# Weierstraß-Institut
## für Angewandte Analysis und Stochastik
### Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

# Statistical Skorohod embedding problem and its generalizations

Denis Belomestny,[1] John G. M. Schoenmakers[2]

submitted: June 5, 2014

[1] Mathematics Department,
Duisburg-Essen University,
Forsthausweg 2
47057 Duisburg, Germany,
E-Mail: denis.belomestny@uni-due.de

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: John.Schoenmakers@wias-berlin.de

No. 1960
Berlin 2014

2010 *Mathematics Subject Classification.* 62F10, 62F25.

*Key words and phrases.* Skorohod embedding problem, Lévy process, Mellin transform, Laplace transform, variance mixture models, time-changed Lévy processes.

Partially supported by the Deutsche Forschungsgemeinschaft through SFB 823 "Statistical modeling of nonlinear dynamic processes", the National Research University Higher School of Economics and by the Laboratory for Structural Methods of Data Analysis in Predictive Modeling, MIPT, RF government grant, ag. 11.G34.31.0073.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

**Abstract**

Given a Lévy process $L$, we consider the so-called statistical Skorohod embedding problem of recovering the distribution of an independent random time $T$ based on i.i.d. sample from $L_T$. Our approach is based on the genuine use of the Mellin and Laplace transforms. We propose consistent estimators for the density of $T$, derive their convergence rates and prove their optimality. It turns out that the convergence rates heavily depend on the decay of the Mellin transform of $T$. We also consider the application of our results to the problem of statistical inference for variance-mean mixture models and for time-changed Lévy processes.

# 1 Introduction

The so called Skorohod embedding (SE) problem or Skorohod stopping problem was first stated and solved by Skorohod in 1961. This problem can be formulated as follows.

**Problem 1.1** (Skorohod Embedding Problem)**.** *For a given probability measure $\mu$ on $\mathbb{R}$, such that $\int |x|d\mu(x) < \infty$ and $\int x d\mu(x) = 0$, find a stopping time $T$ such that $B_T \sim \mu$ and $B_{T\wedge t}$ is a uniformly integrable martingale.*

The SE problem has recently drawn much attention in the literature, see e.g. Obłój, [8], where the list of references consists of more than 100 items. In fact, there is no unique solution to the SE problem and there are currently more than $20$ different solutions available. This means that from a statistical point of view, the SE problem is not well posed. In this paper we first study what we call *statistical Skorohod embedding* (SSE) problem.

**Problem 1.2** (Statistical Skorohod Embedding Problem)**.** *Based on i.i.d. sample $X_1, \ldots, X_n$ from the distribution of $B_T$ consistently estimate the distribution of the random time $T \geq 0$, where $B$ and $T$ are assumed to be independent.*

The independence of $B$ and $T$ is needed to ensure the identifiability of the distribution of $T$ from the distribution of $B_T$. It is shown that the SSE problem is closely related to the multiplicative deconvolution problem. Using the Mellin transform technique, we construct a consistent estimator for the density of $T$ and derive its convergence rates in different norms. Furthermore, we show that the obtained rates are optimal in minimax sense. Next, we generalize the SSE problem by replacing the standard Brownian motion with a general Lévy process. The generalized SSE problem turns out to be much more involved and its solution requires some new ideas. Using a genuine combination of the Laplace and Mellin transforms, we construct a consistent estimator, derive its minimax convergence rates and prove that these rates basically coincide with the rates in the SSE problem.

Some particular cases of generalized statistical Skorohod embedding problem have been already studied in the literature. For example, the case of the stopped Poisson process was considered in the recent paper of Comte and Genon-Catalot, [5].

# 2 Statistical Skorohod embedding problem

Let $B$ be a Brownian motion and let a random variable $T \geq 0$ be independent of $B$. We then have,

$$X := B_T \sim \sqrt{T}\, B_1 \tag{1}$$

and the problem of reconstructing $T$ is related to a multiplicative deconvolution problem. While for additive deconvolution problems the Fourier transform plays an important role, here we can conveniently use the Mellin transform.

**Definition 2.1.** Let $\xi$ be a non-negative random variable with a probability density $p_\xi$, then the *Mellin transform* of $p_\xi$ is defined via

$$\mathcal{M}[p_\xi](z) := \mathbb{E}[\xi^{z-1}] = \int_0^\infty p_\xi(x) x^{z-1}\, dx \tag{2}$$

for all $z \in \mathcal{S}_\xi$ with $\mathcal{S}_\xi = \{z \in \mathbb{C} : \mathbb{E}[\xi^{\mathrm{Re} z - 1}] < \infty\}$.

Since $p_\xi$ is a density it is integrable and so at least $\{z \in \mathbb{C} : \mathrm{Re}(z) = 1\} \subset \mathcal{S}_\xi$. Under mild assumptions on the growth of $p_\xi$ near the origin one obtains

$$\{z \in \mathbb{C} : 0 \leq a_\xi < \mathrm{Re}(z) < b_\xi\} \subset \mathcal{S}_\xi$$

for some $0 \leq a_\xi < 1 \leq b_\xi$. Then the Mellin transform (2) exists and is analytic in the strip $a_\xi < \mathrm{Re}\, z < b_\xi$. For example, if $p_\xi$ is essentially bounded in a right-hand neighborhood of zero, we may take $a_\xi = 0$. The role of the Mellin transform in probability theory is mainly related to the product of independent random variables: in fact it is well-known that the probability density of the product of two independent random variables is given by the Mellin convolution of the two corresponding densities. Due to (1), the SSE problem is closely connected to the Mellin convolution. Suppose that the random time $T$ has a density $p_T$ and that we may take $0 \leq a_T < 1 \leq b_T$. Since $\mathcal{S}_{|B_1|} \supset \{z \in \mathbb{C} : \mathrm{Re}(z) > 0\}$, we derive for $\max(2a_T - 1, 0) < \mathrm{Re}(z) < 2b_T - 1$,

$$\begin{aligned}\mathcal{M}[p_{|X|}](z) &= \mathbb{E}\big[|B_1|^{z-1}\big]\mathbb{E}\big[T^{(z-1)/2}\big] \\ &= \mathcal{M}[p_{|B_1|}](z)\mathcal{M}[p_T]((z+1)/2) = \frac{2^{(z-1)/2}}{\sqrt{\pi}}\Gamma(z/2)\mathcal{M}[p_T]((z+1)/2).\end{aligned}$$

As a result

$$\mathcal{M}[p_T](z) = \frac{\sqrt{\pi}}{2^{z-1}} \frac{\mathcal{M}[p_{|X|}](2z-1)}{\Gamma(z-1/2)}, \quad \max(a_T, 1/2) < \mathrm{Re}(z) < b_T$$

and the Mellin inversion formula yields

$$\begin{aligned} p_T(x) &= \frac{1}{2\pi} \int_{\gamma-\mathrm{i}\infty}^{\gamma+\mathrm{i}\infty} x^{-\gamma-\mathrm{i}v} \mathcal{M}[p_T](\gamma+\mathrm{i}v)\, dv \\ &= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^{-\gamma-\mathrm{i}v} \frac{\mathcal{M}[p_{|X|}](2(\gamma+\mathrm{i}v)-1)}{2^{\gamma+\mathrm{i}v}\Gamma(\gamma+\mathrm{i}v-1/2)}\, dv \quad \text{for} \quad \max(a_T,1/2) < \gamma < b_T, \quad x > 0. \end{aligned} \tag{3}$$

Furthermore, the Mellin transform of $p_{|X|}$ can be directly estimated from the data $X_1, \ldots, X_n$ via the empirical Mellin transform:

$$\mathcal{M}_n[p_{|X|}](z) := \frac{1}{n} \sum_{k=1}^{n} |X_k|^{z-1}, \quad \mathrm{Re}(z) > 1/2, \tag{4}$$

where the condition $\mathrm{Re}(z) > 1/2$ guarantees that the variance of the estimator (4) is finite. Note however that the integral in (3) may fail to exist if we replace $\mathcal{M}[p_{|X|}]$ by $\mathcal{M}_n[p_{|X|}]$. We so need to regularize the inverse Mellin operator. To this end, let us consider a kernel $K(\cdot) \geq 0$ supported on $[-1,1]$ and a sequence of bandwidths $h_n > 0$ tending to $0$ as $n \to \infty$. Then we define, in view of (4), for some $\max(a_T, 3/4) < \gamma < b_T$,

$$p_{T,n}(x) := \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} x^{-\gamma-\mathrm{i}v} K(vh_n) \frac{\mathcal{M}_n[p_{|X|}](2(\gamma+\mathrm{i}v)-1)}{2^{\gamma+\mathrm{i}v}\Gamma(\gamma-1/2+\mathrm{i}v)}\, dv. \tag{5}$$

For our convergence analysis, we will henceforth take the simplest kernel

$$K(y) = 1_{[-1,1]}(y),$$

but note that in principle other kernels may be considered as well. The next theorem states that $p_{T,n}$ converges to $p_T$ at a polynomial rate, provided the Mellin transform of $p_T$ decays exponentially fast.

**Theorem 2.2.** *For any $\beta > 0$, $\gamma > 0$ and $L > 0$, introduce the class of functions*

$$\mathcal{C}(\beta,\gamma,L) = \left\{ f : \int_{-\infty}^{\infty} |\mathcal{M}[f](\gamma+\mathrm{i}v)|\, e^{\beta|v|}\, dv < L \right\}.$$

*Assume that $p_T \in \mathcal{C}(\beta,\gamma,L)$ for some $\beta > 0$ and $L > 0$, and*

$$\max((a_T+1)/2, 3/4) < \gamma < b_T. \tag{6}$$

*Then for some constant $C_{\gamma,L}$ depending on $\gamma$ and $L$ only, it holds*

$$\sup_{x\geq 0} \mathbb{E}\Big[ \{x^{\gamma}|p_T(x) - p_{T,n}(x)|\}^2 \Big] \leq C_{\gamma,L} \times \begin{cases} e^{-2\beta/h_n} + \frac{1}{n} h_n^{2(\gamma-1)} e^{\pi/h_n}, & \gamma < 1, \\ e^{-2\beta/h_n} + \frac{1}{n} e^{\pi/h_n}, & \gamma \geq 1. \end{cases} \tag{7}$$

*By next choosing*

$$h_n = \begin{cases} \frac{\pi+2\beta}{\log n - 2(1-\gamma)\log\log n}, & \gamma < 1, \\ (\pi + 2\beta)/\log n, & \gamma \geq 1, \end{cases} \tag{8}$$

*we arrive at the rate*

$$\sup_{x\geq 0} \sqrt{\mathbb{E}\Big[\big\{x^\gamma |p_T(x) - p_{T,n}(x)|\big\}^2\Big]} \lesssim \begin{cases} n^{-\frac{\beta}{\pi+2\beta}} \log^{\frac{2(1-\gamma)\beta}{\pi+2\beta}} n, & \gamma < 1, \\ n^{-\frac{\beta}{\pi+2\beta}}, & \gamma \geq 1 \end{cases} \tag{9}$$

*as* $n \to \infty$.

