it satisfies for  $l = 1, \dots, m-2$ :

$$\begin{aligned} w_k\left(\frac{u}{U}\right) &= U\mathcal{F}(W_k(Ux))(u), & \int_{\mathbb{R}} W_k(x) dx &= w_k(0) = 1, \\ \int_{\mathbb{R}} x^l W_k(x) dx &= (-i)^l w_k^{(l)}(0) = 0, & \int_{\mathbb{R}} |x|^l |W_k(x)| dx &< \infty. \end{aligned}$$

We split the risk into a deterministic error, an error caused by  $\gamma$  and a stochastic error,

$$\begin{aligned} &\mathbb{E}_{\mathcal{P}}[\|\hat{k}_e - k_e\|_{L^2, \tau}^2] \\ &= \mathbb{E}_{\mathcal{P}}[\|\mathcal{F}^{-1}\left(-i\tilde{\psi}'(u)w_k\left(\frac{u}{U}\right)\right) - k_e\|_{L^2, \tau}^2] \\ &\leq \mathbb{E}_{\mathcal{P}}\left[\int_{\mathbb{R} \setminus [-\tau, \tau]} 3\left|\mathcal{F}^{-1}\left((-\gamma - i\psi'(u))w_k\left(\frac{u}{U}\right)\right)(x) - k_e(x)\right|^2\right. \\ &\quad \left.+ 3\left|\mathcal{F}^{-1}\left(\gamma w_k\left(\frac{u}{U}\right)\right)(x)\right|^2 + 3\left|\mathcal{F}^{-1}\left((-i\tilde{\psi}'(u) + i\psi'(u))w_k\left(\frac{u}{U}\right)\right)(x)\right|^2 dx\right] \\ &= 3\int_{\mathbb{R} \setminus [-\tau, \tau]} \left|\mathcal{F}^{-1}\left(\mathcal{F}k_e(u)w_k\left(\frac{u}{U}\right)\right)(x) - k_e(x)\right|^2 dx + 3|\gamma|^2 \int_{\mathbb{R} \setminus [-\tau, \tau]} |UW_k(Ux)|^2 dx \\ &\quad + 3\mathbb{E}\left[\int_{\mathbb{R} \setminus [-\tau, \tau]} \left|\mathcal{F}^{-1}\left((\tilde{\psi}'(u) - \psi'(u))w_k\left(\frac{u}{U}\right)\right)(x)\right|^2 dx\right] \\ &=: D + G + S. \end{aligned}$$

Using  $w_k \in C^m(\mathbb{R})$ , we infer  $|W_k(x)| \lesssim |x|^{-m}$  for  $x \rightarrow \infty$  and thus, the addend  $G$  can be bounded by

$$G \lesssim 6U^2|\gamma|^2 \int_{\tau}^{\infty} |Ux|^{-2m} dx \lesssim U^{-2m+2}.$$

The deterministic term  $D$  can be estimated in the spatial domain, where we use the local smoothness of  $k_e$ . For pointwise convergence rates this was done by Belomestny [2]. We decompose

$$\begin{aligned} D &= 3 \int_{\mathbb{R} \setminus [-\tau, \tau]} \left| U(k_e * W_k(U\bullet))(x) - k_e(x) \right|^2 dx \\ &\leq 6 \int_{\mathbb{R} \setminus [-\tau, \tau]} \left| \int_{|y| > U\tau} (k_e(x - y/U) - k_e(x)) W_k(y) dy \right|^2 dx \\ &\quad + 6 \int_{\mathbb{R} \setminus [-\tau, \tau]} \left| \int_{|y| \leq U\tau} (k_e(x - y/U) - k_e(x)) W_k(y) dy \right|^2 dx =: 6(D_1 + D_2). \end{aligned}$$

An application of the Cauchy-Schwarz inequality, of the estimate  $\int_{|y| > U\tau} |W_k(y)| dy \leq (U\tau)^{-m+2} \int_{\mathbb{R}} |y|^{m-2} |W_k(y)| dy \lesssim U^{-m+2}$  and of Fubini's theorem yield

$$\begin{aligned} D_1 &\leq \int_{\mathbb{R} \setminus [-\tau, \tau]} \int_{|y| > U\tau} |W_k(y)| dy \int_{|y| > U\tau} |k_e(x - y/U) - k_e(x)|^2 |W_k(y)| dy dx \\ &\lesssim (U)^{-m+2} \int_{\mathbb{R} \setminus [-\tau, \tau]} \int_{|y| > U\tau} (|k_e(x - y/U)|^2 + |k_e(x)|^2) |W_k(y)| dy dx \\ &\lesssim (U)^{-m+2} \int_{|y| > U\tau} |W_k(y)| \int_{\mathbb{R} \setminus [-\tau, \tau]} |k_e(x - y/U)|^2 + |k_e(x)|^2 dx dy \\ &\lesssim (U)^{-2m+4} \|k_e\|_{L^2}^2. \end{aligned}$$

Using a Taylor expansion, we split  $D_2$  in a polynomial part and a remainder:

$$\begin{aligned} D_2 &\leq 2 \int_{\mathbb{R} \setminus [-\tau, \tau]} \left| \int_{|y| \leq U\tau} \left( \sum_{j=0}^{s-1} \frac{k_e^{(j)}(x)}{j! U^j} (-y)^j \right) W_k(y) dy \right|^2 dx \\ &\quad + 2 \int_{\mathbb{R} \setminus [-\tau, \tau]} \left| \int_{|y| \leq U\tau} \int_x^{x-y/U} \frac{k_e^{(s)}(z)(x-y/U-z)^{s-1}}{(s-1)!} dz W_k(y) dy \right|^2 dx \\ &=: 2D_{2P} + 2D_{2R}. \end{aligned}$$

We estimate

$$\begin{aligned} D_{2P} &\leq s(U\tau)^{-2m+4} \sum_{j=0}^{s-1} \frac{\tau^{2j}}{(j!)^2} \int_{\mathbb{R} \setminus [-\tau, \tau]} |k_e^{(j)}(x)|^2 dx \left( \int_{\mathbb{R}} |y|^{m-2} |W_k(y)| dy \right)^2 \\ &\lesssim U^{-2m+4} \sum_{j=0}^{s-1} \|k_e^{(j)}\|_{L^2}^2. \end{aligned}$$

With twofold usage of Cauchy-Schwarz and with Fubini's theorem we obtain

$$\begin{aligned}
D_{2R} &= \int_{|x|>\tau} \left| \int_{|y|\leq U\tau} \int_0^{y/U} \frac{k_e^{(s)}(x-z)(z-\frac{y}{U})^{s-1}}{(s-1)!} dz W_k(y) dy \right|^2 dx \\
&\leq \int_{|x|>\tau} \left( \int_{|y|\leq U\tau} \left( \int_0^{|y/U|} |k_e^{(s)}(x-\operatorname{sgn}(y)z)|^2 dz \right)^{1/2} \right. \\
&\quad \cdot \left. \left( \int_0^{|y/U|} \frac{(z-\frac{y}{U})^{2s-2}}{((s-1)!)^2} dz \right)^{1/2} |W_k(y)| dy \right)^2 dx \\
&\leq \int_{|x|>\tau} \int_{|y|\leq U\tau} \int_0^{|y/U|} |k_e^{(s)}(x-\operatorname{sgn}(y)z)|^2 dz |W_k(y)| dy \\
&\quad \cdot \int_{|y|\leq U\tau} \frac{|y/U|^{2s-1}}{(2s-1)((s-1)!)^2} |W_k(y)| dy dx \\
&\lesssim U^{-(2s-1)} \int_{|y|\leq U\tau} \int_0^{|y/U|} \int_{|x|>\tau} |k_e^{(s)}(x-\operatorname{sgn}(y)z)|^2 dx dz |W_k(y)| dy \\
&\leq U^{-(2s-1)} \|k_e^{(s)}\|_{L^2}^2 \int_{|y|\leq U\tau} |y/U| |W_k(y)| dy \\
&\lesssim U^{-2s}.
\end{aligned}$$

Therefore, we have  $D + G \lesssim U^{-2s}$ .

To estimate the stochastic error  $S$ , we bound the term  $|\tilde{\psi}'(u) - \psi'(u)|$ . Let us introduce the notation

$$\begin{aligned}
\tilde{\varphi}_T(u-i) &:= v_{\kappa(u)}(1 + (iu - u^2)\mathcal{F}\tilde{\mathcal{O}}(u)), \\
\tilde{\varphi}'_T(u-i) &:= (i - 2u)\mathcal{F}\tilde{\mathcal{O}}(u) - (u + iu^2)\mathcal{F}(x\tilde{\mathcal{O}}(x))(u), \quad u \in \mathbb{R}.
\end{aligned}$$

For all  $u \in \mathbb{R}$  where  $|\tilde{\varphi}_T(u-i)| > \kappa(u)$  we obtain  $\tilde{\varphi}_T(u-i) = 1 + (iu - u^2)\mathcal{F}\tilde{\mathcal{O}}(u)$ . For  $|\tilde{\varphi}_T(u-i)| = \kappa(u)$  the estimate  $|\tilde{\varphi}_T(u-i) - \varphi_T(u-i)| \geq 2\kappa(u)$  follows from (4.2). This yields

$$\begin{aligned}
|\tilde{\varphi}_T(u-i) - \varphi_T(u-i)| &\leq |1 + (iu - u^2)\mathcal{F}\tilde{\mathcal{O}}(u) - \varphi_T(u-i)| + \kappa(u) \\
&\leq |1 + (iu - u^2)\mathcal{F}\tilde{\mathcal{O}}(u) - \varphi_T(u-i)| + \frac{1}{2}|\tilde{\varphi}_T(u-i) - \varphi_T(u-i)|.
\end{aligned}$$

Therefore,  $|\tilde{\varphi}_T(u-i) - \varphi_T(u-i)| \leq 2|1 + (iu - u^2)\mathcal{F}\tilde{\mathcal{O}}(u) - \varphi_T(u-i)|$  holds for all  $u \in \mathbb{R}$ . We obtain a similar decomposition as Kappus and Reiß [16],

$$\begin{aligned}
|\tilde{\psi}'(u) - \psi'(u)| &= \frac{1}{T} \left| \frac{\tilde{\varphi}'_T(u-i)}{\tilde{\varphi}_T(u-i)} - \frac{\varphi'_T(u-i)}{\varphi_T(u-i)} \right| \\
&\leq \frac{1}{T|\tilde{\varphi}_T(u-i)|} \left( |\tilde{\varphi}'_T(u-i) - \varphi'_T(u-i)| + T|\psi'(u)||\varphi_T(u-i) - \tilde{\varphi}_T(u-i)| \right)
\end{aligned}$$

$$\leq \frac{1}{2T\kappa(u)} \left( ((1+4u^2)^{1/2} + 2T|\psi'(u)|(u^2+u^4)^{1/2}) |\mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O})(u)| \right. \\ \left. + (u^2+u^4)^{1/2} |\mathcal{F}(x(\tilde{\mathcal{O}} - \mathcal{O})(x))(u)| \right).$$

Since  $|\psi'(u)| \leq |\gamma| + \|k_e\|_{L^1} \leq 2R$ , we have

$$|\tilde{\psi}'(u) - \psi'(u)| \lesssim \frac{1}{\kappa(u)} \left( (1+u^2) |\mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O})(u)| + (u^2+u^4)^{1/2} |\mathcal{F}(x(\tilde{\mathcal{O}} - \mathcal{O})(x))(u)| \right).$$

It follows with Plancherel's equality

$$S \leq 3\mathbb{E} \left[ \|\mathcal{F}^{-1}((\tilde{\psi}'(u) - \psi'(u))w_k(u/U))\|_{L^2}^2 \right] = \frac{3}{2\pi} \int_{\mathbb{R}} \mathbb{E}[|\tilde{\psi}'(u) - \psi'(u)|^2] |w_k(u/U)|^2 du \\ \lesssim \int_{-U}^U \frac{u^4}{|\kappa(u)|^2} \left( \mathbb{E}[|\mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O})(u)|^2] + \mathbb{E}[|\mathcal{F}(x(\tilde{\mathcal{O}} - \mathcal{O})(x))(u)|^2] \right) |w_k(u/U)|^2 du \\ =: S_1 + S_2.$$

Both terms can be estimated similarly. Thus, we only write it down for  $S_2$ , where stronger conditions are needed. Lemma 7.3 and  $\|\mathcal{F}(xb_j(x))\|_{\infty} \leq 2\Delta(x_j + \Delta)$ ,  $j = 1, \dots, N$ , yield

$$S_2 \leq \int_{-U}^U \frac{u^4}{|\kappa(u)|^2} \left( \|x(O_l - O)(x)\|_{\infty}^2 + \text{Var}(\mathcal{F}(x\tilde{\mathcal{O}}(x))(u)) \right) |w_k(u/U)|^2 du \\ \lesssim \int_{-U}^U |u|^{2T\bar{\alpha}+4} \left( \Delta^4 + \sum_{j=1}^N \delta_j^2 |\mathcal{F}(xb_j(x))(u)|^2 \right) du \\ \lesssim (\Delta^4 + \Delta^2 \|(x_j \delta_j)\|_{l^2}^2 + \Delta^4 \|\delta_j\|_{l^2}^2) U^{2T\bar{\alpha}+5} \lesssim \varepsilon^2 U^{2T\bar{\alpha}+5}.$$

Therefore, we have shown  $\mathbb{E}[\|\hat{k}_e - k_e\|_{L^2, \tau}^2] \lesssim U^{-2s} + \varepsilon^2 U^{2T\bar{\alpha}+5}$ . The claim follows from the asymptotic optimal choice  $U = U_{\bar{\alpha}} = \varepsilon^{-2/(2s+2T\bar{\alpha}+5)}$ .