With a little bit more effort one can prove the strong uniform convergence of the estimate $p_{n,T}$.

**Theorem 2.3.** *Under conditions of Theorem 2.2 and for* $\gamma < 1$

$$\sup_{p_T \in \mathcal{C}(\beta,\gamma,L)} \sup_{x\geq 0} \big\{x^\gamma |p_{T,n}(x) - p_T(x)|\big\} = O_{a.s.}\Big(n^{-\frac{\beta}{\pi+2\beta}} \log^{\frac{2(1-\gamma)\beta}{\pi+2\beta}} n\Big).$$

Let us turn now to some examples.

*Example* 2.4. Consider the class of gamma densities

$$p_T(x;\alpha) = \frac{x^{\alpha-1}\cdot e^{-x}}{\Gamma(\alpha)}, \quad x \geq 0$$

for $\alpha > 0$. Since

$$\mathcal{M}[p_T](z) = \frac{\Gamma(z+\alpha-1)}{\Gamma(\alpha)}, \quad \mathrm{Re}(z) > 0,$$

we derive that $p_T \in \mathcal{C}(\beta,\gamma,L)$ for all $0 < \beta < \pi/2$ and $\gamma > 0$ due to the asymptotic properties of the Gamma function (see Lemma 7.3 in Appendix). As a result, Theorem 2.2 implies

$$\sup_{x\geq 0} \mathbb{E}\Big[\big\{x^\gamma |p_T(x) - p_{T,n}(x)|\big\}^2\Big] \lesssim n^{-\rho}, \quad n \to \infty$$

for any $\rho < 1/2$, provided $\gamma \geq 1$.

*Example* 2.5. Let us look at the family of densities

$$p_T(x;q) = \frac{q\sin(\pi/q)}{\pi}\frac{1}{1+x^q}, \quad q \geq 2, \quad x \geq 0.$$

We have

$$\mathcal{M}[p_T](z) = \frac{\sin(\pi/q)}{\sin(\pi z/q)}, \quad 0 < \mathrm{Re}(z) < q.$$

Therefore $p_T \in \mathcal{C}(\beta,\gamma,L)$ for all $0 < \beta < \pi/q$ and $\gamma > 0$, and

$$\sup_{x\geq 0} \mathbb{E}\Big[\big\{x^\gamma |p_T(x) - p_{T,n}(x)|\big\}^2\Big] \lesssim n^{-\rho}, \quad n \to \infty$$

for any $\rho < 1/(1+q/2)$, provided $\gamma \geq 1$.

**Theorem 2.6.** *Consider the class of functions*

$$\mathcal{D}(\beta,\gamma,L)=\left\{f:\int_{-\infty}^{\infty}|\mathcal{M}[f](\gamma+\mathrm{i}v)|\,(1+|v|^{\beta})\,dv<L\right\},$$

*and assume that $p_T\in\mathcal{D}(\beta,\gamma,L)$ for some $\beta>0$ and $L>0$ and $\gamma$ as in (6). Then for some constant $D_{\gamma,L}$, it holds*

$$\sup_{x\geq 0}\mathbb{E}\left[\left\{x^{\gamma}|p_T(x)-p_{T,n}(x)|\right\}^2\right]\leq D_{\gamma,L}\times\begin{cases}h_n^{2\beta}+\frac{1}{n}h_n^{2(\gamma-1)}e^{\pi/h_n}, & \gamma<1,\\ h_n^{2\beta}+\frac{1}{n}e^{\pi/h_n}, & \gamma\geq 1.\end{cases}\tag{10}$$

*By choosing*

$$h_n=\frac{\pi}{\log n-2\,(\beta+1-\gamma)\log\log n},\tag{11}$$

*if $\gamma<1$ and*

$$h_n=\frac{\pi}{\log n-2\beta\log\log n}\tag{12}$$

*for $\gamma\geq 1$, we arrive at*

$$\sup_{x\geq 0}\sqrt{\mathbb{E}\left[\left\{x^{\gamma}|p_T(x)-p_{T,n}(x)|\right\}^2\right]}\lesssim\log^{-\beta}(n),\quad n\to\infty.\tag{13}$$

*Remark* 2.7. Due to the relation

$$\mathcal{M}[p_T](\gamma+\mathrm{i}v)=\mathcal{F}[e^{\gamma\cdot}p_T(e^{\cdot})](v),\quad a_T<\gamma<b_T,$$

the conditions $p_T\in\mathcal{C}(\beta,\gamma,L)$ and $p_T\in\mathcal{D}(\beta,\gamma,L)$ are closely related to the smoothness properties of the function $e^{\gamma x}p_T(e^x)$. For example, if $p_T\in\mathcal{C}(\beta,\gamma,L)$, then

$$\int_{-\infty}^{\infty}|\mathcal{F}[e^{\gamma\cdot}p_T(e^{\cdot})](v)|\,e^{\beta|v|}\,dv<L$$

and the function $e^{\gamma x}p_T(e^x)$ is called supersmooth in this case, see Meister [7] for the discussion on different smoothness classes in the context of the additive deconvolution problems.

The rates of Theorem 2.2 and Theorem 2.6 summarized in Table 1 are in fact optimal (up to a logarithmic factor) in minimax sense for the classes $\mathcal{C}(\beta,\gamma,L)$ and $\mathcal{D}(\beta,\gamma,L)$, respectively.

**Theorem 2.8.** *Fix some $\beta>1$. There are $\varepsilon>0$ and $x>0$ such that*

$$\liminf_{n\to\infty}\inf_{p_n}\sup_{p_T\in\mathcal{C}(\beta,\gamma,L)}\mathrm{P}_{p_T}^{\otimes n}\left(|p_T(x)-p_n(x)|\geq\varepsilon\,n^{-\frac{\beta}{\pi+2\beta}}\log^{-\rho}(n)\right)>0,$$

$$\liminf_{n\to\infty}\inf_{p_n}\sup_{p_T\in\mathcal{D}(\beta,\gamma,L)}\mathrm{P}_{p_T}^{\otimes n}\left(|p_T(x)-p_n(x)|\geq\varepsilon\log^{-\beta}(n)\right)>0,$$

*for some $\rho>0$, where the infimum is taken over all estimators (i.e. all measurable functions of $X_1,\ldots,X_n$) of $p_T$ and $\mathrm{P}_{p_T}^{\otimes n}$ is the distribution of the i.i.d. sample $X_1,\ldots,X_n$ with $X_1\sim W_T$ and $T\sim p_T$.*

| $\mathcal{C}(\beta,\gamma,L)$ | | $\mathcal{D}(\beta,\gamma,L)$ |
|---|---|---|
| $\gamma < 1$ | $\gamma \geq 1$ | |
| $n^{-\frac{\beta}{\pi+2\beta}} \log^{\frac{2(1-\gamma)\beta}{\pi+2\beta}}(n)$ | $n^{-\frac{\beta}{\pi+2\beta}}$ | $\log^{-\beta}(n)$ |

Table 1: Minimax rates of convergence for the classes $\mathcal{C}(\beta,\gamma,L)$ and $\mathcal{D}(\beta,\gamma,L)$.

# 3 Generalised statistical Skorohod embedding problem

In this section we generalize the statistical Skorohod embedding problem to the case of Lévy processes. In particular, we consider the following problem.

**Problem 3.1.** *Based on i.i.d. sample $X_1,\ldots,X_n$ from the distribution of $\mu$ estimate the distribution of the random time $T \geq 0$ independent of a Lévy process $L$ such that $L_T \sim \mu$.*

Note that the situation here is much more difficult than before, since the Lévy processes do not have, in general, the scaling property (1). Hence the approach based on the Mellin deconvolution technique can not be applied any longer. Let $(L_t,\, t \geq 0)$ be a Lévy process with the triplet $(\mu,\sigma^2,\nu)$. Define a curve in $\mathbb{C}$

$$\ell := \Big\{\mathrm{Re}(\psi(u)) + \mathrm{i}\,\mathrm{Im}(\psi(u)),\, u \in \mathbb{R}_+\Big\},$$

where $\psi(u) = -t^{-1}\log(\mathbb{E}(\exp(\mathrm{i}uL_t)))$. Our approach to reconstruct the distribution of $T$ is based on the simple identity

$$\mathcal{F}[p_X](\lambda) = \mathbb{E}[\exp(\mathrm{i}\lambda L_T)] = \mathcal{L}[p_T](\psi(\lambda)). \tag{14}$$

It is well known that the Laplace transform of $\mathcal{L}[p_T](u)$ is analytic in the domain $\big\{\mathrm{Re}(u) > 0\big\}$. The following proposition shows that the object $\mathcal{M}[\mathcal{L}[p_T]](z)$ is well defined and that it can be related to the Fourier transform of $p_X$, which in turn can be estimated from the data.

**Proposition 3.2.** *Let us assume that $\mathrm{Re}(\psi(u)) \to \infty$ as $u \to \infty$ and that*

$$\frac{|\mathrm{Im}(\psi(u))|}{\mathrm{Re}(\psi(u))} < A < \infty \tag{15}$$

*for all $u > 0$ and some $A > 0$. Moreover, let $p_T$ be (essentially) bounded. Then, for $0 < \mathrm{Re}(z) < 1$ it holds that*

$$\mathcal{M}[\mathcal{L}[p_T]](z) = \int_0^\infty u^{z-1}\mathcal{L}[p_T](u)du = \int_\ell w^{z-1}\mathcal{L}[p_T](w)dw.$$

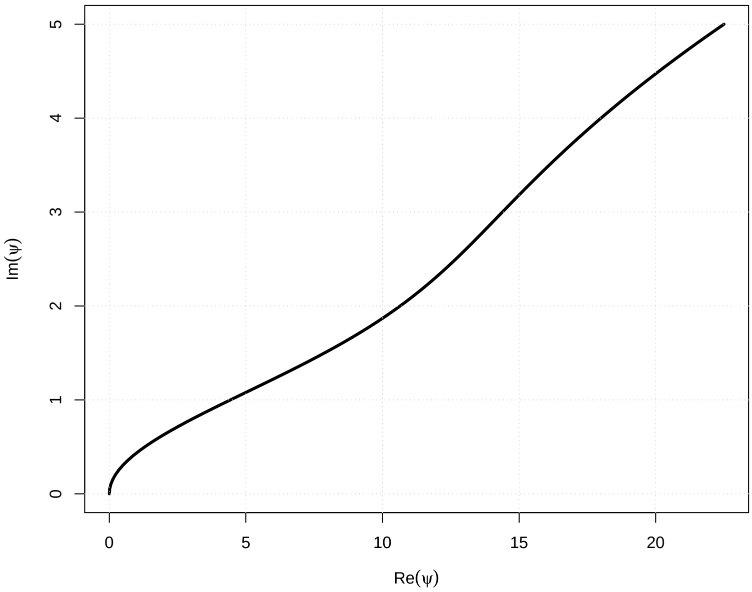



Figure 1: A typical shape of the contour $\ell$.