### 7.3. Proof of Proposition 5.1

*Step 1:* We consider deterministic approximate of  $\alpha$ . Let  $(a_\varepsilon)_{\varepsilon>0}$  be such that there is a constant  $C > 0$  with  $|a_\varepsilon - \alpha| \leq C|\log \varepsilon|^{-1}$ . Let the estimator  $\hat{\alpha}_0$  use the cut-off value  $U_\varepsilon := \tilde{U}_{a_\varepsilon}$  and the trimming parameter  $\kappa_\varepsilon := \tilde{\kappa}_{\bar{a}_\varepsilon}$ , with  $\bar{a}_\varepsilon := a_\varepsilon + C|\log \varepsilon|^{-1}$ , as defined in (5.1) and (5.2). Then we can show the asymptotic risk bound

$$\sup_{\mathcal{P} \in \mathcal{G}_s(R, \alpha)} \mathbb{E}_{\mathcal{P}}[|\hat{\alpha}_0 - \alpha|^2]^{1/2} \lesssim \varepsilon^{2(s-1)/(2s+2T\alpha+1)}$$

as follows: By construction holds  $\alpha \leq \bar{a}_\varepsilon$ . Hence,  $\kappa_\varepsilon$  fulfills condition (4.2) for each pair  $\mathcal{P} \in \mathcal{G}_s(R, \alpha)$ . Therefore, we deduce from Theorem 4.2:

$$\mathbb{E}_{\mathcal{P}}[|\hat{\alpha}_0 - \alpha|^2] \lesssim U_\varepsilon^{-2(s-1)} + U_\varepsilon^{2T\alpha+3} \varepsilon^2 + U_\varepsilon^{4T\bar{a}_\varepsilon+8} \varepsilon^4 \\ = \varepsilon^{4(s-1)/(2s+2T\alpha+1)} \left( 1 + \varepsilon^{4T(a_\varepsilon - \alpha)/(2s+2T\alpha+1)} + \varepsilon^{(4s-8+8T(a_\varepsilon - \bar{a}_\varepsilon))/(2s+2T\alpha+1)} \right). \quad (7.6)$$

The first factor has the claimed order, which follows from

$$\begin{aligned}
& (\alpha - a_\varepsilon) \log \varepsilon \leq C \\
\Rightarrow & (2s + 2T\alpha + 1) \log \varepsilon \leq (2s + 2Ta_\varepsilon + 1) \log \varepsilon + 2TC \\
\Rightarrow & \frac{4(s-1)}{2s+2Ta_\varepsilon+1} \log \varepsilon \leq \frac{4(s-1)}{2s+2T\alpha+1} \log \varepsilon + \frac{8(s-1)T}{(2s+1)^2} C \\
\Rightarrow & \varepsilon^{4(s-1)/(2s+2Ta_\varepsilon+1)} \lesssim \varepsilon^{4(s-1)/(2s+2T\alpha+1)}.
\end{aligned}$$

Thus, the claim follows once we have bound the sum in the bracket of equation (7.6). For the second summand this is implied by

$$\left| \frac{4T(a_\varepsilon - \alpha)}{2s + 2Ta_\varepsilon + 1} \log \varepsilon \right| \leq \frac{4T|(a_\varepsilon - \alpha) \log \varepsilon|}{2s + 1} \leq \frac{4TC}{2s + 1}.$$

To estimate the third term, we obtain from  $s \geq 2$  and  $\varepsilon < 1$

$$\frac{4s - 8 + 8T(a_\varepsilon - \bar{a}_\varepsilon)}{2s + 2Ta_\varepsilon + 1} \log \varepsilon \leq \frac{-8TC|\log \varepsilon|^{-1}}{2s + 1} \log \varepsilon \leq \frac{8TC}{2s + 1}.$$

*Step 2:* Let  $\mathcal{P} \in \mathcal{G}_s(R, \alpha)$ . Note that  $\kappa_\varepsilon$  satisfies the condition (4.2) on the set  $\{|\hat{\alpha}_{pre} - \alpha| < |\log \varepsilon|^{-1}\}$ . Using the independence of  $\hat{\alpha}_{pre}$  and  $O_j$ , the almost sure bound  $\tilde{\alpha}_0 \leq \bar{\alpha}$  and the concentration of  $\hat{\alpha}_{pre}$ , we deduce from step 1:

$$\begin{aligned}
\mathbb{E}_{\mathcal{P}, \hat{\alpha}_{pre}} [|\tilde{\alpha}_0 - \alpha|^2] & \leq \mathbb{E}_{\mathcal{P}, \hat{\alpha}_{pre}} \left[ \mathbb{E}_{\mathcal{P}, \hat{\alpha}_{pre}} [|\tilde{\alpha}_0 - \alpha|^2 | \hat{\alpha}_{pre}] \mathbf{1}_{\{|\hat{\alpha}_{pre} - \alpha| < |\log \varepsilon|^{-1}\}} \right] \\
& \quad + 4\bar{\alpha}^2 \mathbb{P}_{\hat{\alpha}_{pre}} (|\hat{\alpha}_{pre} - \alpha| \geq |\log \varepsilon|^{-1}) \\
& \lesssim \varepsilon^{4(s-1)/(2s+2T\alpha+1)} + 4\bar{\alpha}^2 d\varepsilon^2.
\end{aligned}$$

Since the second term decreases faster than the first one for  $\varepsilon \rightarrow 0$ , we obtain the claimed rate.

## 7.4. Proof of Proposition 5.2

Let  $\varepsilon < 1$ . Recall that the cut-off value of  $\hat{\alpha}_0$  is given by  $U = \varepsilon^{-2/(2s+2T\bar{\alpha}+1)}$ . For  $\kappa > 0$  we obtain from the definition of the estimator and the decomposition of the stochastic error into linear part and remainder:

$$\begin{aligned}
\mathbb{P}(|\hat{\alpha}_0 - \alpha| \geq \kappa) & = \mathbb{P}\left(\left|\int_{-U}^U \operatorname{Re}(\rho + \tilde{\psi} - \psi)(u) w_{\alpha_0}^U(u) du\right| \geq \kappa\right) \\
& \leq \mathbb{P}\left(\left|\int_{-U}^U \rho(u) w_{\alpha_0}^U(u) du\right| \geq \frac{\kappa}{3}\right) + \mathbb{P}\left(\left|\int_{-U}^U \operatorname{Re}(\mathcal{L}(u)) w_{\alpha_0}^U(u) du\right| \geq \frac{\kappa}{3}\right) \\
& \quad + \mathbb{P}\left(\left|\int_{-U}^U \mathcal{R}(u) w_{\alpha_0}^U(u) du\right| \geq \frac{\kappa}{3}\right) \\
& =: P_1 + P_2 + P_3.
\end{aligned}$$

We will bound all three probabilities separately. To that end, let  $c_j, j \in \mathbb{N}$ , be suitable non-negative constants not depending on  $\kappa, \varepsilon$  and  $N$ .

The event in  $P_1$  is deterministic. Hence, the same estimate on the deterministic error as in Theorem 4.2

$$\left| \int_{-U}^U \rho(u) w_{\alpha_0}^U(u) du \right| \leq c_1 U^{-(s-1)} = c_1 \varepsilon^{2(s-1)/(2s+2T\bar{\alpha}+1)}$$

yields  $P_1 = 0$  for all  $\varepsilon < \varepsilon^{(1)} := (\kappa/(3c_1))^{(2s+2T\bar{\alpha}+1)/(2s-2)}$ .

To bound  $P_2$  we infer from the definition of  $\mathcal{L}$ , the linear appearance of the errors in  $\tilde{\mathcal{O}} = \mathcal{O}_l + \sum_{j=1}^N \delta_j \varepsilon_j b_j$  and from the estimate of the term  $|\mathcal{L}_b|$  in Theorem 4.2:

$$\begin{aligned} \left| \int_{-U}^U \operatorname{Re}(\mathcal{L}(u)) w_{\alpha_0}^U(u) du \right| &= \left| \int_{-U}^U \operatorname{Re} \left( \frac{(i-u)u}{T\varphi_T(u-i)} \mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O})(u) \right) w_{\alpha_0}^U(u) du \right| \\ &\leq \int_{-U}^U \frac{(u^4 + u^2)^{1/2}}{T|\varphi_T(u-i)|} |\mathcal{F}(\mathcal{O}_l - \mathcal{O})(u) w_{\alpha_0}^U(u)| du \\ &\quad + \left| \int_{-U}^U \operatorname{Re} \left( \frac{(i-u)u}{T\varphi_T(u-i)} \sum_{j=1}^N \delta_j \varepsilon_j \mathcal{F}b_j(u) \right) w_{\alpha_0}^U(u) du \right| \\ &\leq c_2 \Delta^2 U^{T\bar{\alpha}+2} + \left| \sum_{j=1}^N \delta_j \varepsilon_j \int_{-U}^U \operatorname{Re} \left( \frac{(i-u)u}{T\varphi_T(u-i)} \mathcal{F}b_j(u) \right) w_{\alpha_0}^U(u) du \right| \\ &\leq c_2 \varepsilon^{2(s-1)/(2s+2T\bar{\alpha}+1)} + \left| \sum_{j=1}^N a_j \varepsilon_j \right| \end{aligned}$$

where the coefficients are given by

$$a_j := \delta_j \int_{-U}^U \operatorname{Re} \left( \frac{(i-u)u}{T\varphi_T(u-i)} \mathcal{F}b_j(u) \right) w_{\alpha_0}^U(u) du, \quad j = 1, \dots, N.$$

To apply (5.4), we deduce from  $\|\mathcal{F}b_j\|_\infty \leq 2\Delta$ , the weight function property (4.1) and the assumption  $\Delta \|\delta\|_{l^2}^2 \lesssim \|\delta\|_{l^\infty}^2$ :

$$\begin{aligned} \sum_{j=1}^N a_j^2 &\leq \sum_{j=1}^N \delta_j^2 \left( \int_{-U}^U \frac{(u^4 + u^2)^{1/2}}{T|\varphi_T(u-i)|} |\mathcal{F}b_j(u)| |w_{\alpha_0}^U(u)| du \right)^2 \leq c_3 \Delta^2 U^{2T\bar{\alpha}+4} \|\delta\|_{l^2}^2 \\ &\leq c_4 \varepsilon^2 U^{2T\bar{\alpha}+4} = c_4 \varepsilon^{2(s-1)/(2s+2T\bar{\alpha}+1)}. \end{aligned}$$

This implies through the concentration inequality of  $(\varepsilon_j)$

$$\begin{aligned} P_2 &\leq \mathbb{P} \left( \left| \sum_{j=1}^N a_j \varepsilon_j \right| \geq \frac{\kappa}{6} \right) + \mathbb{P} \left( c_2 \varepsilon^{2(s-1)/(2s+2T\bar{\alpha}+1)} \geq \frac{\kappa}{6} \right) \\ &\leq C_1 \exp \left( - \frac{C_2}{36c_4} \kappa^2 \varepsilon^{-2(s-1)/(2s+2T\bar{\alpha}+1)} \right) \end{aligned}$$

for all  $\varepsilon < \varepsilon^{(2)} := (\kappa/(6c_2))^{(2s+2T\bar{\alpha}+1)/(2s-2)}$ .

It remains to estimate probability  $P_3$ . The bound of  $\mathcal{R}$  in Proposition 7.2 ii) yields

$$\begin{aligned} & \left| \int_{-U}^U \mathcal{R}(u) w_{\alpha_0}^U(u) du \right| \leq \int_{-U}^U \frac{u^4 + u^2}{T\kappa(u)^2} |\mathcal{F}(\tilde{\mathcal{O}} - \mathcal{O})(u)|^2 |w_{\alpha_0}^U(u)| du \\ & \leq 2 \int_{-U}^U \frac{u^4 + u^2}{T\kappa(u)^2} |\mathcal{F}(\mathcal{O}_l - \mathcal{O})(u)|^2 |w_{\alpha_0}^U(u)| du + 2 \int_{-U}^U \frac{u^4 + u^2}{T\kappa(u)^2} \left| \sum_{j=1}^N \delta_j \varepsilon_j \mathcal{F}b_j(u) \right|^2 |w_{\alpha_0}^U(u)| du. \end{aligned}$$

The first addend gets small owing to Proposition 7.2 i):

$$\begin{aligned} \int_{-U}^U \frac{u^4 + u^2}{T\kappa(u)^2} |\mathcal{F}(\mathcal{O}_l - \mathcal{O})(u)|^2 |w_{\alpha_0}^U(u)| du & \leq \|\mathcal{F}(\mathcal{O}_l - \mathcal{O})\|_\infty^2 \int_{-U}^U \frac{u^4 + u^2}{T\kappa(u)^2} |w_{\alpha_0}^U(u)| du \\ & \leq c_5 \Delta^4 U^{2T\bar{\alpha}+4} \leq c_5 \varepsilon^{2(s-1)/(2s+2T\bar{\alpha}+1)}. \end{aligned}$$

For the second one we obtain

$$\left| \sum_{j=1}^N \delta_j \varepsilon_j \mathcal{F}b_j(u) \right|^2 = \sum_{j=1}^N \delta_j^2 \varepsilon_j^2 |\mathcal{F}b_j(u)|^2 + 2 \sum_{j=2}^N \sum_{k=1}^{j-1} \delta_j \delta_k \varepsilon_j \varepsilon_k \operatorname{Re}(\mathcal{F}b_j(u) \mathcal{F}b_k(-u)).$$

Thus,

$$\begin{aligned} & \left| \int_{-U}^U \mathcal{R}(u) w_{\alpha_0}^U(u) du \right| \\ & \leq 2c_5 \varepsilon^{(4s-6)/(2s+2T\bar{\alpha}+1)} + 2 \sum_{j=1}^N \delta_j^2 \varepsilon_j^2 \xi_{j,j}(U) + 4 \sum_{j=2}^N \sum_{k=1}^{j-1} \delta_j \delta_k \varepsilon_j \varepsilon_k \xi_{j,k}(U) \end{aligned}$$

with

$$\xi_{j,k}(U) := \int_{-U}^U \frac{u^4 + u^2}{T\kappa(u)^2} \operatorname{Re}(\mathcal{F}b_j(u) \mathcal{F}b_k(-u)) |w_{\alpha_0}^U(u)| du.$$

Denoting the diagonal term and the cross term as

$$D_N := \sum_{j=1}^N \delta_j^2 \varepsilon_j^2 \xi_{j,j}(U) \quad \text{and} \quad U_N := \sum_{j=2}^N \sum_{k=1}^{j-1} \delta_j \delta_k \varepsilon_j \varepsilon_k \xi_{j,k}(U),$$

respectively, we obtain

$$P_3 \leq \mathbb{P}\left(2c_5 \varepsilon^{2(s-1)/(2s+2T\bar{\alpha}+1)} \geq \frac{\kappa}{9}\right) + \mathbb{P}\left(2D_N \geq \frac{\kappa}{9}\right) + \mathbb{P}\left(4U_N \geq \frac{\kappa}{9}\right).$$

The first summand vanishes for  $\varepsilon < \varepsilon^{(3)} := (\kappa/(18c_5))^{(2s+2T\bar{\alpha}+1)/(2s-2)}$ . To estimate the probabilities on  $D_N$  and  $U_N$ , we establish the bound

$$|\xi_{j,k}(U)| \leq \|\mathcal{F}b_j\|_\infty \|\mathcal{F}b_k\|_\infty \int_{-U}^U \frac{u^4 + u^2}{T\kappa(u)^2} |w_{\alpha_0}^U(u)| du \leq c_6 \Delta^2 U^{2T\bar{\alpha}+4} \quad (7.7)$$

for  $j, k = 1, \dots, N$ . Hence,

$$\left| \sum_{j=1}^N \delta_j^2 \xi_{j,j}(U) \right| \leq c_6 \Delta^2 \|\delta\|_{l^2}^2 U^{2T\bar{\alpha}+4} \leq c_7 \varepsilon^2 U^{2T\bar{\alpha}+4} \leq c_7 \varepsilon^{2(s-1)/(2s+2T\bar{\alpha}+1)},$$

which yields together with (5.4)

$$\begin{aligned} \mathbb{P}\left(D_N \geq \frac{\kappa}{18}\right) &\leq \mathbb{P}\left(\sup_{k=1, \dots, N} |\varepsilon_k|^2 \left| \sum_{j=1}^N \delta_j^2 \xi_{j,j}(U) \right| \geq \frac{\kappa}{18}\right) \\ &\leq \mathbb{P}\left(\sup_{k=1, \dots, N} |\varepsilon_k|^2 \geq \frac{\kappa}{18c_7} \varepsilon^{-2(s-1)/(2s+2T\bar{\alpha}+1)}\right) \\ &\leq C_1 N \exp\left(-\frac{C_2}{18c_7} \kappa \varepsilon^{-2(s-1)/(2s+2T\bar{\alpha}+1)}\right). \end{aligned}$$

To derive an exponential inequality for the U-statistic  $U_N$ , we apply the martingale idea of Houdré and Reynaud-Bouret [14]: Because of the independence and the centering of the  $(\varepsilon_j)$ , the process  $(U_N)_{N \geq 1}$  is a martingale with respect to its natural filtration  $(\mathcal{F}_N^U)$  (setting  $U_1 = 0$ ):

$$\mathbb{E}[U_N - U_{N-1} | \mathcal{F}_{N-1}^U] = \mathbb{E}\left[\sum_{k=1}^{N-1} \delta_N \delta_k \varepsilon_N \varepsilon_k \xi_{N,k}(U) | \mathcal{F}_{N-1}^U\right] = 0.$$

We will apply the following martingale version of the Bernstein inequality:

**Proposition 7.4** (Bernstein's inequality). *Let  $(M_n, \mathcal{F}_n)$  be a martingale with  $M_0 = 0$ . For arbitrary  $t, Q, S > 0$  the following holds true:*

$$\mathbb{P}(|M_n| \geq t) \leq 2\mathbb{P}(\langle M \rangle_n > Q) + 2\mathbb{P}\left(\max_{k=1, \dots, n} |M_k - M_{k-1}| > S\right) + 2 \exp\left(-\frac{t^2}{4(Q + tS)}\right).$$

Hence, we consider the increment,  $N \geq 2$ ,

$$|U_N - U_{N-1}| = |\varepsilon_N| \left| \sum_{k=1}^{N-1} \underbrace{\delta_N \delta_k \xi_{N,k}(U)}_{=: a_{N,k}} \varepsilon_k \right|$$

for which we estimate using (7.7)

$$\begin{aligned} \sum_{k=1}^{N-1} a_{N,k}^2 &= \delta_N^2 \sum_{k=1}^{N-1} \delta_k^2 \xi_{N,k}(U)^2 \leq c_6^2 \Delta^4 U^{4T\bar{\alpha}+8} \delta_N^2 \|\delta\|_{l^2}^2 \\ &\leq c_6^2 \Delta^4 \|\delta\|_{l^2}^4 U^{4T\bar{\alpha}+8} \leq c_7^2 \varepsilon^4 U^{4T\bar{\alpha}+8} \leq c_7^2 \varepsilon^{4(s-1)/(2s+2T\bar{\alpha}+1)}. \end{aligned} \quad (7.8)$$



Thus, by Assumption (5.4) we obtain for all  $S > 0$

$$\begin{aligned} \mathbb{P}(|U_N - U_{N-1}| > S) &= \mathbb{P}\left(|\varepsilon_N| \left| \sum_{k=1}^{N-1} a_{N,k} \varepsilon_k \right| > S\right) \\ &\leq \mathbb{P}\left(|\varepsilon_N| > \sqrt{S} \varepsilon^{-(s-1)/(2s+2T\bar{\alpha}+1)}\right) + \mathbb{P}\left(\left| \sum_{k=1}^{N-1} a_{N,k} \varepsilon_k \right| > \sqrt{S} \varepsilon^{-(s-1)/(2s+2T\bar{\alpha}+1)}\right) \\ &\leq C_1 \exp\left(-C_2 S \varepsilon^{-2(s-1)/(2s+2T\bar{\alpha}+1)}\right) + C_1 \exp\left(-\frac{C_2}{c_7^2} S \varepsilon^{-2(s-1)/(2s+2T\bar{\alpha}+1)}\right). \end{aligned}$$

The quadratic variation of  $U_N$  is given by

$$\langle U \rangle_N - \langle U \rangle_{N-1} = \mathbb{E}[(U_N - U_{N-1})^2 | \mathcal{F}_{N-1}^U] = \delta_N^2 \left( \sum_{k=1}^{N-1} \delta_k \varepsilon_k \xi_{N,k}(U) \right)^2.$$

W.l.o.g. we can assume  $\sum_{j=2}^N \delta_j^2 > 0$ . Otherwise follows  $\sum_{j=2}^N \delta_j^2 = 0$  which implies  $\delta_j = 0$  for all  $j = 2, \dots, N$  and thus  $\langle U \rangle_N = \sum_{j=2}^N (\langle U \rangle_j - \langle U \rangle_{j-1}) = 0$ . Then  $\mathbb{P}(\langle U \rangle_N > Q) = 0$  would hold for  $Q > 0$ . Hence, we obtain:

$$\begin{aligned} \mathbb{P}(\langle U \rangle_N > Q) &= \mathbb{P}\left(\sum_{j=2}^N (\langle U \rangle_j - \langle U \rangle_{j-1}) > Q\right) \leq \sum_{j=2}^N \mathbb{P}\left(\langle U \rangle_j - \langle U \rangle_{j-1} > \frac{\delta_j^2}{\sum_{k=2}^N \delta_k^2} Q\right) \\ &\leq \sum_{j=2}^N \mathbb{P}\left(\|\delta\|_{l^2} \sum_{k=1}^{j-1} \delta_k \varepsilon_k \xi_{j,k}(U) > \sqrt{Q}\right). \end{aligned}$$

To apply inequality (5.4) we estimate  $\|\delta\|_{l^2}^2 \sum_{k=1}^{j-1} \delta_k^2 \xi_{j,k}(U)^2 \leq c_6^2 \Delta^4 \|\delta\|_{l^2}^4 U^{4T\bar{\alpha}+8} \leq c_7^2 \varepsilon^{4(s-1)/(2s+2T\bar{\alpha}+1)}$  analogous to (7.8) and obtain

$$\mathbb{P}(\langle U \rangle_N > Q) \leq C_1 N \exp\left(-\frac{C_2}{c_7^2} Q \varepsilon^{-4(s-1)/(2s+2T\bar{\alpha}+1)}\right).$$

We deduce from Bernstein's inequality:

$$\begin{aligned} &\mathbb{P}\left(U_N \geq \frac{\kappa}{36}\right) \\ &\leq 2\mathbb{P}(\langle U \rangle_N > Q) + 2\mathbb{P}\left(\max_{k=2, \dots, N} |U_k - U_{k-1}| > S\right) + 2 \exp\left(-\frac{\kappa^2}{144(36Q + \kappa S)}\right) \\ &\leq 2C_1 N \exp\left(-\frac{C_2}{c_7^2} Q \varepsilon^{-4(s-1)/(2s+2T\bar{\alpha}+1)}\right) \\ &\quad + 4C_1 N \exp\left(-\frac{C_2}{c_7^2 \vee 1} S \varepsilon^{-2(s-1)/(2s+2T\bar{\alpha}+1)}\right) + 2 \exp\left(-\frac{\kappa^2}{144(36Q + \kappa S)}\right). \end{aligned}$$

By choosing  $Q = \kappa S$  and  $S = \sqrt{\kappa} \varepsilon^{-(s-1)/(2s+2T\bar{\alpha}+1)}$  we get

$$\mathbb{P}\left(U_N \geq \frac{\kappa}{36}\right) \leq (6C_1 N + 2) \exp\left(-c_8 \min_{q=1,3} (\kappa^{1/2} \varepsilon^{-(s-1)/(2s+2T\bar{\alpha}+1)})^q\right).$$

For all  $\varepsilon < \varepsilon^{(3)}$  we have  $\kappa \varepsilon^{-2(s-1)/(2s+2T\bar{\alpha}+1)} > \kappa (\varepsilon^{(3)})^{-2(s-1)/(2s+2T\bar{\alpha}+1)} \sim 1$  and hence,

$$P_3 \leq \mathbb{P}(D_N \geq \frac{\kappa}{18}) + \mathbb{P}(U_N \geq \frac{\kappa}{36}) \leq (7C_1N + 2) \exp\left(-c_8 \kappa^{1/2} \varepsilon^{-(s-1)/(2s+2T\bar{\alpha}+1)}\right).$$

Putting the bounds of  $P_1, P_2$  and  $P_3$  together yields for a constant  $c \in (0, \infty)$  and all  $\varepsilon < \varepsilon_0 \wedge 1$  with  $\varepsilon_0 := \min\{\varepsilon^{(1)}, \varepsilon^{(2)}, \varepsilon^{(3)}\}$

$$\mathbb{P}(|\hat{\alpha}_0 - \alpha| \geq \kappa) \leq (7C_1N + C_1 + 2) \exp\left(-c(\kappa^2 \wedge \kappa^{1/2}) \varepsilon^{-(s-1)/(2s+2T\bar{\alpha}+1)}\right).$$

## Appendix A: Proof of Lemma 2.1

*Part i)* The martingale condition yields

$$|\varphi_T(u - i)| = \exp\left(T \int_{-\infty}^{\infty} (\cos(ux) - 1) \frac{e^x k(x)}{|x|} dx\right).$$

W.l.o.g. we assume  $T = 1$ ,  $\alpha > 0$  and  $u \geq 1$  because of the symmetry of the cosine.