*Remark* 3.3. The condition (15) is fulfilled if, for example, the diffusion part of $L$ is nonzero or if $\psi$ is real and $\psi(u) \to \infty$ as $u \to \infty$.

Under the assumptions of Proposition 3.2 we may write,

$$\mathcal{M}[\mathcal{L}[p_T]](z) = \int_0^\infty [\psi(\lambda)]^{z-1} \mathcal{L}[p_T](\psi(\lambda))\psi'(\lambda)d\lambda,$$

where $\mathcal{L}[p_T](\psi(\lambda)) = \mathcal{F}[p_X](\lambda)$ due to (14). On other hand, one may straightforwardly derive,

$$\mathcal{M}[\mathcal{L}[p_T]](z) = \mathcal{M}[p_T](1-z)\Gamma(z), \quad 0 < \mathrm{Re}(z) < 1,$$

i.e.,

$$\mathcal{M}[p_T](z) = \frac{\mathcal{M}[\mathcal{L}[p_T]](1-z)}{\Gamma(1-z)} = \frac{\int_0^\infty [\psi(\lambda)]^{-z} \mathcal{F}[p_X](\lambda)\psi'(\lambda)d\lambda}{\Gamma(1-z)}, \quad 0 < \mathrm{Re}(z) < 1. \tag{16}$$

In principle one can now replace the Fourier transform of $p_X$ in (16) by its empirical counterpart based on the data. However, in this case we need to regularize the estimate of $\mathcal{M}[p_T](z)$ to perform the inverse Mellin transform. To this end consider the approximation

$$\mathcal{M}[\mathcal{L}[p_T]](z) \approx \frac{1}{n}\sum_{k=1}^n \int_0^{A_n} [\psi(\lambda)]^{z-1} e^{\mathrm{i}X_k\lambda}\psi'(\lambda)d\lambda =: \frac{1}{n}\sum_{k=1}^n \Phi_n(z, X_k)$$

and define in view of (16),

$$p_{T,n}(x) := \frac{1}{2\pi n}\sum_{k=1}^{n}\int_{-U_n}^{U_n}\frac{\Phi_n(1-\gamma-\mathrm{i}v, X_k)}{\Gamma(1-\gamma-\mathrm{i}v)}x^{-\gamma-\mathrm{i}v}dv, \quad \text{for } 0<\gamma<1 \tag{17}$$

where $U_n, A_n \to \infty$ in a suitable way as $n\to\infty$. Note that in many cases the function $\Phi_n$ can be found in closed form. For example, consider the case of subordinated stable Lévy process with $\psi(\lambda)=|\lambda|^\alpha$. It then holds for $\mathrm{Re}(z)>0$,

$$\begin{aligned}\Phi_n(z,x) &= \int_0^{A_n}[\psi(\lambda)]^{z-1}e^{\mathrm{i}x\lambda}\psi'(\lambda)d\lambda\\ &= \alpha\int_0^{A_n}\lambda^{\alpha(z-1)}e^{\mathrm{i}x\lambda}\lambda^{\alpha-1}\,d\lambda\\ &= \alpha\int_0^{A_n}\lambda^{\alpha z-1}e^{\mathrm{i}x\lambda}d\lambda\\ &= \frac{A_n^{\alpha z}}{z}F_1(\alpha z;1+\alpha z;\mathrm{i}A_n x),\end{aligned}$$

where $F_1$ is Kummer's function. In the next two theorems we prove a remarkable result showing that the estimate $p_{T,n}(x)$ converges to $p(x)$ at the same rate (up to a logarithmic factor in the polynomial case) as in the case of the time-changed Brownian motion.

**Theorem 3.4.** *Suppose that $\psi$ satisfies the conditions of Proposition 3.2, and that moreover $\int_{\{|x|>1\}}|x|\nu(dx)<\infty$. Furthermore suppose that there is a $1/2<\gamma<1$ such that $p_T\in\mathcal{C}(\beta,\gamma,L)$ (cf. Theorem 2.2) for some $\beta>0$, and*

$$\int_1^\infty\frac{1}{\lambda^{2\gamma-1-\varepsilon}}\left|\mathcal{F}[p_X](\lambda)\right|d\lambda<\infty, \tag{18}$$

*for some $\varepsilon>0$. Then under the choice*

$$A_n = n^{\frac{1}{4(1-\gamma)+2\varepsilon}} \tag{19}$$

*and*

$$U_n = \frac{\varepsilon}{(2-2\gamma+\varepsilon)(2\beta+\pi)}\log n - \frac{2\gamma-1}{2\beta+\pi}\log\log n, \tag{20}$$

*we get*

$$\sup_{x\geq 0}\sqrt{\mathrm{E}\left[x^{2\gamma}\left|p_n(x)-p_T(x)\right|^2\right]}\lesssim n^{-\frac{\beta}{2\beta+\pi}\frac{\varepsilon}{2(1-\gamma)+\varepsilon}}\log^{\beta\frac{2\gamma-1}{2\beta+\pi}}n, \quad n\to\infty. \tag{21}$$

*Thus for $\gamma\to 1$ we recover the rates of Theorem 2.2 up to a logarithmic factor.*

*Remark* 3.5. Since

$$\int_1^\infty\frac{1}{\lambda^{2\gamma-1-\varepsilon}}\left|\mathcal{F}[p_X](\lambda)\right|d\lambda = \int_1^\infty\frac{1}{\lambda^{2\gamma-1-\varepsilon}}\left|\mathcal{L}[p_T](\psi(\lambda))\right|d\lambda,$$

the condition (18) is, for example, fulfilled for some $\varepsilon>0$ if $\mathrm{Re}[\psi(\lambda)]\gtrsim\lambda$ for $\lambda\to+\infty$ and $p_T$ is continuous in $0$ with $p_T(0)<\infty$.

In the case $p_T \in \mathcal{D}(\beta, \gamma, L)$ we get exactly the same logarithmic rate as in Theorem 2.6.

**Theorem 3.6.** *Suppose that $\psi$ and $\gamma$ are as in Theorem 3.4, and that now $p_T \in \mathcal{D}(\beta, \gamma, L)$ (cf. Theorem 2.6) for some $\beta > 0$. Further suppose that (18) holds. Then under the choice*

$$A_n = n^{\frac{1}{4(1-\gamma)+2\varepsilon}} \tag{22}$$

*(hence the same as in Theorem 3.4) and*

$$U_n = \frac{\varepsilon}{\pi\,(2 - 2\gamma + \varepsilon)} \log n - \frac{2\beta + 2\gamma - 1}{\pi} \log\log n, \tag{23}$$

*we get*

$$\sup_{x \geq 0} \sqrt{\mathrm{E}\left[x^{2\gamma}\,|p_n(x) - p_T(x)|^2\right]} \lesssim \log^{-\beta}(n), \quad n \to \infty.$$

**Discussion** The rates in Theorem 3.4 and Theorem 3.6 are optimal in minimax sense, since they are basically coincides (up to a logarithmic factor) with the rates in Theorem 2.2 and Theorem 2.6, respectively. As can be seen from the proof of Theorem 2.8 and Remark 3.5, the lower bonds continue to hold true under the additional assumption (18). Let us also stress that the class $\mathcal{C}(\beta, \gamma, L)$ is quite large and contains the well known families of distributions such as Gamma, Beta and Weibull families. It follows from Theorem 3.4 that for all these families our estimator $p_{n,T}$ converges at a polynomial rate.

# 4 Applications

## 4.1 Estimation of the variance-mean mixture models

The variance-mean mixture of the normal distribution is defined as

$$p(x) = \int_0^\infty (2\pi\sigma^2 u)^{-1/2} \exp(-(x - \mu u)^2/(2\sigma^2 u))\, g(u) du,$$

where $g(u)$ is a mixing density on $\mathbb{R}_+$. The variance-mean mixture models play an important role in both the theory and the practice of statistics. In particular, such mixtures appear as limit distributions in asymptotic theory for dependent random variables and they are useful for modeling data stemming from heavy-tailed and skewed distributions, see, e.g. [1] and [3]. As can be easily seen, the variance-mean mixture distribution $p$ coincides with the distribution of the random variable $\sigma W_T + \mu T$, where $T$ is the random variable with density $g$, which is independent of $W$. The class of variance-mean mixture models is rather large. For example, the class of the normal variance mixture distributions ($\mu = 0$) can be described as follows: $p$ is the density of a normal variance mixture (equivalently $p$ is the density of $W_T$) if and only if $\mathcal{F}[p](\sqrt{u})$ is a completely monotone function in $u$. The problem of statistical inference for variance-mean mixture models has been already considered in the literature. For example, Korsholm, [6] proved the consistency of the non-parametric maximum likelihood estimator for the parameters $\sigma$ and $\mu$, $g$ being treated as an infinite dimensional nuisance parameter. In Zhang [11] the problem of estimating the mixing density in location (mean) mixtures was studied. To the best of our knowledge, we here address, for the first time, the problem of non-parametric inference for the mixing density $g$ in full generality and derive the minimax convergence rates. In fact, Theorem 3.4 and Theorem 3.6 directly apply not only to normal variance-mean mixture models, but also to stable variance-mean mixtures.