*Step 1:* Let  $k(0-) = 0$ . We split the integral domain into three parts:

$$|\varphi_1(u - i)| = \exp\left(\left(\int_0^1 + \int_1^u + \int_u^\infty\right) (\cos x - 1) \frac{e^{x/u} k(\frac{x}{u})}{x} dx\right).$$

Using the monotonicity of  $k$  and the constant  $C_1 := \int_0^1 \frac{1 - \cos x}{x} dx \in (0, \infty)$ , we estimate

$$\int_0^1 (\cos x - 1) \frac{e^{x/u} k(\frac{x}{u})}{x} dx \geq e^{1/u} k(0+) \int_0^1 \frac{\cos x - 1}{x} dx \geq -C_1 e k(0+).$$

In the second part the dependence on  $u$  comes into play. The Taylor series of the exponential function together with dominated convergence yield:

$$\begin{aligned} \int_1^u (\cos x - 1) \frac{e^{x/u} k(\frac{x}{u})}{x} dx &\geq k(0+) \left( \int_1^u \frac{\cos x}{x} - \frac{1}{x} dx + \sum_{k=1}^{\infty} \int_1^u (\cos x - 1) \frac{x^{k-1}}{u^k k!} dx \right) \\ &\geq k(0+) \underbrace{\left( \min_{v \geq 1} \int_1^v \frac{\cos x}{x} dx - \log(u) - 2 \sum_{k=1}^{\infty} \frac{1}{k! k} (1 - u^{-k}) \right)}_{=: -C_2 \leq 0} \geq k(0+) (-C_2 - 2e - \log(u)). \end{aligned}$$

Note that the constant  $C_2$  is finite since  $x \mapsto \frac{\cos(x)}{x}$  is Riemann integrable on  $[1, \infty)$ . We obtain for the third part:

$$\int_1^\infty (\cos x - 1) \frac{e^x k(x)}{x} dx \geq -2 \int_1^\infty e^x k(x) dx.$$

This achieves the estimate

$$|\varphi_1(u - i)| \geq \exp \left( k(0+)(-(C_1 + 2)e - C_2) - 2 \int_1^\infty e^x k(x) dx \right) u^{-k(0+)}.$$

*Step 2:* Suppose  $k(0+) = 0$ . By substituting  $x = -y$  we derive similarly:

$$|\varphi_1(u - i)| \geq \exp \left( -k(0-)(C_1 + C_2) - 2 \int_{-\infty}^{-1} e^x k(x) dx \right) u^{-k(0-)}.$$

*Step 3:* Let  $k(0-) > 0$  and  $k(0+) > 0$ . We split the integral domain into  $\mathbb{R}_+$  and  $\mathbb{R}_-$  and deduce from steps 1 and 2 the estimate  $|\varphi_T(u - i)| \geq C_\varphi(T, \|e^x k(x)\|_{L^1}, \alpha) u^{-T\alpha}$  for  $|u| \geq 1$  and with  $C_\varphi(T, R, \alpha) := \exp(T\alpha(-(C_1 + 2)e + C_2) - 2TR)$ . *Part ii)* follows immediately from the explicit choice of  $C_\varphi$ .

## Appendix B: White noise model and lower bounds

To establish asymptotic lower bounds for the convergence rates of the estimators  $\hat{\gamma}, \hat{\alpha}_j$ ,  $j = 0, \dots, s - 2$  and  $\hat{k}_e$ , we consider the continuous white noise model

$$dZ_{\mathcal{P}}(x) = \mathcal{O}_{\mathcal{P}}(x) dx + \frac{1}{\sqrt{N}} \lambda_N(x) dW(x), \quad x \in [-A_N, A_N], \quad (\text{B.1})$$

with a two sided Brownian motion  $W$ , an option function  $\mathcal{O}_{\mathcal{P}}$  induced by the pair  $\mathcal{P} \in \mathcal{G}_s(R, \bar{\alpha})$  and  $A_N > 0$  growing in  $N$ . Under certain conditions, we can apply the results of Brown and Low [5] to show the asymptotic equivalence of the above considered regression model

$$\mathcal{O}_j = \mathcal{O}(x_j) + \delta_j \varepsilon_j, \quad j = 1, \dots, N, \quad (\text{B.2})$$

and the white noise model (B.1) for  $N \rightarrow \infty$ . Setting  $N$  and  $\varepsilon$  in relation to each other allows us to derive lower bounds in terms of  $\varepsilon$ .

To that end, we state the situation of (B.2) more precisely. Let  $N \in \mathbb{N}$  and  $H_N : [-A_N, A_N] \rightarrow [0, 1]$  be an increasing, absolutely continuous distribution function with

$$h_N := H'_N > 0 \text{ a.e. on } [-A_N, A_N] \quad (\text{B.3})$$

such that the strikes are given by  $x_j = H_N^{-1}(j/(N + 1))$  for  $j = 1, \dots, N$ . Furthermore, let  $\delta_j = \delta(x_j)$  for some function  $\delta \in L^\infty(\mathbb{R})$ . We suppose that  $\delta$  is absolutely continuous satisfying the technical condition

$$\left| \frac{d}{dx} \log \delta(x) \right| \leq C_\delta, \quad x \in \mathbb{R}, \quad (\text{B.4})$$

for some constant  $C_\delta < \infty$ . From Belomestny and Reiß [3, Prop. 2.1] we know that  $\mathcal{O}'$  is continuous on  $\mathbb{R} \setminus \{0\}$  with a jump at zero of height -1.  $\|\mathcal{O}''\|_{L^1} \leq 3$  implies  $|\mathcal{O}'(x)| = \left| \int_0^x \mathcal{O}''(x) dx \right| \leq 3$  for all  $x \neq 0$  and thus,  $\mathcal{O}'$  is uniformly bounded. Especially,  $\mathcal{O}$  is Lipschitz continuous for any  $\mathcal{P} \in \mathcal{G}_s(R, \bar{\alpha}) \cup \mathcal{H}_s(R, \bar{\alpha})$ . The equivalence result follows immediately from Brown and Low [5, Thm. 4.1, Cor. 4.2].

**Proposition B.1.** *Under the assumptions (B.3) and (B.4) the nonparametric regression model (B.2) with  $\varepsilon_j \sim N(0, 1)$  iid. and the white noise model (B.1) are asymptotically equivalent in Le Cam's sense if  $\lambda_N^2(x) = \delta^2(x)/h_N(x)$  for  $x \in [-A_N, A_N]$ .*

**Remark B.2.** Grama and Nussbaum [13] showed the asymptotic equivalence of location type regression models with non-Gaussian noise to regression models with Gaussian noise. Hence, the condition on the distribution of  $\varepsilon_j$  in the preceding proposition can be relaxed to error laws with regular densities.

By Proposition B.1 asymptotic lower bounds in the regression model can be proved in the white noise setting equivalently. To this end we need a uniform lower bound of the noise level  $\lambda_N/\sqrt{N}$  in terms of  $\varepsilon$ . Let the strike distribution be polynomial:

$$h_N(x) \sim C_N^{-1}(|x| + 1)^{-q}, \quad x \in [-A_N, A_N], \quad (\text{B.5})$$

for some  $q \geq 0$  and normalization constant  $C_N$ . This is reasonable since in practice most of the traded options are almost at-the-money whereas only less ones are far in- or out-of-the-money. Moreover, we suppose a minimal noise through

$$\delta^2(x) \gtrsim \|\delta\|_\infty^2 (|x| + 1)^{-p}, \quad x \in \mathbb{R}, \quad (\text{B.6})$$

where  $p > 1$  is necessary. The restriction on  $p$  is because the condition  $\Delta\|(\delta_j)\|_{l^2}^2 \lesssim \|(\delta_j)\|_{l^\infty}^2$  in the Theorem 4.2 implies necessarily  $\delta \in L^2(\mathbb{R})$  which is due to Fatou's lemma:

$$\|\delta\|_{L^2}^2 \leq \liminf_{N \rightarrow \infty} \left( \Delta\delta^2(x_1) + \sum_{j=2}^N (x_j - x_{j-1})\delta^2(x_j) \right) \leq \liminf_{N \rightarrow \infty} \Delta\|(\delta_j)\|_{l^2}^2 \lesssim \|\delta\|_\infty^2.$$

In the same way  $\Delta\|(x_j\delta_j)\|_{l^2}^2 \lesssim \|(\delta_j)\|_{l^\infty}^2$  implies even  $x\delta(x) \in L^2(\mathbb{R})$  in the situation of Theorem 4.5. In view of the condition  $e^{-A_N} \lesssim \Delta^2 \sim \varepsilon^4$  and  $A_N e^{-A_N} \lesssim \Delta^2$ , respectively, we assume additionally  $A_N \sim \log \varepsilon^{-1}$ .

**Lemma B.3.** *Let the properties (B.5) and (B.6) be satisfied. If  $A_N \sim \log \varepsilon^{-1}$  holds then we obtain  $\frac{\lambda_N^2(x)}{N} \gtrsim \varepsilon^2 (\log \varepsilon^{-1})^{-\beta}$  for  $\beta := \max\{p, q\}$ .*

The proof of the lemma is straight forward and thus omitted. In the sequel we assume always that the model (B.1) satisfies the conditions of this lemma.

**Remark B.4.** For the estimators  $\hat{\gamma}$  and  $\hat{\alpha}_j$  with  $s \geq 3$  we can consider the case  $p \in [0, 1)$  because of Remark 4.3. In the situation of  $0 = p = q$  we obtain  $\beta = 0$  and hence, the lower bounds proved in the next theorems imply that our estimation procedure achieves exact minimax rates for the parameters. With regard to  $\hat{k}_e$  the condition  $\Delta\|(x_j\delta_j)\|_{l^2}^2 \lesssim \|(\delta_j)\|_{l^\infty}^2$  yields  $\beta \geq p > 3$ .

**Theorem B.5.** *Let  $s \in \mathbb{N}$ ,  $s \geq 2$ ,  $R, \bar{\alpha} > 0$  and  $j = 0, \dots, s-2$ . We obtain in the observation model (B.1) the asymptotic risk lower bounds*

$$\begin{aligned} \inf_{\hat{\gamma}} \sup_{\mathcal{P} \in \mathcal{G}_s(R, \bar{\alpha})} \mathbb{E}_{\mathcal{P}}[|\hat{\gamma} - \gamma|^2]^{1/2} &\gtrsim (\varepsilon(\log \varepsilon^{-1})^{-\beta/2})^{2s/(2s+2T\bar{\alpha}+1)}, \\ \inf_{\hat{\alpha}_j} \sup_{\mathcal{P} \in \mathcal{G}_s(R, \bar{\alpha})} \mathbb{E}_{\mathcal{P}}[|\hat{\alpha}_j - \alpha_j|^2]^{1/2} &\gtrsim (\varepsilon(\log \varepsilon^{-1})^{-\beta/2})^{2(s-1-j)/(2s+2T\bar{\alpha}+1)} \quad \text{and} \\ \inf_{\hat{k}_e} \sup_{\mathcal{P} \in \mathcal{H}_s(R, \bar{\alpha})} \mathbb{E}_{\mathcal{P}}[\|\hat{k}_e - k_e\|_{L^2, \tau}^2]^{1/2} &\gtrsim (\varepsilon(\log \varepsilon^{-1})^{-\beta/2})^{2s/(2s+2T\bar{\alpha}+5)} \end{aligned}$$

where the infimum is taken over all estimators, i.e. all measurable functions of the observation  $Z$ . The bound for  $k_e$  holds for  $s = 1$  as well.