## 4.2 Estimation of time-changed Lévy models

Let $L = (L_t)_{t\geq 0}$ be a one-dimensional Lévy process and let $\mathcal{T} = (\mathcal{T}(s))_{s\geq 0}$ be a non-negative, non-decreasing stochastic process independent of $X$ with $\mathcal{T}(0) = 0$. A time-changed Lévy process $Y = (Y_s)_{s\geq 0}$ is then defined as $Y_s = X_{\mathcal{T}(s)}$. The process $\mathcal{T}$ is usually referred to as time change or subordinator. Consider the problem of statistical inference on the distribution of the time change $\mathcal{T}$ based on the low-frequency observations of the time-changed Lévy process $X_t = L_{\mathcal{T}(t)}$. Suppose that $n$ observations of the Lévy process $L_t$ at times $t_j = j\Delta$, $j = 0, \ldots, n$, are available. If the sequence $\mathcal{T}(t_j) - \mathcal{T}(t_{j-1})$, $j = 1, \ldots, n$, is strictly stationary with the invariant stationary distribution $\pi$, then for any bounded "test function" $f$

$$\frac{1}{n}\sum_{j=1}^{n} f\left(L_{\mathcal{T}(t_j)} - L_{\mathcal{T}(t_{j-1})}\right) \to \mathbb{E}_\pi[f(L_{\mathcal{T}(\Delta)})], \quad n \to \infty, \tag{24}$$

The limiting expectation in (24) is then given by

$$\mathbb{E}_\pi[f(L_{\mathcal{T}(\Delta)})] = \int_0^\infty \mathbb{E}[f(L_s)]\,\pi(ds).$$

Taking $f(z) = f_u(z) = \exp(\mathrm{i}u^\top z)$, $u \in \mathbb{R}^d$, we arrive at the the following representation for the c.f. of $L_{\mathcal{T}(s)}$:

$$\mathbb{E}\left[\exp\left(\mathrm{i}uL_{\mathcal{T}(\Delta)}\right)\right] = \int_0^\infty \exp(t\psi(u))\,\pi(dt) = \mathcal{L}_\pi(\psi(u)), \tag{25}$$

where $\psi(u) := -t^{-1}\log(\phi_t(u))$ with $\phi_t(u) = \mathbb{E}\exp(\mathrm{i}u^\top L_t)$ being the characteristic exponent of the Lévy process $L$ and $\mathcal{L}_\pi$ is the Laplace transform of $\pi$. Suppose we want to estimate the invariant measure $\pi$ (or its density) from the discrete time observations of $L_{\mathcal{T}}$, then we are in the setting of the generalized statistical Skorohod embedding with the only difference that the elements of the sample $L_{\mathcal{T}(t_1)} - L_{\mathcal{T}(t_0)}, \ldots, L_{\mathcal{T}(t_n)} - L_{\mathcal{T}(t_{n-1})}$ are not necessarily independent. However, under appropriate mixing properties of the sequence $\mathcal{T}(t_j) - \mathcal{T}(t_{j-1})$, $j = 1, \ldots, n$, one can easily generalize the results of Section 3 to the case of dependent data (see, e.g. [2] for similar results). Let us note that the statistical inference for time-changed Lévy processes based on high-frequency observations of $Y$ has been the subject of many studies, see, e.g. Bull, [4] and Todorov and Tauchen, [9] and the references therein.

# 5 Numerical examples

Barndorff-Nielsen et al. [1] consider a class of variance-mean mixtures of normal distributions which they call generalized hyperbolic distributions. The univariate and symmetric members of this family appear as normal scale mixtures whose mixing distribution is the generalized inverse Gaussian distribution with density

$$p_T(v) = \frac{(\varkappa/\delta)^\lambda}{2K_\lambda(\delta\varkappa)} v^{\lambda-1} \exp\left(-\frac{1}{2}\left(\varkappa^2 v + \frac{\delta^2}{v}\right)\right), \quad v > 0, \tag{26}$$

where $K$ is a modified Bessel function. The resulting normal scale mixture has probability density function

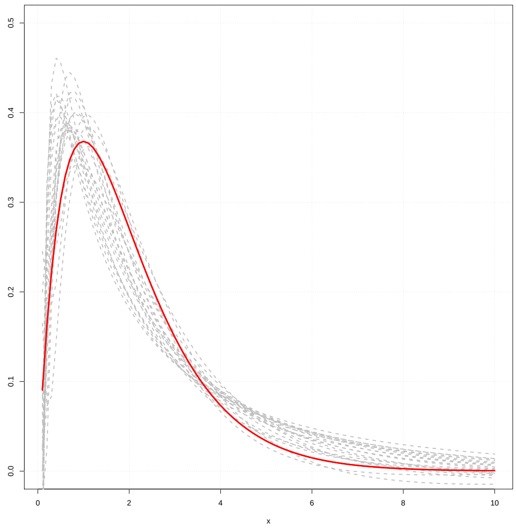
$$p_X(x) = \frac{\varkappa^{1/2}}{(2\pi)^{1/2}\delta^\lambda} K_\lambda(\delta\varkappa)(\delta^2 + x^2)^{\frac{1}{2}(\lambda-\frac{1}{2})} K_{\lambda-\frac{1}{2}}\big(\varkappa(\delta^2+x^2)^{1/2}\big).$$

Let us start with a simple example, Gamma density $p_T(x) = x\exp(-x)$, $x \geq 0$, which is a special case of (26) for $\delta = 0$, $\lambda = 2$ and $\varkappa = \sqrt{2}$. We simulate a sample of size $n$ from the distribution of $X$, and construct the estimate (5) with the bandwidth $h_n$ given (up to a constant not depending on $n$) by (8) and $\gamma = 0.8$. In Figure 2 (left), one can see $50$ estimated densities based on $50$ independent samples from $W_T$ of size $n = 1000$, together with $p_T$ in red. Next we estimate the distribution of the loss $\sup_{x\in[0,10]}\big\{|p_{T,n}(x) - p_T(x)|\big\}$ based on $100$ independent repetitions of the estimation procedure. The corresponding box plots for different $n$ are shown in Figure 2 (right).

Let us now turn to a more interesting example of variance-mean mixtures. We take $X = T + W_T$ and choose $T$ to follow a Gamma distribution with the density $p_T(x) = x\exp(-x)$, $x \geq 0$. The estimate (17) is constructed as follows. First note that $\psi(\lambda) = -\mathrm{i}\lambda + \lambda^2/2$. In order to numerically compute the function $\Phi_n(1-z, X_k)$ for $z = \gamma + \mathrm{i}v$ with $\gamma < 1$, we use the decomposition

$$\begin{aligned} \frac{1}{n}\sum_{k=1}^n \Phi_n(1-z, X_k) &= \int_0^{A_n} [\psi(\lambda)]^{-z}\,[\phi_n(\lambda) - e^{-m_n\psi(\lambda)}]\psi'(\lambda)\,d\lambda \\ &\quad + m_n^{z-1}\Gamma(1-z) + O\big(m_n^{-(1-\gamma)}\exp(-m_n A_n^2/2)\big), \end{aligned} \tag{27}$$

where $\phi_n(\lambda) = \frac{1}{n}\sum_{k=1}^n e^{\mathrm{i}\lambda X_k}$ is the empirical characteristic function and $m_n = \frac{1}{n}\sum_{k=1}^n X_k \to 2$. This decomposition follows from a Cauchy argument similar as in the proof of Proposition 3.2 and is quite useful to reduce the cost of computing the integral in (27), since the integral on the r.h.s. of (27) is much easier to compute due to the asymptotic relation $\phi_n(\lambda) - e^{-m_n\psi(\lambda)} = O(\lambda^2)$, $\lambda \to 0$. Next we take $\gamma = 0.7$, $A_n$ and $h_n$ as in Theorem 3.4 with $\varepsilon = 0.5$ and $\beta = \pi/2$ (see Example 2.4). Figure 3 shows the performance of the estimate defined in (17): on the left-hand side $20$ independent realizations of the estimate $p_{T,n}$ for $n = 1000$ are shown together with the true density $p_T$. The box plots of the loss $\sup_{x\in[0,10]}\big\{|p_{T,n}(x) - p_T(x)|\big\}$ based on $100$ runs of the algorithm are depicted on the right-hand side of Figure 3. By comparing the



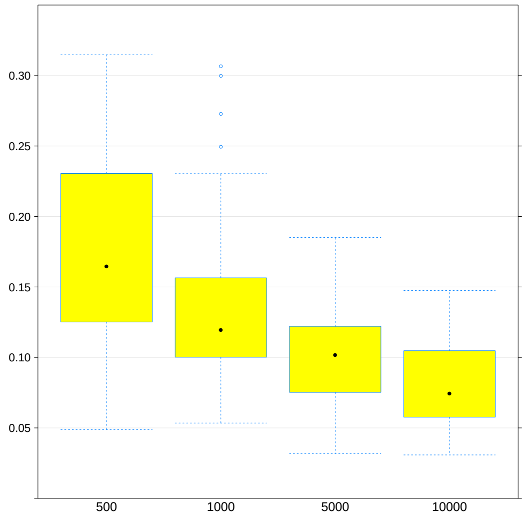



Figure 2: Left: the Gamma density (red) and its $50$ estimates (grey) for the sample size $n = 1000$. Right: the box plots of the loss $\sup_{x\in[0,10]}\left\{|p_{T,n}(x) - p_T(x)|\right\}$ for different sample sizes.

right-hand sides of Figure 2 and Figure 3, we observe that the performances of the estimates (17) and (5) are similar, although the estimate (5) seem to have higher variance. This supports the claim of Theorem 3.4 about the same convergence rates in statistical Skorohod embedding and generalized statistical Skorohod embedding problems, given that $p_T \in \mathcal{C}(\beta, \gamma, L)$.

# 6 Proofs

## 6.1 Proof of Theorem 2.2

First let us estimate the bias of $p_{T,n}$. We have

$$\begin{aligned}\mathbb{E}[p_{T,n}(x)] &= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} x^{-\gamma-\mathrm{i}v} K(vh_n)\frac{\mathcal{M}[p_{|X|}](2(\gamma+\mathrm{i}v)-1)}{2^{\gamma+\mathrm{i}v}\Gamma(\gamma-1/2+\mathrm{i}v)}\,dv \\ &= \frac{1}{2\pi}\int_{-1/h_n}^{1/h_n} x^{-\gamma-\mathrm{i}v}\mathcal{M}[p_T](\gamma+\mathrm{i}v)\,dv.\end{aligned}$$

Hence

$$p_T(x) - \mathbb{E}[p_{T,n}(x)] = \frac{1}{2\pi}\int_{\{|v|\geq 1/h_n\}} \mathcal{M}[p_T](\gamma+\mathrm{i}v)x^{-\gamma-\mathrm{i}v}dv$$

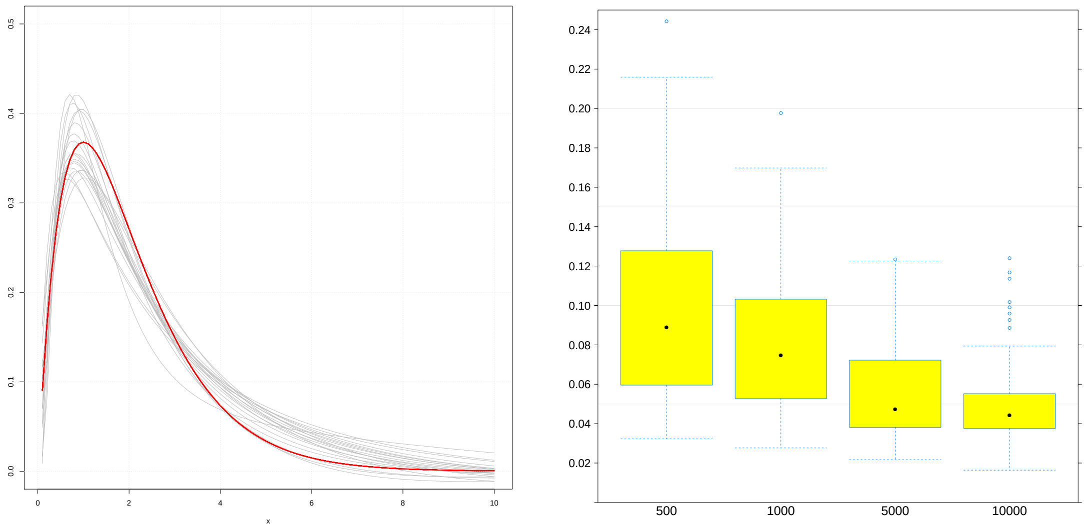



Figure 3: Left: the Gamma density (red) and its $20$ estimates (grey) for the sample size $n = 5000$. Right: the box plots of the loss $\sup_{x\in[0,10]}\{|p_{T,n}(x) - p_T(x)|\}$ for different sample sizes.

and we then have the estimate,

$$\begin{aligned}\sup_{x\geq 0}\{x^\gamma|\mathbb{E}[p_{T,n}(x)] - p_T(x)|\} &\leq \frac{1}{2\pi}\int_{\{|v|\geq 1/h_n\}} |\mathcal{M}[p_T](\gamma + \mathrm{i}v)|\, dv \\ &\leq \frac{e^{-\beta/h_n}}{2\pi}\int_{\{|v|\geq 1/h_n\}} e^{-\beta|v|}\, |\mathcal{M}[p_T](\gamma + \mathrm{i}v)|\, e^{\beta|v|} dv \\ &\leq L\,\frac{e^{-\beta/h_n}}{2\pi}. \end{aligned} \tag{28}$$

As to the variance, by the simple inequality $\mathrm{Var}\left(\int f_t dt\right) \leq \left(\int \sqrt{\mathrm{Var}[f_t]} dt\right)^2$, which holds for any random function $f_t$ with $\int \mathbb{E}[f_t^2]dt < \infty$, we get

$$\begin{aligned}\mathrm{Var}[x^\gamma p_{T,n}(x)] &= \mathrm{Var}\left[\frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} x^{-\mathrm{i}v} K(vh_n)\frac{\mathcal{M}_n[p_{|X|}](2(\gamma+\mathrm{i}v)-1)}{2^{\gamma+iv}\Gamma(\gamma - 1/2 + \mathrm{i}v)}\, dv\right] \\ &\leq \frac{1}{\pi 2^{2\gamma}}\left[\int_{-1/h_n}^{1/h_n} \frac{\sqrt{\mathrm{Var}\left(\mathcal{M}_n[p_{|X|}](2(\gamma+\mathrm{i}v)-1)\right)}}{|\Gamma(\gamma - 1/2 + \mathrm{i}v)|} dv\right]^2 \\ &\leq \frac{1}{2n\pi}\left[\int_{-1/h_n}^{1/h_n} \frac{\sqrt{\mathrm{Var}\left(|X|^{2(\gamma+\mathrm{i}v-1)}\right)}}{|\Gamma(\gamma - 1/2 + \mathrm{i}v)|} dv\right]^2 \\ &\leq \frac{1}{2n\pi}\left[\int_{-1/h_n}^{1/h_n} \frac{\sqrt{\mathbb{E}\left[|W_T|^{4(\gamma-1)}\right]}}{|\Gamma(\gamma - 1/2 + \mathrm{i}v)|}\, dv\right]^2. \end{aligned} \tag{29}$$

Note that

$$\begin{aligned}\mathbb{E}\left[|W_T|^{4(\gamma-1)}\right] &= \int_0^\infty \mathbb{E}\left[|W_t|^{4(\gamma-1)}\right] p_T(t)\, dt \\ &= \mathbb{E}\left[|W_1|^{4(\gamma-1)}\right] \int_0^\infty t^{2(\gamma-1)} p_T(t) dt \\ &=: C_2(\gamma) < \infty,\end{aligned}$$

due to (6). We obtain from (29) due to Corollary 7.4 (see Appendix) and by taking into account (6),

$$\mathrm{Var}[x^\gamma p_{T,n}(x)] \leq \frac{C_2(\gamma)}{2n\pi} C_3 h_n^{2(\gamma-1)} e^{\pi/h_n} = \frac{C_3(\gamma)}{n} h_n^{2(\gamma-1)} e^{\pi/h_n}.$$

and so (7) follows with $C_{\gamma,L} = \max(C_3(\gamma), \frac{L^2}{4\pi^2})$. Finally, by plugging (8) into (7) we get (9) and the proof is finished.

## 6.2 Proof of Theorem 2.6

The proof is analog to the one of Theorem 2.2 , the only difference is the bias estimate (28) that now becomes

$$\sup_{x \geq 0} \{ x^\gamma |\mathbb{E}[p_{T,n}(x)] - p_T(x)| \} \leq \frac{L}{2\pi} h_n^\beta,$$

which gives (10) with a constant $D_{\gamma,L} = \max(C_3(\gamma), \frac{L^2}{4\pi^2})$ again. Next with the choice (11) we obtain from (10) the logarithmic rate (13).

## 6.3 Proof of Theorem 2.8

Our construction relies on the following basic result (see [10] for the proof).

**Theorem 6.1.** *Suppose that for some $\varepsilon > 0$ and $n \in \mathbb{N}$ there are two densities $p_{0,n}, p_{1,n} \in \mathcal{G}$ such that*

$$d(p_{0,n}, p_{1,n}) > 2\varepsilon v_n.$$

*If the observations in model $n$ follow the product law $\mathsf{P}_{p,n} = \mathsf{P}_p^{\otimes n}$ under the density $p \in \mathcal{G}$ and*

$$\chi^2(p_{1,n} \,|\, p_{0,n}) \leq n^{-1} \log(1 + (2 - 4\delta)^2)$$

*holds for some $\delta \in (0, 1/2)$, then the following lower bound holds for all density estimators $\hat{p}_n$ based on observations from model $n$:*

$$\inf_{\hat{p}_n} \sup_{p \in \mathcal{G}} \mathrm{P}_p^{\otimes n} \left( d(\hat{p}_n, p) \geq \varepsilon v_n \right) \geq \delta.$$

*If the above holds for fixed $\varepsilon, \delta > 0$ and all $n \in \mathbb{N}$, then the optimal rate of convergence in a minimax sense over $\mathcal{G}$ is not faster than $v_n$.*

### 6.3.1 Proof of a lower bound for the class $\mathcal{C}(\beta, \gamma, L)$

Let us start with the construction of the densities $p_{0,n}$ and $p_{1,n}$. Define for any $\nu > 1$ and $M > 0$ two auxiliary functions

$$q(x) = \frac{\nu \sin(\pi/\nu)}{\pi} \frac{1}{1 + x^{\nu}}, \quad x \geq 0$$

and

$$\rho_M(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\log^2(x)}{2}} \frac{\sin(M \log(x))}{x}, \quad x \geq 0.$$

The properties of the functions $q$ and $\rho_M$ are collected in the following lemma.

**Lemma 6.2.** *The function $q$ is a probability density on $\mathbb{R}_+$ with the Mellin transform*

$$\mathcal{M}[q](z) = \frac{\sin(\pi/\nu)}{\sin(\pi z/\nu)}, \quad \text{Re}[z] > 0.$$

*The Mellin transform of the function $\rho_M$ is given by*

$$\mathcal{M}[\rho_M](u + \mathrm{i}v) = \frac{1}{2} \left[ e^{(u-1+\mathrm{i}(v+M))^2/2} - e^{(u-1+\mathrm{i}(v-M))^2/2} \right]. \tag{30}$$

*Hence*

$$\int_0^{\infty} \rho_M(x) dx = \mathcal{M}[\rho_M](1) = 0.$$

Set now for any $M > 0$

$$q_{0,M}(x) := q(x), \quad q_{1,M}(x) := q(x) + (q \vee \rho_M)(x),$$

where $f \vee g$ stands for the multiplicative convolution of two functions $f$ and $g$ on $\mathbb{R}_+$ defined as

$$(f \vee g)(x) := \int_0^{\infty} \frac{f(t) g(x/t)}{t} dt, \quad x \geq 0.$$

The following lemma describes some properties of $q_{0,M}$ and $q_{1,M}$.

**Lemma 6.3.** *For any $M > 0$ the function $q_{1,M}$ is a probability density satisfying*

$$\|q_{0,M} - q_{1,M}\|_{\infty} = \sup_{x \in \mathbb{R}_+} |q_{0,M}(x) - q_{1,M}(x)| \gtrsim \exp(-M\pi/\nu), \quad M \to \infty.$$

*Moreover, $q_{0,M}$ and $q_{1,M}$ are in $\mathcal{C}(\beta, \gamma, L)$ for all $0 < \beta < \pi/\nu$ and $\gamma > 0$ with $L$ depending on $\gamma$.*

*Proof.* First note that

$$\int_0^\infty q_{1,M}(x)dx = 1 + \int_0^\infty (q \vee \rho_M)(x) = 1 + \mathcal{M}[q](1)\mathcal{M}[\rho_M](1) = 1.$$

Furthermore, due to the Parseval identity

$$\begin{aligned}
(q \vee \rho_M)(y) &= \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{\log^2(x)}{2}} \frac{\sin(M\log(x))}{x^2} \frac{1}{1+(y/x)^\nu} dx \\
&= \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} \sin(Mv) \frac{e^{-v}}{1+e^{-\nu(v-y_l)}} dv \\
&= e^{-\log(y)} \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} \sin(Mv) \frac{e^{(y_l - v)}}{1+e^{\nu(\log(y)-v)}} dv \\
&= \frac{e^{-\log(y)}}{2\pi} \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} \sin(Mv) \frac{e^{(\log(y)-v)}}{1+e^{\nu(\log(y)-v)}} dv \\
&= \frac{e^{-\log(y)}}{2\pi} \int_{-\infty}^\infty e^{-\mathrm{i}u\log(y)} \left[\frac{H(u+M)-H(u-M)}{2}\right] \mathcal{F}[R](u) du,
\end{aligned}$$

where $R(x) = \frac{e^x}{1+e^{\nu x}}$ and $H(x) = e^{-x^2/2}$. Note that

$$\mathcal{F}[R](u) = \int_{-\infty}^\infty \frac{e^{x+\mathrm{i}ux}}{1+e^{\nu x}} dx = \frac{1}{\nu} \int_{-\infty}^\infty \frac{e^{v/\nu + \mathrm{i}uv/\nu}}{1+e^v} dx = \frac{1}{\nu} \Gamma\left(\frac{1+\mathrm{i}u}{\nu}\right) \Gamma\left(1 - \frac{1+\mathrm{i}u}{\nu}\right).$$