As noticed these lower bounds are a logarithmic factor better than the convergence rates in the last section in the case  $\beta > 0$ . The following corollary extends the result to the parameter  $\alpha_{s-1} := k^{(s-1)}(0+) + k^{(s-1)}(0-)$ . Especially, the estimation of  $\alpha$  in the case  $s = 1$  is of interest.

**Corollary B.6.** *Consider the situation of the Theorem B.5 where  $s = 1$  is also possible. For  $\alpha_{s-1}$  we obtain asymptotically the risk lower bound*

$$\inf_{\hat{\alpha}_{s-1}} \sup_{\mathcal{P} \in \mathcal{G}_s(R, \bar{\alpha})} \mathbb{E}_{\mathcal{P}}[|\hat{\alpha}_{s-1} - \alpha_{s-1}|^2]^{1/2} \gtrsim 1$$

in the observation model (B.1). Hence, we cannot estimate  $\alpha_{s-1}$  consistently in the  $L^2$ -sense.

## Appendix C: Proofs of lower bounds

First, we are interested in the distance of  $Z_{\mathcal{P}_0}$  and  $Z_{\mathcal{P}_1}$  with  $\mathcal{P}_0, \mathcal{P}_1 \in \mathcal{G}_0(R, \bar{\alpha})$ . Girsanov's theorem implies the equivalence of the laws of  $Z_{\mathcal{P}_\bullet}$  and the likelihood ratio for  $\mathcal{P}_0$  with respect to  $\mathcal{P}_1$ , given by Liptser and Shiryaev [17, Theorem 7.18], is:

$$\begin{aligned} \Lambda(\mathcal{P}_0, \mathcal{P}_1) &= \exp \left( \int_{-A_N}^{A_N} (\mathcal{O}_{\mathcal{P}_1} - \mathcal{O}_{\mathcal{P}_0})(x) \sqrt{N} \lambda_N(x)^{-1} dW(x) \right. \\ &\quad \left. - \frac{1}{2} \int_{-A_N}^{A_N} |\mathcal{O}_{\mathcal{P}_1} - \mathcal{O}_{\mathcal{P}_0}|^2(x) N \lambda_N(x)^{-2} dx \right). \end{aligned}$$

Hence, we can bound the Kullback-Leibler divergence using the uniform bound of  $N/\lambda_N^2$ , Parseval's identity and the pricing formula

$$\begin{aligned}
KL(\mathcal{P}_1|\mathcal{P}_0) &= \mathbb{E}_{\mathcal{P}_1} [\log (\Lambda(\mathcal{P}_1, \mathcal{P}_0))] = \frac{1}{2} \int_{-A_N}^{A_N} |\mathcal{O}_{\mathcal{P}_1} - \mathcal{O}_{\mathcal{P}_0}|^2(x) N \lambda_N(x)^{-2} dx \\
&\sim \varepsilon^{-2} (\log \varepsilon^{-1})^\beta \int_{-\infty}^{\infty} |\mathcal{F}(\mathcal{O}_{\mathcal{P}_1} - \mathcal{O}_{\mathcal{P}_0})(u)|^2 du \\
&\sim \varepsilon^{-2} (\log \varepsilon^{-1})^\beta \int_{-\infty}^{\infty} \left| \frac{\varphi_{T, \mathcal{P}_0}(u-i) - \varphi_{T, \mathcal{P}_1}(u-i)}{u(u-i)} \right|^2 du. \tag{C.1}
\end{aligned}$$

### C.1. Lower bound for $\gamma$ :

Following the standard approach, we perturb a pair  $\mathcal{P}_0 \in \mathcal{G}_s(R, \bar{\alpha})$ . Let  $\mathcal{P}_0 = (\gamma_0, k_0)$  satisfies all conditions where norms and constants are strictly smaller than  $R$  and with  $\alpha = \bar{\alpha}$ . Furthermore, let  $k_0$  do not decrease too rapidly, i.e., we assume  $k_0 \gtrsim |x|^{-p}$  and  $|k'_0| \gtrsim |x|^{-q}$  for some  $p, q > 0$ , and  $|\varphi_T(u-i)| \lesssim |u|^{-T\alpha}$  hold exactly. Certainly such a pair exists.

Let  $\delta > 0$  and consider  $\mathcal{P}_1 = (\gamma_1, k_1)$  given by

$$\gamma_1 = \gamma_0 + \delta \quad \text{and} \quad \mathcal{F}\left(\frac{k_1(x) - k_0(x)}{|x|} e^x\right)(u) = -\delta i(u-i) e^{-u^{2m}/U^{2m}}$$

where we will choose  $U > 1$  and  $m \in \mathbb{N}$  properly. In the following we call the difference of the exponentially scaled jump measures  $g(x) := \frac{k_1(x) - k_0(x)}{|x|} e^x$ . By construction  $g$  is real valued and the martingale condition is valid:

$$\begin{aligned}
\gamma_1 + \int_{\mathbb{R}} (e^x - 1) \frac{k_1(x)}{|x|} dx &= \gamma_0 + \delta + \int_{\mathbb{R}} (e^x - 1) \frac{k_0(x)}{|x|} dx + \mathcal{F}g(0) - \mathcal{F}g(i) \\
&= \gamma_0 + \delta + \int_{\mathbb{R}} (e^x - 1) \frac{k_0(x)}{|x|} dx - \delta = 0.
\end{aligned}$$

Also the moment assumption can be checked straight forward for  $\delta$  small enough:

$$\begin{aligned}
\mathbb{E}_{\mathcal{P}_1}[e^{2X_T}] &= \mathbb{E}_{\mathcal{P}_0}[e^{2X_T}] \cdot \exp\left(2\delta + \int_{-\infty}^{\infty} (e^x - e^{-x})g(x) dx\right) \\
&= \mathbb{E}_{\mathcal{P}_0}[e^{2X_T}] \cdot \exp\left(2\delta(1 - e^{(-1)^{m+1}/U^{2m}})\right) < e^{2\delta} \mathbb{E}_{\mathcal{P}_0}[e^{2X_T}] \leq R.
\end{aligned}$$

Using the Schwartz-functions  $\zeta_1 := \mathcal{F}^{-1}(-iue^{-u^{2m}})$ ,  $\zeta_2 := \mathcal{F}^{-1}(-e^{-u^{2m}}) \in \mathcal{S}(\mathbb{R})$  the inversion and scaling properties of the Fourier transform yield for  $u \in \mathbb{R}$

$$\mathcal{F}g(u) = \delta U \mathcal{F}\zeta_1\left(\frac{u}{U}\right) + \delta \mathcal{F}\zeta_2\left(\frac{u}{U}\right) = \mathcal{F}(\delta U^2 \zeta_1(Ux) + \delta U \zeta_2(Ux))(u).$$

Hence, we have  $g(x) = \delta U^2 \zeta_1(Ux) + \delta U \zeta_2(Ux) \in \mathcal{S}(\mathbb{R})$ . Even  $e^{-x}g(x) \in \mathcal{S}(\mathbb{R})$  holds because of  $\mathcal{F}(e^{-x}\zeta_j(x))(u) = \mathcal{F}\zeta_j(u+i)$  for  $u \in \mathbb{R}$  and  $j \in \{1, 2\}$ .

If  $\delta U^2 \lesssim 1$  the disturbance  $|x|e^{-x}g(x)$  and its derivative are bounded for all  $U \geq 1$ , owing to the rescaling with  $U$ :

$$|Uxe^{-x}\zeta_j(Ux)| \leq |Ux(1 \vee e^{-Ux})\zeta_j(Ux)| \leq \|x(1 \vee e^{-x}\zeta_j(x))\|_\infty \leq \infty.$$

Since additionally  $|x|e^xg(x)$  is fast decreasing and  $k_0$  and  $k'_0$  are bounded from below, the k-function  $k_1$  is non-negative and satisfies the monotonicity conditions provided  $\delta$  is small enough.

The continuity and polynomial decrease of  $\zeta_j$  imply  $\|(1 \vee e^{-x})\zeta_j^{(l)}\|_{L^1} < \infty, j = 1, 2$ . Furthermore, by construction  $(k_1 - k_0)(0) = 0$  and the derivatives of the disturbance is given by

$$\begin{aligned} (k_1 - k_0)^{(m)}(x) &= \frac{d^m}{dx^m} |x|e^{-x}g(x) = \sum_{l=0}^m \binom{m}{l} (-1)^l (|x| - l \operatorname{sgn} x) e^{-x} g^{(m-l)}(x) \\ &= \delta \sum_{l=0}^m \binom{m}{l} (-1)^l U^{m-l+2} (|x| - l \operatorname{sgn} x) e^{-x} (\zeta_1^{(m-l)}(Ux) + U^{-1} \zeta_2^{(m-l)}(Ux)) \end{aligned}$$

for  $m = 1, \dots, s$  and  $x \neq 0$ . Hence,  $(k_0 - k_1)^{(m)}(0+) + (k_0 - k_1)^{(m)}(0-) = 0$  for  $m = 0, \dots, s$  and substituting  $y = Ux$  yield

$$\begin{aligned} &\|(1 \vee e^x)(k_0 - k_1)^{(s)}(x)\|_{L^1} \\ &\leq \delta \left( U^s \|(1 \vee e^{-y})|y|\zeta_1^s(y)\|_{L^1} + \sum_{l=1}^s \binom{s}{l} U^{s-l+1} \|(1 \vee e^{-y})(|y| + l)\zeta_1^{s-l}(y)\|_{L^1} \right. \\ &\quad \left. + U^{s-1} \|(1 \vee e^{-y})|y|\zeta_2^s(y)\|_{L^1} + \sum_{l=1}^s \binom{s}{l} U^{s-l} \|(1 \vee e^{-y})(|y| + l)\zeta_2^{s-l}(y)\|_{L^1} \right) \\ &\lesssim \delta U^s \end{aligned}$$

and even better bounds for derivatives of lower order. Thus, the norm restrictions are fulfilled by choosing  $U \sim \delta^{-1/s}$ . Additionally,  $\delta U^2 \sim \delta^{(s-2)/s} \lesssim 1$  is valid. Therefore,  $\mathcal{P}_1 \in \mathcal{G}_s(R, \alpha_{max})$  holds. From Tsybakov [26, Thm. 2.2] follows the lower bound

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

once we have shown that the Kullback-Leibler divergence is asymptotically bounded. We deduce from equation (C.1), the estimate  $|1 - e^z| \leq 2|z|$  for all  $z \in \mathbb{C}$  in a small ball around 0 and the assumed decrease  $|\varphi_T(u - i)| \lesssim |u|^{-T\alpha}$  for  $m \geq T\alpha + 1$

$$\begin{aligned} KL(\mathcal{P}_1|\mathcal{P}_0) &\lesssim \frac{|\log \varepsilon|^\beta}{\varepsilon^2} \int_{-\infty}^{\infty} |\varphi_{T, \mathcal{P}_0}(u - i)|^2 |i\delta(u - i) + \mathcal{F}g(u) - \mathcal{F}g(i)|^2 (u^4 + u^2)^{-1} du \\ &= \varepsilon^{-2} (\log \varepsilon^{-1})^\beta \delta^2 \int_{-\infty}^{\infty} |\varphi_{T, \mathcal{P}_0}(u - i)|^2 |1 - e^{-u^2m/U^{2m}}|^2 u^{-2} du. \end{aligned}$$

$$\begin{aligned} &\lesssim \varepsilon^{-2} (\log \varepsilon^{-1})^\beta \delta^2 U^{-2T\alpha-1} \int_{-\infty}^{\infty} |u|^{-2T\alpha-2} |1 - e^{-u^{2m}}|^2 du \\ &\lesssim \varepsilon^{-2} (\log \varepsilon^{-1})^\beta \delta^{(2s+2T\alpha+1)/s} \end{aligned}$$

Hence,  $KL(\mathcal{P}_1|\mathcal{P}_0)$  remains bounded if  $\delta \sim (\varepsilon^2 (\log \varepsilon^{-1})^{-\beta})^{s/(2s+2T\alpha+1)}$ .

## C.2. Lower bound for $\alpha_j$ :

We will need the following auxiliary lemma which is shown in Section C.4 separately:

**Lemma C.1.** *Let  $j \in \{0, \dots, s-1\}$  and  $m \in \mathbb{N}$ . There is a family of functions  $(g_U)_{U \geq 1} \subset C^s(\mathbb{R} \setminus \{0\})$  or  $(g_U)_{U \geq 1} \subset C^{j-1}(\mathbb{R}) \cap C^s(\mathbb{R} \setminus \{0\})$  for  $j = 0$  or  $j \geq 1$ , respectively, such that each  $g_U$  has compact support and satisfies the following conditions for some constants  $S, U_{\min} > 1$  and for all  $U \geq U_{\min}$ :*

- i)  $\int_{-\infty}^{\infty} \frac{e^{x/U} - 1}{|x|} g_U(x) dx = 0$ ,
- ii)  $\int_{-\infty}^{\infty} \left| \frac{e^{2x} - 1}{x} g_U(x) \right| dx \leq S$ ,
- iii)  $g_U(0+) + g_U(0-) \geq 0$ ,  $g_U^{(j)}(0+) + g_U^{(j)}(0-) = 1$  and  $|g_U^{(l)}(0+) + g_U^{(l)}(0-)| \leq S$  for  $l = 0, \dots, s-1$ ,
- iv)  $\|g_U\|_\infty \leq S$  and  $\int_{-\infty}^{\infty} (1 \vee e^x) |g_U^{(l)}(x)| dx \leq S$  for  $l = 0, \dots, s$  as well as
- v)  $\left| u^{-m} \int_{-\infty}^{\infty} \frac{e^{iux} - 1}{|x|} e^{x/U} g_U(x) dx \right| \leq S$ ,  $u \in [-1, 1]$ .

Now, we are in position to prove lower bounds for  $\alpha_j$ . Since the convergence rates of  $\alpha_j$  decrease for rising  $j$  and because of the recursion formula in Lemma 3.1, it is sufficient to consider  $k^{(j)}(0+) + k^{(j)}(0-)$  instead of  $\alpha_j$ . Fix a  $j \in \{0, \dots, s-2\}$ .

We argue analogously to the proof for the estimation of  $\gamma$ : Again we perturb a pair  $\mathcal{P}_0 = (\gamma_0, k_0) \in \mathcal{G}_s(R, \bar{\alpha})$  with exactly the same properties as above. To disturb  $\mathcal{P}_0$  in a suitable way, we choose a family of functions  $(g_U)_{U \geq 1} \subset C^s(\mathbb{R} \setminus \{0\})$  or  $(g_U)_{U \geq 1} \subset C^{j-1}(\mathbb{R}) \cap C^s(\mathbb{R} \setminus \{0\})$  for  $j = 0$  or  $j \geq 1$ , respectively, with the properties i) - v) from Lemma C.1 for some constants  $S, U_{\min} > 1$  and  $m \in \mathbb{N}, m \geq T\alpha + 2$ . For  $\delta > 0$  define  $\mathcal{P}_1 = (\gamma_1, k_1)$  as

$$\gamma_1 := \gamma_0 \quad \text{and} \quad k_1(x) := k_0(x) - \delta U^{-j} g_U(Ux), \quad x \in \mathbb{R}.$$

From i) follows the martingale condition as for  $\gamma$ :

$$0 = \gamma_0 + \int_{\mathbb{R}} (e^x - 1) \frac{k_0(x)}{|x|} dx - \delta U^{-j} \int_{\mathbb{R}} \frac{e^{x/U} - 1}{|x|} g_U(x) dx = \gamma_1 + \int_{\mathbb{R}} (e^x - 1) \frac{k_1(x)}{|x|} dx.$$

As long as  $\delta U^{-j+1}$  is bounded the perturbation and its derivative are bounded in  $U$  such that the necessary monotonicity and non-negativity conditions of  $k_1$  follow as in the proof before. We derive the moment assumption using ii)

$$\mathbb{E}_{\mathcal{P}_1}[e^{2X_T}] = \exp\left(-\delta U^{-j} \int_{\mathbb{R}} \frac{e^{2x} - 1}{|x|} g_U(Ux) dx\right) \mathbb{E}_{\mathcal{P}_0}[e^{2X_T}] \leq e^{S\delta U^{-j}} \mathbb{E}_{\mathcal{P}_0}[e^{2X_T}].$$



Furthermore, the smoothness of  $g_U$  and condition iii) yield

$$\begin{aligned} (k_0 - k_1)^{(j)}(0+) + (k_0 - k_1)^{(j)}(0-) &= \delta \quad \text{and} \\ (k_0 - k_1)^{(s-1)}(0+) + (k_0 - k_1)^{(s-1)}(0-) &\leq S\delta U^{s-1-j}. \end{aligned}$$

Using the integrability condition iv), we estimate

$$\begin{aligned} \|(1 \vee e^x)(k_1 - k_0)^{(s)}\|_{L^1} &= \delta U^{-j} \|(1 \vee e^x)(g_U(Ux))^{(s)}\|_{L^1} \\ &= \delta U^{s-1-j} \|(1 \vee e^{x/U})g_U^{(s)}(x)\|_{L^1} \leq S\delta U^{s-1-j} \end{aligned}$$

and better bounds for derivatives of lower order. Thus,  $\mathcal{P}_1 \in \mathcal{G}_s(R, \alpha_{max})$  if we choose  $U \sim \delta^{-1/(s-1-j)}$  with a constant small enough (note  $\delta U^{-j+1} \sim \delta^{(s-2)/(s-1-j)} \lesssim 1$ ).

With respect to condition v), the uniform boundedness of  $g_U$ , their compact support and the estimate (C.1) the Kullback-Leibler divergence is bounded by

$$\begin{aligned} &KL(\mathcal{P}_1|\mathcal{P}_0) \\ &\lesssim T^2 \frac{|\log \varepsilon|^\beta}{\varepsilon^2} \int_{-\infty}^{\infty} |\varphi_{T, \mathcal{P}_0}(u-i)|^2 \left| \int_{-\infty}^{\infty} \frac{e^{iux} - 1}{|x|} e^x (k_1 - k_0)(x) dx \right|^2 (u^4 + u^2)^{-1} du \\ &\lesssim \frac{|\log \varepsilon|^\beta}{\varepsilon^2} \delta^2 U^{-2j} \int_{-\infty}^{\infty} |\varphi_{T, \mathcal{P}_0}(u-i)|^2 \left| \int_{-\infty}^{\infty} \frac{e^{iux} - 1}{|x|} e^x g_U(Ux) dx \right|^2 (u^4 + u^2)^{-1} du \\ &\lesssim \frac{|\log \varepsilon|^\beta}{\varepsilon^2} \delta^2 U^{-2j} \int_{-\infty}^{\infty} |u|^{-2T\alpha-4} \left| \int_{-\infty}^{\infty} \frac{e^{iux/U} - 1}{|x|} e^{x/U} g_U(x) dx \right|^2 du \\ &\lesssim \frac{|\log \varepsilon|^\beta}{\varepsilon^2} \delta^2 U^{-2T\alpha-2j-3} \int_{-\infty}^{\infty} |u|^{-2T\alpha-4} \left| \int_{-\infty}^{\infty} \frac{e^{iux} - 1}{|x|} e^{x/U} g_U(x) dx \right|^2 du \\ &\lesssim \varepsilon^{-2} (\log \varepsilon^{-1})^\beta \delta^{(2s+2T\alpha+1)/(s-1-j)}. \end{aligned}$$

Hence, the choice  $\delta \sim (\varepsilon^2 (\log \varepsilon^{-1})^{-\beta})^{(s-1-j)/(2s+2T\alpha+1)}$  yields the claim.

Form this proof we can conclude *Corollary B.6* as follows:

Using the same perturbation  $\mathcal{P}_1$  of a pair  $\mathcal{P}_0 \in \mathcal{G}_s(R, \bar{\alpha})$  we obtain bounds for  $(k_0 - k_1)^{(s-1)}(0+) + (k_0 - k_1)^{(s-1)}(0-)$  and  $\|(1 \vee e^x)(k_1 - k_0)^{(s)}\|_{L^1}$  which depend only on  $\delta$ . Thus, we choose  $U > 1$  and  $\delta$  independently from each other and estimate the Kullback-Leibler-divergence as in the theorem:

$$KL(\mathcal{P}_1|\mathcal{P}_0) \lesssim \varepsilon^{-2} (\log \varepsilon^{-1})^\beta \delta^2 U^{-2s-2T\alpha-1}.$$

Therefore, for a small constant  $\delta$  and  $U \sim (\varepsilon^{-2} (\log \varepsilon^{-1})^\beta)^{1/(2s+2T\alpha+1)}$  the Kullback-Leibler-divergence is bounded.

### C.3. Lower bound for $k_e$ :

Choose some  $\mathcal{P}_0 = (\gamma_0, k_0) \in \mathcal{H}_s(R, \bar{\alpha})$  such that the corresponding characteristic function decreases as  $|u|^{-T\bar{\alpha}}$  but all integral norms and constants are strictly smaller than

R. For  $j \in \mathbb{N}$  let  $\psi_j \in C^\infty(\mathbb{R})$  satisfy  $\text{supp } \psi = [0, 1]$  and  $\|\psi_j\|_{L^2} = 1$  as well as

$$\int_{\mathbb{R}} \psi_j(x) dx = \int_{\mathbb{R}} \psi_j(x) e^{-2^{-j}x} dx = 0 \quad \text{and} \quad \int_{\mathbb{R}} |\mathcal{F}\psi_j(u) u^{-\lambda}|^2 du < \infty$$

for some  $\lambda > T\bar{\alpha} + 2$  and have uniformly bounded norms  $\|\psi_j^{(m)}\|_{L^2} < C$  for a constant  $C > 0$  and all  $j \geq 1, m = 0, \dots, s$ . Such functions exist, for instance the last property holds if  $\psi_j$  is the  $\lambda$ th derivative of an  $L^2$ -function. Defining

$$\psi_{jl}(x) := 2^{j/2} \psi_j(2^j x - l) \quad \text{for } j \geq 1, l = 2^{j-1}, \dots, 2^j - 1,$$

we obtain  $\mathcal{F}\psi_{jl}(0) = \mathcal{F}\psi_{jl}(i) = 0$ . For any  $r = (r_{2^{j-1}}, \dots, r_{2^j-1}) \in \{-1, 1\}^{2^j-1}$  we consider the perturbed pair  $\mathcal{P}_r = (\gamma_0, k_r)$  with

$$k_r(x) - k_0(x) = \delta e^{-x} |x| \sum_{l=2^{j-1}}^{2^j-1} r_l \psi_{jl}(x), \quad x \in \mathbb{R}.$$

Hence,  $k_{e,r}(x) - k_{e,0}(x) = \delta x \sum_{l=2^{j-1}}^{2^j-1} r_l \psi_{jl}(x)$  and  $\mathcal{P}_r$  satisfies the martingale condition and  $(k_r - k_0)(0+) + (k_r - k_0)(0-) = 0$ . Assumption 1 with  $C_2 \leq R$  and  $\mathbb{E}_{\mathcal{P}_r}[X_T e^{X_T}] = -i\varphi'_{T,\mathcal{P}_r}(-i) \leq R$  hold for  $\delta$  sufficiently small. Using the disjoint support of  $\psi_{jl}$  for different  $l$ , we calculate for  $m = 0, \dots, s$ :

$$\begin{aligned} \|k_{e,r}^{(m)} - k_{e,0}^{(m)}\|_{L^2}^2 &= \delta^2 \sum_{l=2^{j-1}}^{2^j-1} \left\| \frac{d^m}{dx^m} x \psi_{jl}(x) \right\|_{L^2}^2 = \delta^2 \sum_{l=2^{j-1}}^{2^j-1} \left\| x \psi_{jl}^{(m)}(x) + m \psi_{jl}^{(m-1)}(x) \right\|_{L^2}^2 \\ &= \delta^2 2^{2jm} \sum_{l=2^{j-1}}^{2^j-1} \left\| 2^{-j}(y+l) \psi_j^{(m)}(y) + m 2^{-j} \psi_j^{(m-1)}(y) \right\|_{L^2}^2. \end{aligned}$$

We note that  $y = 2^j x - l \in [0, 1]$  implies  $x = 2^{-j}(y+l) \in [1/2, 1]$  and estimate

$$\|k_{e,r}^{(m)} - k_{e,0}^{(m)}\|_{L^2}^2 \leq \delta^2 2^{2jm+j-1} (2 \|\psi_j^{(m)}\|_{L^2}^2 + 2m^2 \|\psi_j^{(m-1)}\|_{L^2}^2) \lesssim \delta^2 2^{j(2m+1)}.$$

Since  $\|k_{e,r} - k_{e,0}\|_{L^1} \lesssim \delta 2^{j/2}$  follows in the same way, choosing  $\delta \sim 2^{-j(s+1/2)}$  ensures  $\mathcal{P}_r \in \mathcal{H}_s(R, \bar{\alpha})$ .