Hence due to (53)

$$\sup_{y\in\mathbb{R}_+} |q_{0,M}(y) - q_{1,M}(y)| = \sup_{y\in\mathbb{R}_+} |(q \vee \rho_M)(y)| \gtrsim \exp(-M\pi/\nu), \quad M \to \infty.$$

The second statement of the lemma follows from Lemma 6.2 and the fact that $\mathcal{M}[q \vee \rho_M] = \mathcal{M}[q]\mathcal{M}[\rho_M]$. □

Let $T_{0,M}$ and $T_{1,M}$ be two random variables with densities $q_{0,M}$ and $q_{1,M}$, respectively. Then the density of the r.v. $|W_{T_{i,M}}|$, $i = 0, 1$, is given by

$$p_{i,M}(x) := \frac{2}{\sqrt{2\pi}} \int_0^\infty \lambda^{-1/2} e^{-\frac{x^2}{2\lambda}} q_{i,M}(\lambda)\, d\lambda \quad i = 0, 1.$$

For the Mellin transform of $p_{i,M}$ we get

$$\begin{aligned}
\mathcal{M}[p_{i,M}](z) &= \mathbb{E}\big[|W_1|^{z-1}\big]\mathbb{E}\big[T_{i,M}^{(z-1)/2}\big] \\
&= \mathbb{E}\big[|W_1|^{z-1}\big]\mathcal{M}[q_{i,M}]((z+1)/2) \\
&= \frac{2^{z/2}}{\sqrt{2\pi}} \Gamma(z/2)\mathcal{M}[q_{i,M}]((z+1)/2), \quad i = 0, 1. \qquad (31)
\end{aligned}$$

**Lemma 6.4.** *The* $\chi^2$*-distance between the densities* $p_{0,M}$ *and* $p_{1,M}$ *fulfills*

$$\chi^2(p_{1,M}|p_{0,M}) = \int \frac{(p_{1,M}(x) - p_{0,M}(x))^2}{p_{0,M}(x)} dx \lesssim e^{-M\pi(1+2/\nu)}, \quad M \to \infty.$$

*Proof.* First note that $p_{0,M}(x) > 0$ on $[0, \infty)$. Since

$$\begin{aligned}
p_{0,M}(x) &= \frac{2}{\sqrt{2\pi}} \frac{\nu \sin(\pi/\nu)}{\pi} \int_0^\infty \lambda^{-1/2} e^{-\frac{x^2}{2\lambda}} \frac{1}{1+\lambda^\nu} d\lambda \\
/y = 1/\lambda/ &= \frac{2}{\sqrt{2\pi}} \frac{\nu \sin(\pi/\nu)}{\pi} \int_0^\infty y^{1/2} e^{-y\frac{x^2}{2}} \frac{1}{y^2(1+y^{-\nu})} dy \\
&= \frac{2}{\sqrt{2\pi}} \frac{\nu \sin(\pi/\nu)}{\pi} \int_0^\infty e^{-y\frac{x^2}{2}} \frac{y^{\nu-1/2-1}}{(1+y^\nu)} dy \\
&\asymp \frac{2}{\sqrt{2\pi}} \frac{\nu \sin(\pi/\nu)}{\pi} \Gamma(\nu - 1/2) x^{-2\nu+1}, \quad x \to \infty,
\end{aligned}$$

we have $p_{0,M}(x) \gtrsim x^{-2\nu+1}$, $x \to \infty$. Furthermore, due to (31) and the Parseval identity

$$\int_0^\infty x^{2\nu-1} |p_{0,M}(x) - p_{1,M}(x)|^2 dx =$$
$$\frac{2^{-4+2\nu}}{\pi} \int_{\gamma-i\infty}^{\gamma+i\infty} \mathcal{M}[q \vee \rho_M] \left(\frac{z+1}{2}\right) \Gamma\left(\frac{z}{2}\right) \mathcal{M}[q \vee \rho_M] \left(\frac{2\nu - z + 1}{2}\right) \Gamma\left(\frac{2\nu - z}{2}\right) dz, \tag{32}$$

where $\mathcal{M}[q \vee \rho_M](z) = \mathcal{M}[q](z)\mathcal{M}[\rho_M](z)$. Due to (30)

$$|\mathcal{M}[\rho_M](u + \mathrm{i}v)| \le e^{\frac{(u-1)^2}{2}} \frac{\phi(v+M) + \phi(v-M)}{2} \tag{33}$$

with $\phi(v) = e^{-\frac{v^2}{2}}$. Combining (53) (Appendix), (32) and (33), we derive

$$\begin{aligned}
\chi^2(p_{1,M}|p_{0,M}) &= \int \frac{(p_{1,M}(x) - p_{0,M}(x))^2}{p_{0,M}(x)} dx \\
&\lesssim \int_0^\infty (p_{1,M}(x) - p_{0,M}(x))^2 dx + \int_0^\infty x^{2\nu-1} (p_{1,M}(x) - p_{0,M}(x))^2 dx \\
&\lesssim \int_{-\infty}^\infty |v|^{\nu-1} e^{-|v|\pi/2 - |v|\pi/\nu} \left(\phi(v/2 + M) + \phi(v/2 - M)\right)^2 dv \\
&\lesssim M^{\nu-1} e^{-M\pi(1+2/\nu)}, \quad M \to \infty.
\end{aligned}$$

□

Fix some $\kappa \in (0, 1/2)$. Due to Lemma 6.4, the inequality

$$n\chi^2(p_{1,M}|p_{0,M}) \le \kappa$$

holds for $M$ large enough, provided

$$M = \frac{1+\varepsilon}{\pi(1+2/\nu)}(\log(n) + (\nu-1)\log\log(n))$$

for arbitrary small $\varepsilon > 0$. Hence Lemma 6.3 and Theorem 6.1 imply

$$\inf_{\hat{p}_n} \sup_{p\in\mathcal{C}(\beta,\gamma,L)} \mathsf{P}_{p,n}\big(\|\hat{p}_n - p\|_\infty \geq cv_n\big) \geq \delta.$$

for any $\beta < \pi/\nu < \pi$, any $\gamma > 0$, some constants $c > 0, \delta > 0$ and $v_n = n^{-\beta/(\pi+2\beta)} \log^{-\frac{\pi-\beta}{\pi+2\beta}}(n)$.

### 6.3.2 Proof of a lower bound for the class $\mathcal{D}(\beta, \gamma, L)$

Define for any $\nu > 1$, $\alpha > 0$ and $M > 0$,

$$q(x) = [2\Gamma(\nu)]^{-1} \times \begin{cases} \log^{\nu-1}(1/x), & 0 \leq x \leq 1, \\ x^{-2}\log^{\nu-1}(x), & x > 1 \end{cases}$$

and

$$\rho_M(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\log^2(x)}{2}} \frac{\sin(M\log(x))}{x\log(x)}, \quad x \geq 0.$$

The properties of the functions $q$ and $\rho_M$ can be found in the next lemma.

**Lemma 6.5.** *The function $q$ is a probability density on $\mathbb{R}_+$ with the Mellin transform*

$$\mathcal{M}[q](z) = \frac{1}{2}\left[z^{-\nu} + (2-z)^{-\nu}\right], \quad 0 < \mathsf{Re}[z] < 2.$$

*The Mellin transform of the function $\rho_M$ is given by*

$$\mathcal{M}[\rho_M](u + \mathrm{i}v) = e^{\frac{(u-1)^2}{2}} \frac{G(u, v+M) - G(u, v-M)}{2}, \tag{34}$$

*where $G(u, v) = \int_{-\infty}^{v} e^{-\frac{x^2}{2} + \mathrm{i}x(u-1)} dx$. Hence*

$$\zeta_M := \int_0^\infty \rho_M(x)dx = \mathcal{M}[\rho_M](1) = \int_{-M}^{M} e^{-\frac{x^2}{2}} dx.$$

Set now for any $M > 0$

$$q_{0,M}(x) := q(x), \quad q_{1,M}(x) := (1-\zeta_M)q(x) + (q \vee \rho_M)(x),$$

where $f \vee g$ stands for the multiplicative convolution of two functions $f$ and $g$ on $\mathbb{R}_+$ defined via

$$(f \vee g)(x) := \int_0^\infty \frac{f(t)g(x/t)}{t} dt.$$

**Lemma 6.6.** *For any $M > 0$, the function $q_{1,M}$ is a probability density satisfying*

$$\sup_{x\in(1-\delta,1+\delta)} |q_{0,M}(x) - q_{1,M}(x)| \asymp |\cos(\pi\nu/2)|M^{-\nu+1}, \quad M \to \infty,$$

*where $\delta > 0$ is a fixed number. Moreover, $q_{0,M}$ and $q_{1,M}$ are in $\mathcal{D}(\beta, \gamma, L)$ for all $\beta < \nu - 1$ and $\gamma \in (0, 2)$.*

*Proof.* First note that

$$\int_0^\infty q_{1,M}(x)dx = 1 + \int_0^\infty (q \vee \rho_M)(x) - \zeta_M = 1 + \mathcal{M}[\rho_M](1) \times \mathcal{M}[q](1) - \zeta_M = 1.$$

Furthermore, $(q \vee \rho_M)(y) = [2\Gamma(\nu)]^{-1} [I_1(y) + I_2(y)]$ with

$$\begin{aligned} I_1(y) &= \int_y^\infty e^{-\frac{\log^2(x)}{2\alpha}} x^{-2} \frac{\sin(M\log(x))}{\log(x)} \log^{\nu-1}(x/y)dx \\ &= \int_{\log(y)}^\infty e^{-\frac{z^2}{2\alpha}-z} \frac{\sin(Mz)}{z}(z - \log(y))^{\nu-1}dz \end{aligned}$$

and

$$\begin{aligned} I_2(y) &= \int_0^y e^{-\frac{\log^2(x)}{2\alpha}} y^{-2} \frac{\sin(M\log(x))}{\log(x)} \log^{\nu-1}(y/x)dx \\ &= \int_{-\infty}^{\log(y)} e^{-\frac{z^2}{2\alpha}+z} y^{-2} \frac{\sin(Mz)}{z}(\log(y) - z)^{\nu-1}dz. \end{aligned}$$

By taking $y = \exp(A)$, we get for $I_1(y)$

$$\begin{aligned} I_1(y) &= \int_0^\infty e^{-\frac{(z+A)^2}{2\alpha}-(z+A)} \frac{\sin(M(z+A))}{z+A} z^{\nu-1}dz \\ &= \cos(AM) \int_0^\infty \frac{e^{-\frac{(z+A)^2}{2\alpha}-(z+A)}}{z+A} \sin(Mz)z^{\nu-1}dz \\ &\quad + \sin(AM) \int_0^\infty \frac{e^{-\frac{(z+A)^2}{2\alpha}-(z+A)}}{z+A} \cos(Mz)z^{\nu-1}dz. \end{aligned}$$