Let  $r, r' \in \{-1, 1\}^{2^j-1}$  with Hemming distance equal to one, that is  $r_l = r'_l$  except for one  $l_0$ . Then,  $\psi_{j,l_0}(x) = 0$  for  $|x| < 1/2$  implies

$$\|k_{e,r} - k_{e,r'}\|_{L^2,\tau}^2 = 4\delta^2 \|x \psi_{j,l_0}(x)\|_{L^2}^2 = 4\delta^2 \|2^{-j}(y+l_0) \psi_j(y)\|_{L^2}^2 \geq \delta^2 \|\psi_j\|_{L^2}^2.$$

We will apply Assouad's lemma (see Tsybakov [26, Lem. 2.12]), which yields

$$\inf_{\hat{k}_e} \sup_{\mathcal{P} \in \mathcal{H}_s(R, \bar{\alpha})} \mathbb{E}_{\mathcal{P}}[|\hat{k}_e - k_e|^2] \gtrsim 2^{j-1} \|k_{e,r} - k_{e,r'}\|_{L^2,\tau}^2 \gtrsim 2^{-2js}$$

if the Kullback-Leibler divergence for two alternatives with Hamming distance one remains bounded. This holds true by choosing  $2^{-j} \sim (\varepsilon(\log \varepsilon^{-1})^{-\beta/2})^{2/(2s+2T\bar{\alpha}+5)}$  since

$$\begin{aligned} KL(\mathcal{P}_{r'}|\mathcal{P}_r) &\lesssim \varepsilon^{-2}(\log \varepsilon^{-1})^\beta \int_{-\infty}^{\infty} \left| \varphi_{T,\mathcal{P}_r}(u-i) \mathcal{F}\left(\frac{e^x(k_{r'}-k_r)(x)}{|x|}\right)(u) \right|^2 (u^4+u^2)^{-1} du \\ &\lesssim \varepsilon^{-2}(\log \varepsilon^{-1})^\beta \delta^2 \int_{-\infty}^{\infty} |u|^{-2T\bar{\alpha}-4} |\mathcal{F}\psi_{j_{l_0}}(u)|^2 du \\ &= \varepsilon^{-2}(\log \varepsilon^{-1})^\beta \delta^2 2^{-j} \int_{-\infty}^{\infty} |u|^{-2T\bar{\alpha}-4} |\mathcal{F}\psi_j(2^{-j}u)|^2 du \\ &\lesssim \varepsilon^{-2}(\log \varepsilon^{-1})^\beta 2^{-j(2s+2T\bar{\alpha}+5)}. \end{aligned}$$

#### C.4. Proof of Lemma C.1

Let  $j \geq 0$  and define  $C^{-1}(\mathbb{R}) := \mathbb{R}^{\mathbb{R}}$ . For a constant  $T > 1$  consider the set

$$\begin{aligned} \mathcal{C}_T := \{ f \in C^s(\mathbb{R} \setminus \{0\}) \cap C^{j-1}(\mathbb{R}) \mid \text{supp } f = [-1, 1], f \text{ satisfies condition iii),} \\ \|f^{(l)}\|_\infty < T, l = 0, \dots, s \} \subset L^2(\mathbb{R}). \end{aligned}$$

Certainly,  $\mathcal{C}_T$  is non-empty if  $T$  is greater than a minimal value. Thus, each  $g \in \mathcal{C}_T$  satisfies the properties ii) - iv) for some  $S > T$  big enough. To handle conditions i) and v), we define the functions  $\chi_l^U : [-1, 1] \rightarrow \mathbb{R}, x \mapsto \text{sgn}(x)x^{l-1}e^{x/U}$ , with  $l = 1, \dots, m-1$ , and consider  $g_U \in \mathcal{C}_T$  which satisfies  $g_U \perp \chi_l^U, l = 1, \dots, m-1$ , where we write  $f \perp g$  if  $f, g \in L^2[-1, 1]$  are orthogonal in the sense of  $L^2[-1, 1]$ . Then dominated convergence yields for  $u \in [-1, 1]$

$$\begin{aligned} \left| \int_{-\infty}^{\infty} \frac{e^{iux} - 1}{|x|} e^{x/U} g_U(x) dx \right| &\leq \sum_{k=m}^{\infty} \frac{|u|^k}{k!} \left| \int_{-\infty}^{\infty} \text{sgn}(x)x^{k-1}e^{x/U} g_U(x) dx \right| \\ &\leq |u|^m \int_{-\infty}^{\infty} \sum_{k=m-1}^{\infty} \frac{|x|^k}{(k+1)!} e^{x/U} |g_U(x)| dx \leq |u|^m \int_{-\infty}^{\infty} e^{|x|} (1 \vee e^x) |g_U(x)| dx \leq 2e^2 T |u|^m. \end{aligned}$$

Hence, condition v) is satisfied if  $S \geq 2e^2 T$ . Furthermore, the orthogonality of  $g_U$  and  $\chi_l^U, l = 1, \dots, m-1$ , implies

$$\left\langle \frac{e^{x/U} - 1}{|x|}, g_U \right\rangle = \left\langle \underbrace{- \sum_{k=m}^{\infty} \frac{\text{sgn}(x)x^{k-1}}{k!(-U)^k} e^{x/U}}_{=: (-U)^{-m} \chi_m^U(x)}, g_U \right\rangle.$$

Therefore,  $g_U$  satisfies condition i) if  $g_U \perp \chi_m^U$ . It remains to prove

$$\mathcal{C}_T \cap \mathcal{A}_U^\perp \neq \emptyset \quad \text{where } \mathcal{A}_U := \{\chi_l^U \mid l = 1, \dots, m\}.$$

This can be done by contradiction using a separation argument:

Assume  $\mathcal{C}_T \cap \mathcal{A}_U^\perp = \emptyset$ . Since  $\mathcal{C}_T \neq \emptyset$  and  $\mathcal{A}_U^\perp$  are convex, the Hahn-Banach separation theorem and the Fréchet-Riesz representation theorem imply the existence of a function  $\xi_U$  in the Hilbert space  $(L^2[-1, 1], \langle \cdot, \cdot \rangle)$  such that  $\langle f, \xi_U \rangle \leq \langle g, \xi_U \rangle$  for all  $f \in \mathcal{C}_T, \forall g \in \mathcal{A}_U^\perp$ . Because  $\mathcal{A}_U^\perp \subset L^2[-1, 1]$  is a subspace and the linear functional  $\mathcal{A}_U^\perp \ni g \mapsto \langle g, \xi_U \rangle \in \mathbb{R}$  is bounded from below,  $\langle g, \xi_U \rangle = 0$  holds for all  $g \in \mathcal{A}_U^\perp$ . Since  $\mathcal{A}_U$  is finite, we conclude further

$$\xi_U \in \overline{\text{lin } \mathcal{A}_U} = \text{lin } \mathcal{A}_U \quad \text{and} \quad \langle f, \xi_U \rangle \leq 0, \quad \forall f \in \mathcal{C}_T.$$

This leads to a contradiction if there exists an  $f \in \mathcal{C}_T$  such that  $\langle f, \xi_U \rangle > 0$ . To show this, we define  $\chi_l \in L^2[-1, 1]$  via the pointwise limits  $\chi_l(x) := \lim_{U \rightarrow \infty} \chi_l^U(x)$  for  $x \in \mathbb{R} \setminus \{0\}, l = 1, \dots, m$ . These limits are linearly independent and non-zero. By the compactness of the interval and the uniform boundedness of the functions  $\chi_l^U$  this is also an  $L^2$  limit. Let  $\xi_U = \sum_{l=1}^m a_l \chi_l^U$  for some  $a_l \in \mathbb{R}, l = 1, \dots, m$ . Using the Cauchy-Schwarz inequality, we obtain for all  $f \in \mathcal{C}_T$ :

$$\begin{aligned} \langle f, \xi_U \rangle &= \sum_{l=1}^m a_l (\langle f, \chi_l \rangle - \langle f, \chi_l - \chi_l^U \rangle) \geq \sum_{l=1}^m |a_l| (\langle f, \text{sgn}(a_l) \chi_l \rangle - \|f\|_{L^2} \|\chi_l - \chi_l^U\|_{L^2}) \\ &\geq \sum_{l=1}^m |a_l| (\langle f, \text{sgn}(a_l) \chi_l \rangle - \sqrt{2}T \|\chi_l - \chi_l^U\|_{L^2}) \end{aligned}$$

Thus, for some  $\tau > 0$  we choose  $T$  big enough such that for all  $e = (e_l)_{l=1}^m \in \{-1, 1\}^m$  there exists an  $f_e \in \mathcal{C}_T$  such that  $\min_{l=1, \dots, m} \langle f_e, e_l \chi_l \rangle > \tau$ . (This can be done by choosing  $f_e$  as a polynomial with  $2s + 2$  conditions at  $\pm 1$  and 0 as well as  $m$  linear restrictions on the coefficients of the polynomial since we can calculate  $\langle f_e, e_l \chi_l \rangle$  explicitly.) Choosing a  $U_{min} > 1$  such that  $\tau > \sqrt{2}T \max_{l=1, \dots, m} \|\chi_l^{U_{min}} - \chi_l\|_{L^2}$  ensures  $\langle f_e, \xi_U \rangle > 0$  with  $e = (\text{sgn}(a_l))_{l=1}^m$ .

## References

- [1] BELOMESTNY, D. (2010). Spectral estimation of the fractional order of a Lévy process. *Ann. Statist.* **38** 317-351.
- [2] BELOMESTNY, D. (2011). Statistical inference for time-changed Lévy processes via composite characteristic function estimation. *Ann. Statist.* To appear.
- [3] BELOMESTNY, D. and REISS, M. (2006). Spectral calibration of exponential Lévy models. *Finance Stoch* **10** 449-474.
- [4] BELOMESTNY, D. and SCHOENMAKERS, J. (2011). A jump-diffusion Libor model and its robust calibration. *Quant. Finance* **11** 529-546.
- [5] BROWN, L. D. and LOW, M. G. (1996). Asymptotic equivalence of nonparametric regression and white noise. *Ann. Statist.* **24** 2384-2398.
- [6] CARR, P. and MADAN, D. B. (1999). Option valuation using the fast Fourier transform. *J. Comput. Finance* **2** 61-73.

- [7] CARR, P., GEMAN, H., MADAN, D. B. and YOR, M. (2002). The Fine Structure of Asset Returns: An Empirical Investigation. *J. Bus.* **75** 305-332.
- [8] CARR, P., GEMAN, H., MADAN, D. B. and YOR, M. (2007). Self-decomposability and option pricing. *Math. Finance* **17** 31-57.
- [9] CHERNOZHUKOV, V., FERNÁNDEZ-VAL, I. and GALICHON, A. (2009). Improving point and interval estimators of monotone functions by rearrangement. *Biometrika* **96** 559-575.
- [10] CONT, R. and TANKOV, P. (2004). Non-parametric calibration of jump-diffusion option pricing models. *J. Comput. Finance* **7** 1-49.
- [11] EBERLEIN, E., KELLER, U. and PRAUSE, K. (1998). New Insights Into Smile, Mispricing and Value At Risk: The Hyperbolic Model. *J. Bus.* **71** 371-406.
- [12] EBERLEIN, E. and MADAN, D. B. (2009). Sato processes and the valuation of structured products. *Quant. Finance* **9** 27-42.
- [13] GRAMA, I. and NUSSBAUM, M. (2002). Asymptotic equivalence for nonparametric regression. *Math. Methods Statist.* **11** 1-36.
- [14] HOUDRÉ, C. and REYNAUD-BOURET, P. (2003). Exponential inequalities, with constants, for U-statistics of order two. In *Stochastic inequalities and applications.*, (Evariste Giné and Christian Houdré and David Nualart, ed.). *Progress in Probability* 55-69. Birkhäuser.
- [15] JONGBLOED, G., VAN DER MEULEN, F. H. and VAN DER VAART, A. W. (2005). Nonparametric inference for Lvy-driven Ornstein-Uhlenbeck processes. *Bernoulli* **11** 759-791.
- [16] KAPPUS, J. and REISS, M. (2010). Estimation of the characteristics of a Lévy process observed at arbitrary frequency. *Stat. Neerl.* **64** 314-328.
- [17] LIPTSER, R. S. and SHIRYAEV, A. N. (2001). *Statistics of Random Processes* **1**. Springer.
- [18] MADAN, D. B., CARR, P. P. and CHANG, E. C. (1998). The Variance Gamma Process and Option Pricing. *Europ. Finance Rev.* **2** 79-105.
- [19] MADAN, D. B. and SENETA, E. (1990). The Variance Gamma (VG) Model for Share Market Returns. *J. Bus.* **63** 511-524.
- [20] MERTON, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *J. Finan. Econ.* **3** 125 - 144.
- [21] SATO, K.-I. (1991). Self-similar processes with independent increments. *Probab. Theory Related Fields* **89** 285-300.
- [22] SATO, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.
- [23] SÖHL, J. (2010). Polar sets for anisotropic Gaussian random fields. *Statist. Probab. Lett.* **80** 840 - 847.
- [24] SÖHL, J. (2011). Confidence sets in nonparametric calibration of exponential Lévy models SFB 649 Discussion Paper report, Sonderforschungsbereich 649, Humboldt Universität zu Berlin, Germany. To appear.
- [25] TANKOV, P. (2011). Pricing and Hedging in Exponential Lévy Models: Review of Recent Results. In *Paris-Princeton Lectures on Mathematical Finance 2010. Lecture Notes in Mathematics* **2003** 319-359. Springer Berlin / Heidelberg.