The well known Erdélyi lemma implies

$$\int_0^\infty \frac{e^{-\frac{(z+A)^2}{2\alpha}-(z+A)}}{z+A} \sin(Mz)z^{\nu-1}dz \asymp \frac{e^{-\frac{A^2}{2\alpha}-A}}{A}\Gamma(\nu)\sin(\pi\nu/2)M^{-\nu}, \quad M \to \infty$$

and

$$\int_0^\infty \frac{e^{-\frac{(z+A)^2}{2\alpha}-(z+A)}}{z+A} \cos(Mz)z^{\nu-1}dz \asymp \frac{e^{-\frac{A^2}{2\alpha}-A}}{A}\Gamma(\nu)\cos(\pi\nu/2)M^{-\nu}, \quad M \to \infty.$$

Hence

$$I_1(e^A) \asymp \frac{e^{-\frac{A^2}{2\alpha}-A}}{A}\Gamma(\nu)\sin(AM+\pi\nu/2)M^{-\nu}, M\to\infty. \tag{35}$$

Analogously

$$\begin{aligned}
I_2(e^A) &= e^{-2A}\int_{-\infty}^{A} e^{-\frac{z^2}{2\alpha}+z}\frac{\sin(Mz)}{z}(A-z)^{\nu-1}dz \\
&= e^{-2A}\int_0^{\infty} e^{-\frac{(A-z)^2}{2\alpha}+A-z}\frac{\sin(M(A-z))}{A-z}z^{\nu-1}dz \\
&= e^{-2A}\sin(AM)\int_0^{\infty} e^{-\frac{(A-z)^2}{2\alpha}+A-z}\frac{\cos(Mz)}{A-z}z^{\nu-1}dz \\
&\quad -e^{-2A}\cos(AM)\int_0^{\infty} e^{-\frac{(A-z)^2}{2\alpha}+A-z}\frac{\sin(Mz)}{A-z}z^{\nu-1}dz \\
&\asymp \frac{e^{-\frac{A^2}{2\alpha}-A}}{A}\Gamma(\nu)\sin(AM-\pi\nu/2)M^{-\nu}.
\end{aligned}$$

Combining the previous estimates, we arrive at

$$I_2(e^A)+I_1(e^A) = 2\frac{e^{-\frac{A^2}{2\alpha}-A}}{A}\Gamma(\nu)\sin(AM)\cos(\pi\nu/2)M^{-\nu}.$$

It remains to note that the maximum of r.h.s of (35) is attained for $A\in\{\pi/2M, 3\pi/2M\}$ and

$$\sup_A[I_2(e^A)+I_1(e^A)] \asymp \Gamma(\nu)|\cos(\pi\nu/2)|M^{-\nu+1}.$$

The property $q_{1,M}\in\mathcal{D}(\beta,\gamma,L)$ for all $\beta<\nu-1$ and $\gamma\in(0,2)$ with $L$ depending on $\gamma$, follows from the identity $\mathcal{M}[q_{1,M}](z)=\mathcal{M}[q](z)(1-\zeta_M)+\mathcal{M}[\rho_M](z)\mathcal{M}[q](z)$ and (34). □

Let $T_{0,M}$ and $T_{1,M}$ be two random variables with densities $q_{0,M}$ and $q_{1,M}$ respectively. The the density of the r.v. $|W_{T_{i,M}}|$, $i=0,1$, is given by

$$p_{i,M}(x) := \frac{2}{\sqrt{2\pi}}\int_0^{\infty}\lambda^{-1/2}e^{-\frac{x^2}{2\lambda}}q_{i,M}(\lambda)d\lambda, \quad i=0,1.$$

For the Mellin transform of $p_{i,M}$, we have

$$\begin{aligned}
\mathcal{M}[p_{i,M}](z) &= \mathbb{E}\big[|W_1|^{z-1}\big]\mathbb{E}\big[T_{i,M}^{(z-1)/2}\big] \\
&= \mathbb{E}\big[|W_1|^{z-1}\big]\mathcal{M}[q_{i,M}]((z+1)/2) \\
&= \frac{2^{z/2}}{\sqrt{2\pi}}\Gamma(z/2)\mathcal{M}[q_{i,M}]((z+1)/2).
\end{aligned} \tag{36}$$

**Lemma 6.7.** *The* $\chi^2$*-distance between the densities* $p_{0,M}$ *and* $p_{1,M}$ *satisfies*

$$\chi^2(p_{1,M}|p_{0,M}) := \int \frac{(p_{1,M}(x) - p_{0,M}(x))^2}{p_{0,M}(x)}dx \lesssim e^{-M\pi/2}, \quad M \to \infty.$$

*Proof.* First note that $p_{0,M}(x) > 0$ on $[0, \infty)$. Since

$$\begin{aligned}
\int_0^1 \lambda^{-1/2} e^{-\frac{x^2}{2\lambda}} \log^{\nu-1}(1/\lambda) d\lambda &= \int_0^1 \lambda^{-1/2} e^{-\frac{x^2}{2\lambda}} \log^{\nu-1}(1/\lambda) d\lambda \\
/y = 1/\lambda, \lambda = 1/y/ &= \int_1^\infty y^{-3/2} e^{-x^2 y/2} \log^{\nu-1}(y) dy \\
&= \int_{x^2}^\infty x^{-2} (y/x^2)^{-3/2} e^{-y/2} \log^{\nu-1}(y/x^2) dy \\
&= x \int_{x^2}^\infty y^{-3/2} e^{-y/2} \log^{\nu-1}(y/x^2) dy \lesssim e^{-x^2/2}
\end{aligned}$$

and

$$\begin{aligned}
\int_1^\infty \lambda^{-3/2} e^{-\frac{x^2}{2\lambda}} \log^{\nu-1}(\lambda) d\lambda &= \int_0^1 y^{-1/2} e^{-\frac{x^2}{2} y} \log^{\nu-1}(1/y) dy \\
&\asymp \frac{\Gamma(1/2)}{\sqrt{2}} x^{-1} \log^{\nu-1}(x^2).
\end{aligned}$$

we have $p_{0,M}(x) \gtrsim x^{-1}$, $x \to \infty$. Furthermore, due to (36) and the Parseval identity

$$\int_0^\infty x^{a-1} \left| p_{0,M}(x) - p_{1,M}(x) \right|^2 dx =$$
$$\frac{2^{-4+a}}{\pi} \int_{\gamma - \mathrm{i}\infty}^{\gamma + \mathrm{i}\infty} \mathcal{M}[q \vee \rho_M] \left( \frac{z+1}{2} \right) \Gamma\left( \frac{z}{2} \right) \mathcal{M}[q \vee \rho_M] \left( \frac{a-z+1}{2} \right) \Gamma\left( \frac{a-z}{2} \right) dz, \quad (37)$$

where $\mathcal{M}[q \vee \rho_M](z) = \mathcal{M}[q](z)\mathcal{M}[\rho_M](z)$. Due to (34)

$$|\mathcal{M}[\rho_M](u + \mathrm{i}v)| \le e^{\frac{(u-1)^2}{2}} \frac{\Phi(v+M) + \Phi(v-M)}{2} \quad (38)$$

with $\Phi(v) = \int_{-\infty}^v e^{-\frac{x^2}{2}} dx$. Combining (37) with properly chosen $\gamma > 0$, (38) and Lemma 7.3 (see Appendix), we derive

$$\begin{aligned}
\chi^2(p_1|p_0) &= \int \frac{(p_1(x) - p_0(x))^2}{p_0(x)} dx \lesssim \int_0^\infty (p_1(x) - p_0(x))^2 dx + \int_0^\infty x (p_1(x) - p_0(x))^2 dx \\
&\lesssim \int_{-\infty}^\infty e^{-|v|\pi/2} \left( \Phi(v/2 + M) + \Phi(v/2 - M) \right)^2 dv \lesssim e^{-M\pi/2}, \quad M \to \infty.
\end{aligned}$$

□

Fix some $\kappa \in (0, 1/2)$. Due to Lemma 6.7, the inequality

$$n\chi^2(p_{1,M}|p_{0,M}) \leq \kappa$$

holds for $M$ large enough, provided

$$M = \frac{2(1+\varepsilon)}{\pi} \log(n)$$

for arbitrary small $\varepsilon > 0$. Hence Lemma 6.6 and Theorem 6.1 imply

$$\inf_{\hat{p}_n} \sup_{p \in \mathcal{D}(\beta,\gamma,L)} \mathsf{P}_{p,n}\big(\|\hat{p}_n - p\|_\infty \geq c v_n\big) \geq \delta.$$

for any $\beta < \nu - 1$, any $\gamma \in (0, 2)$, some constants $c > 0$, $\delta > 0$ and $v_n = \log^{-\beta}(n)$.

## 6.4 Proof of Proposition 2.3

It holds

$$\begin{aligned} p_{T,n}(x) - \mathbb{E}[p_{T,n}(x)] = \frac{1}{\sqrt{\pi}} \int_{-1/h_n}^{1/h_n} x^{-\gamma - \mathrm{i}v} \frac{K(vh_n)}{2^{\gamma+\mathrm{i}v}} \\ \times \frac{\big\{\mathcal{M}_n[p_{|X|}](2(\gamma+\mathrm{i}v)-1) - \mathcal{M}[p_{|X|}](2(\gamma+\mathrm{i}v)-1)\big\}}{\Gamma((\gamma+\mathrm{i}v)-1/2)}\, dv. \end{aligned}$$

Due to Proposition 7.1

$$\sup_{x \geq 0} \big\{x^\gamma |\mathbb{E}[p_{T,n}(x)] - p_T(x)|\big\} \leq \frac{\Delta_n}{\sqrt{\pi n}} \int_{-1/h_n}^{1/h_n} \frac{A_1}{2^\gamma} \frac{\log(e+|v|)}{\Gamma((\gamma+\mathrm{i}v)-1/2)}\, dv$$

with $\Delta_n = O_{a.s.}(1)$.

## 6.5 Proof of Proposition 3.2

Let $\theta_{\max}$ be such that $A = \tan\theta_{\max}$. At the arc $K_R : w = R\,e^{i\theta}$, $-\theta_{\max} < \theta < \theta_{\max}$, it holds that

$$\begin{aligned} \left|\int_{K_R} w^{z-1} \mathcal{L}[p_T](w) dw\right| &\leq R\theta_{\max} \cdot R^{\mathrm{Re}\, z - 1} \int e^{-xR\cos\theta_{\max}} p_T(x) dx \\ &\leq B\theta_{\max} R^{\mathrm{Re}\, z} \int e^{-xR\cos\theta_{\max}} dx = B\theta_{\max} \frac{R^{\mathrm{Re}\, z-1}}{\cos\theta_{\max}} \to 0, \end{aligned}$$

for $0 < \mathrm{Re}\, z < 1$, where $\sup_{x>0} p_T(x) \leq B$.