- [26] TSYBAKOV, A. B. (2009). *Introduction to Nonparametric Estimation*. Springer.
- [27] VAN DE GEER, S. (2000). *Empirical processes in M-estimation*. Cambridge University Press.

# SFB 649 Discussion Paper Series 2011

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "Localising temperature risk" by Wolfgang Karl Härdle, Brenda López Cabrera, Ostap Okhrin and Weining Wang, January 2011.
- 002 "A Confidence Corridor for Sparse Longitudinal Data Curves" by Shuzhuan Zheng, Lijian Yang and Wolfgang Karl Härdle, January 2011.
- 003 "Mean Volatility Regressions" by Lu Lin, Feng Li, Lixing Zhu and Wolfgang Karl Härdle, January 2011.
- 004 "A Confidence Corridor for Expectile Functions" by Esra Akdeniz Duran, Mengmeng Guo and Wolfgang Karl Härdle, January 2011.
- 005 "Local Quantile Regression" by Wolfgang Karl Härdle, Vladimir Spokoiny and Weining Wang, January 2011.
- 006 "Sticky Information and Determinacy" by Alexander Meyer-Gohde, January 2011.
- 007 "Mean-Variance Cointegration and the Expectations Hypothesis" by Till Strohsal and Enzo Weber, February 2011.
- 008 "Monetary Policy, Trend Inflation and Inflation Persistence" by Fang Yao, February 2011.
- 009 "Exclusion in the All-Pay Auction: An Experimental Investigation" by Dietmar Fehr and Julia Schmid, February 2011.
- 010 "Unwillingness to Pay for Privacy: A Field Experiment" by Alastair R. Beresford, Dorothea Kübler and Sören Preibusch, February 2011.
- 011 "Human Capital Formation on Skill-Specific Labor Markets" by Runli Xie, February 2011.
- 012 "A strategic mediator who is biased into the same direction as the expert can improve information transmission" by Lydia Mechtenberg and Johannes Münster, March 2011.
- 013 "Spatial Risk Premium on Weather Derivatives and Hedging Weather Exposure in Electricity" by Wolfgang Karl Härdle and Maria Osipenko, March 2011.
- 014 "Difference based Ridge and Liu type Estimators in Semiparametric Regression Models" by Esra Akdeniz Duran, Wolfgang Karl Härdle and Maria Osipenko, March 2011.
- 015 "Short-Term Herding of Institutional Traders: New Evidence from the German Stock Market" by Stephanie Kremer and Dieter Nautz, March 2011.
- 016 "Oracally Efficient Two-Step Estimation of Generalized Additive Model" by Rong Liu, Lijian Yang and Wolfgang Karl Härdle, March 2011.
- 017 "The Law of Attraction: Bilateral Search and Horizontal Heterogeneity" by Dirk Hofmann and Salmai Qari, March 2011.
- 018 "Can crop yield risk be globally diversified?" by Xiaoliang Liu, Wei Xu and Martin Odening, March 2011.
- 019 "What Drives the Relationship Between Inflation and Price Dispersion? Market Power vs. Price Rigidity" by Sascha Becker, March 2011.
- 020 "How Computational Statistics Became the Backbone of Modern Data Science" by James E. Gentle, Wolfgang Härdle and Yuichi Mori, May 2011.
- 021 "Customer Reactions in Out-of-Stock Situations – Do promotion-induced phantom positions alleviate the similarity substitution hypothesis?" by Jana Luisa Diels and Nicole Wiebach, May 2011.

**SFB 649, Spandauer Str. 1, D-10178 Berlin**  
**<http://sfb649.wiwi.hu-berlin.de>**

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".



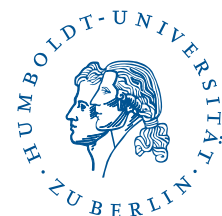
# SFB 649 Discussion Paper Series 2011

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 022 "Extreme value models in a conditional duration intensity framework" by Rodrigo Herrera and Bernhard Schipp, May 2011.
- 023 "Forecasting Corporate Distress in the Asian and Pacific Region" by Russ Moro, Wolfgang Härdle, Saeideh Aliakbari and Linda Hoffmann, May 2011.
- 024 "Identifying the Effect of Temporal Work Flexibility on Parental Time with Children" by Juliane Scheffel, May 2011.
- 025 "How do Unusual Working Schedules Affect Social Life?" by Juliane Scheffel, May 2011.
- 026 "Compensation of Unusual Working Schedules" by Juliane Scheffel, May 2011.
- 027 "Estimation of the characteristics of a Lévy process observed at arbitrary frequency" by Johanna Kappus and Markus Reiß, May 2011.
- 028 "Asymptotic equivalence and sufficiency for volatility estimation under microstructure noise" by Markus Reiß, May 2011.
- 029 "Pointwise adaptive estimation for quantile regression" by Markus Reiß, Yves Rozenholc and Charles A. Cuenod, May 2011.
- 030 "Developing web-based tools for the teaching of statistics: Our Wikis and the German Wikipedia" by Sigbert Klinke, May 2011.
- 031 "What Explains the German Labor Market Miracle in the Great Recession?" by Michael C. Burda and Jennifer Hunt, June 2011.
- 032 "The information content of central bank interest rate projections: Evidence from New Zealand" by Gunda-Alexandra Detmers and Dieter Nautz, June 2011.
- 033 "Asymptotics of Asynchronicity" by Markus Bibinger, June 2011.
- 034 "An estimator for the quadratic covariation of asynchronously observed Itô processes with noise: Asymptotic distribution theory" by Markus Bibinger, June 2011.
- 035 "The economics of TARGET2 balances" by Ulrich Bindseil and Philipp Johann König, June 2011.
- 036 "An Indicator for National Systems of Innovation - Methodology and Application to 17 Industrialized Countries" by Heike Belitz, Marius Clemens, Christian von Hirschhausen, Jens Schmidt-Ehmcke, Axel Werwatz and Petra Zloczynski, June 2011.
- 037 "Neurobiology of value integration: When value impacts valuation" by Soyoung Q. Park, Thorsten Kahnt, Jörg Rieskamp and Hauke R. Heekeren, June 2011.
- 038 "The Neural Basis of Following Advice" by Guido Biele, Jörg Rieskamp, Lea K. Krugel and Hauke R. Heekeren, June 2011.
- 039 "The Persistence of "Bad" Precedents and the Need for Communication: A Coordination Experiment" by Dietmar Fehr, June 2011.
- 040 "News-driven Business Cycles in SVARs" by Patrick Bunk, July 2011.
- 041 "The Basel III framework for liquidity standards and monetary policy implementation" by Ulrich Bindseil and Jeroen Lamoot, July 2011.
- 042 "Pollution permits, Strategic Trading and Dynamic Technology Adoption" by Santiago Moreno-Bromberg and Luca Taschini, July 2011.
- 043 "CRRA Utility Maximization under Risk Constraints" by Santiago Moreno-Bromberg, Traian A. Pirvu and Anthony Réveillac, July 2011.

**SFB 649, Spandauer Str. 1, D-10178 Berlin**  
**<http://sfb649.wiwi.hu-berlin.de>**

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".





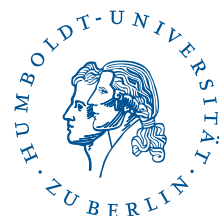
# SFB 649 Discussion Paper Series 2011

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 044 "Predicting Bid-Ask Spreads Using Long Memory Autoregressive Conditional Poisson Models" by Axel Groß-Klußmann and Nikolaus Hautsch, July 2011.
- 045 "Bayesian Networks and Sex-related Homicides" by Stephan Stahlschmidt, Helmut Tausendteufel and Wolfgang K. Härdle, July 2011.
- 046 "The Regulation of Interdependent Markets", by Raffaele Fiocco and Carlo Scarpa, July 2011.
- 047 "Bargaining and Collusion in a Regulatory Model", by Raffaele Fiocco and Mario Gilli, July 2011.
- 048 "Large Vector Auto Regressions", by Song Song and Peter J. Bickel, August 2011.
- 049 "Monetary Policy, Determinacy, and the Natural Rate Hypothesis", by Alexander Meyer-Gohde, August 2011.
- 050 "The impact of context and promotion on consumer responses and preferences in out-of-stock situations", by Nicole Wiebach and Jana L. Diels, August 2011.
- 051 "A Network Model of Financial System Resilience", by Kartik Anand, Prasanna Gai, Sujit Kapadia, Simon Brennan and Matthew Willison, August 2011.
- 052 "Rollover risk, network structure and systemic financial crises", by Kartik Anand, Prasanna Gai and Matteo Marsili, August 2011.
- 053 "When to Cross the Spread: Curve Following with Singular Control" by Felix Naujokat and Ulrich Horst, August 2011.
- 054 "TVICA - Time Varying Independent Component Analysis and Its Application to Financial Data" by Ray-Bing Chen, Ying Chen and Wolfgang K. Härdle, August 2011.
- 055 "Pricing Chinese rain: a multi-site multi-period equilibrium pricing model for rainfall derivatives" by Wolfgang K. Härdle and Maria Osipenko, August 2011.
- 056 "Limit Order Flow, Market Impact and Optimal Order Sizes: Evidence from NASDAQ TotalView-ITCH Data" by Nikolaus Hautsch and Ruihong Huang, August 2011.
- 057 "Optimal Display of Iceberg Orders" by Gökhan Cebiroğlu and Ulrich Horst, August 2011.
- 058 "Optimal liquidation in dark pools" by Peter Kratz and Torsten Schöneborn, September 2011.
- 059 "The Merit of High-Frequency Data in Portfolio Allocation" by Nikolaus Hautsch, Lada M. Kyj and Peter Malec, September 2011.
- 060 "On the Continuation of the Great Moderation: New evidence from G7 Countries" by Wenjuan Chen, September 2011.
- 061 "Forward-backward systems for expected utility maximization" by Ulrich Horst, Ying Hu, Peter Imkeller, Anthony Réveillac and Jianing Zhang.
- 062 "On heterogeneous latent class models with applications to the analysis of rating scores" by Aurélie Bertrand and Christian M. Hafner, October 2011.
- 063 "Multivariate Volatility Modeling of Electricity Futures" by Luc Bauwens, Christian Hafner and Diane Pierret, October 2011.

**SFB 649, Spandauer Str. 1, D-10178 Berlin**  
**<http://sfb649.wiwi.hu-berlin.de>**

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".



# SFB 649 Discussion Paper Series 2011

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 064 "Semiparametric Estimation with Generated Covariates" by Enno Mammen, Christoph Rothe and Melanie Schienle, October 2011.
- 065 "Linking corporate reputation and shareholder value using the publication of reputation rankings" by Sven Tischer and Lutz Hildebrandt, October 2011.
- 066 "Monitoring, Information Technology and the Labor Share" by Dorothee Schneider, October 2011.
- 067 "Minimal Supersolutions of BSDEs with Lower Semicontinuous Generators" by Gregor Heyne, Michael Kupper and Christoph Mainberger, October 2011.
- 068 "Bargaining, Openness, and the Labor Share" by Dorothee Schneider, October 2011.
- 069 "The Labor Share: A Review of Theory and Evidence" by Dorothee Schneider, October 2011.
- 070 "The Power of Sunspots: An Experimental Analysis" by Dietmar Fehr, Frank Heinemann and Aniol Llorente-Saguer, October 2011.
- 071 "Econometric analysis of volatile art markets" by Fabian Y. R. P. Bocart and Christian M. Hafner, October 2011.
- 072 "Financial Network Systemic Risk Contributions" by Nikolaus Hautsch, Julia Schaumburg and Melanie Schienle, October 2011.
- 073 "Calibration of self-decomposable Lévy models" by Mathias Trabs, November 2011.

**SFB 649, Spandauer Str. 1, D-10178 Berlin**  
**<http://sfb649.wiwi.hu-berlin.de>**

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