## 6.6 Proof of Proposition 3.4

By (16) we derive for the bias of $p_{T,n}(x)$, $x > 0$,

$$|\mathbb{E}[p_{T,n}(x)] - p_T(x)| = \left| \frac{1}{2\pi} \int_{-U_n}^{U_n} \frac{\mathrm{E}\left[\Phi_n(1-\gamma-\mathrm{i}v, X_1)\right]}{\Gamma(1-\gamma-\mathrm{i}v)} x^{-\mathrm{i}v} dv - \int_{-\infty}^{\infty} \mathcal{M}[p_T](\gamma+\mathrm{i}v) x^{-\gamma-\mathrm{i}v} dv \right|$$

$$\leq \left| \frac{1}{2\pi} \int_{-U_n}^{U_n} \frac{\int_{A_n}^{\infty} [\psi(\lambda)]^{-\gamma-\mathrm{i}v} \mathcal{F}[p_X](\lambda)\psi'(\lambda)d\lambda}{\Gamma(1-\gamma-\mathrm{i}v)} x^{-\gamma-\mathrm{i}v} dv \right| + \frac{|x|^{-\gamma}}{2\pi} \int_{\{|v|>U_n\}} |\mathcal{M}[p_T](\gamma+\mathrm{i}v)| \, dv$$

$$=: (*)_1 + (*)_2$$

Similar to the proof of Theorem 2.2 we have,

$$(*)_2 \leq \frac{|x|^{-\gamma}}{2\pi} e^{-\beta U_n} \int_{\{|v|>U_n\}} |\mathcal{M}[p_T](\gamma+\mathrm{i}v)| \, e^{\beta|v|} dv \leq e^{-\beta U_n} \frac{|x|^{-\gamma} L}{2\pi},$$

and by Lemma 7.2 and (53)

$$(*)_1 \lesssim \frac{|x|^{-\gamma}}{2\pi} \int_{-U_n}^{U_n} \frac{\int_{A_n}^{\infty} \lambda^{-2\gamma+1} |\mathcal{F}[p_X](\lambda)| \, d\lambda}{|\Gamma(1-\gamma-\mathrm{i}v)|} dv$$

$$\lesssim |x|^{-\gamma} U_n^{\gamma-1/2} e^{U_n \pi/2} \int_{A_n}^{\infty} \frac{\lambda^{-\varepsilon}}{\lambda^{2\gamma-1-\varepsilon}} |\mathcal{F}[p_X](\lambda)| \, d\lambda \lesssim |x|^{-\gamma} \frac{U_n^{\gamma-1/2} e^{U_n\pi/2}}{A_n^{\varepsilon}}.$$

As for the variance

$$\mathrm{Var}(p_{T,n}(x)) = \frac{1}{(2\pi)^2 n} \mathrm{Var}\left[ \int_{-U_n}^{U_n} \frac{\Phi_n(1-\gamma-\mathrm{i}v, X_1)}{\Gamma(1-\gamma-\mathrm{i}v)} x^{-\gamma-\mathrm{i}v} dv \right]$$

$$\leq \frac{1}{(2\pi)^2 n} |x|^{-2\gamma} \left[ \int_{-U_n}^{U_n} \frac{\sqrt{\mathrm{Var}[\Phi_n(1-\gamma-\mathrm{i}v, X_1)]}}{|\Gamma(1-\gamma-\mathrm{i}v)|} dv \right]^2, \qquad (39)$$

where

$$\sqrt{\mathrm{Var}[\Phi_n(1-\gamma-\mathrm{i}v, X_1)]} \leq \int_0^{A_n} \sqrt{\mathrm{Var}[[\psi(\lambda)]^{-\gamma-\mathrm{i}v} \, e^{\mathrm{i}X_1\lambda} \psi'(\lambda)]} d\lambda$$

$$= \int_0^{A_n} |\psi(\lambda)|^{-\gamma} \, |\psi'(\lambda)| \, \sqrt{\mathrm{Var}[e^{\mathrm{i}X_1\lambda}]} d\lambda.$$

Due to Lemma 7.2 we have

$$\int_1^{A_n} |\psi(\lambda)|^{-\gamma} \, |\psi'(\lambda)| \, \sqrt{\mathrm{Var}[e^{\mathrm{i}X_1\lambda}]} d\lambda \lesssim \int_1^{A_n} \lambda^{(1-2\gamma)} d\lambda \leq C_0 \frac{A_n^{2(1-\gamma)}}{1-\gamma}$$

and in any case of Lemma 7.2 it holds

$$\int_0^1 |\psi(\lambda)|^{-\gamma} \, |\psi'(\lambda)| \, \sqrt{\mathrm{Var}[e^{\mathrm{i}X_1\lambda}]} d\lambda \leq \int_0^1 |\psi(\lambda)|^{-\gamma} \, |\psi'(\lambda)| \, d\lambda \leq \frac{C_1}{1-\gamma}$$

for some natural constant $C_0, C_1 > 0$. Hence from (39) we get by (53),

$$|x|^{2\gamma}\mathrm{Var}(p_{T,n}(x)) \leq \frac{1}{(2\pi)^2 n}\left(CU_n^{\gamma-1/2}e^{U_n\pi/2}\frac{A_n^{2(1-\gamma)}}{1-\gamma}\right)^2 =: (*)_3,$$

and by gathering $(*)_1$, $(*)_2$, and $(*)_3$,

$$\sqrt{\mathrm{E}\left[x^{2\gamma}\left|p_n(x)-p(x)\right|^2\right]} \lesssim \frac{C}{2\pi\,(1-\gamma)\sqrt{n}}U_n^{\gamma-1/2}e^{U_n\pi/2}A_n^{2(1-\gamma)} + \frac{U_n^{\gamma-1/2}e^{U_n\pi/2}}{A_n^{\varepsilon}} + e^{-\beta U_n}.$$

Next, the choices (19) and (20) lead to the desired result.

# 7 Appendix

**Proposition 7.1.** *Let $Z_j$, $j = 1, \ldots, n$, be a sequence of independent identically distributed random variables. Fix some $u > 0$ and define*

$$\varphi_n(v) := \frac{1}{n}\sum_{j=1}^{n}\exp\left\{(u+\mathrm{i}v)\,Z_j\right\}, \quad v \in \mathbb{R}.$$

*Furthermore let $w$ be a positive monotone decreasing Lipschitz function on $\mathbb{R}_+$ such that*

$$0 < w(z) \leq \frac{1}{\sqrt{\log(e+|z|)}}, \quad z \in \mathbb{R}_+. \tag{40}$$

*Suppose that $\mathbb{E}\big[e^{puZ}\big] < \infty$ and $\mathbb{E}\big[|Z|^p\big] < \infty$ for some $p > 2$. Then with probability 1*

$$\|\varphi_n - \varphi\|_{L_\infty(\mathbb{R},w)} = O\left(\sqrt{\frac{\log n}{n}}\right). \tag{41}$$

*Proof.* Fix a sequence $\Xi_n \to \infty$ as $n \to \infty$. Denote

$$\begin{aligned}
\mathcal{W}_n^1(v) &:= \frac{w(v)}{n}\sum_{j=1}^{n}\Big(e^{(u+\mathrm{i}v)Z_j}\mathbb{I}\left\{e^{uZ_j} < \Xi_n\right\} - \mathbb{E}\left[e^{(u+\mathrm{i}v)Z}\mathbb{I}\left\{e^{uZ} < \Xi_n\right\}\right]\Big),\\
\mathcal{W}_n^2(v) &:= \frac{w(v)}{n}\sum_{j=1}^{n}\Big(e^{(u+\mathrm{i}v)Z_j}\mathbb{I}\left\{e^{uZ_j} \geq \Xi_n\right\} - \mathbb{E}\left[e^{(u+\mathrm{i}v)Z}\mathbb{I}\left\{e^{uZ} \geq \Xi_n\right\}\right]\Big),
\end{aligned}$$

where $Z$ is a random variable with the same distribution as $Z_1$. The main idea of the proof is to show that

$$|\mathcal{W}_n^1(v)| = O_{a.s.}\left(\sqrt{\frac{\log n}{n}}\right), \tag{42}$$

$$|\mathcal{W}_n^2(v)| = O_{a.s.}\left(\sqrt{\frac{\log n}{n}}\right) \tag{43}$$

under a proper choice of the sequence $\Xi_n$.

**Step 1.** The aim of the first step is to show (42). Consider the sequence $A_k = e^k$, $k \in \mathbb{N}$ and cover each interval $[-A_k, A_k]$ by $M_k = (\lfloor 2A_k/\gamma \rfloor + 1)$ disjoint small intervals $\Lambda_{k,1}, \ldots, \Lambda_{k,M_k}$ of the length $\gamma$. Let $v_{k,1}, \ldots, v_{k,M_k}$ be the centers of these intervals. We have for any natural $K > 0$

$$\max_{k=1,\ldots,K} \sup_{A_{k-1}<|v|\leq A_k} |\mathcal{W}_n^1(v)| \leq \max_{k=1,\ldots,K} \max_{1\leq m\leq M_k} \sup_{v\in\Lambda_{k,m}} |\mathcal{W}_n^1(v) - \mathcal{W}_n^1(v_{k,m})| + \max_{k=1,\ldots,K} \max_{\left\{\substack{1\leq m\leq M_k:\\ |v_{k,m}|>A_{k-1}}\right\}} |\mathcal{W}_n^1(v_{k,m})|.$$

Hence for any positive $\lambda$,

$$\mathrm{P}\left(\max_{k=1,\ldots,K} \sup_{A_{k-1}<|v|\leq A_k} |\mathcal{W}_n^1(v)| > \lambda\right) \leq \mathrm{P}\left(\sup_{|v_1-v_2|<\gamma} |\mathcal{W}_n^1(v_1) - \mathcal{W}_n^1(v_2)| > \lambda/2\right) + \sum_{k=1}^{K} \sum_{\left\{\substack{1\leq m\leq M_k:\\ |v_{k,m}|>A_{k-1}}\right\}} \mathrm{P}(|\mathcal{W}_n^1(v_{k,m})| > \lambda/2). \quad (44)$$

We proceed with the first summand in (44). It holds for any $v_1, v_2 \in \mathbb{R}$

$$\begin{aligned} |\mathcal{W}_n^1(v_1) - \mathcal{W}_n^1(v_2)| &\leq 2\,\Xi_n\,|w(v_1) - w(v_2)| + \frac{1}{n}\sum_{j=1}^{n}\Big[\Big|e^{(u+\mathrm{i}v_1)Z_j} - e^{(u+\mathrm{i}v_2)Z_j}\Big|\, I\left\{e^{uZ_j} < \Xi_n\right\}\Big] \\ &\quad + \Big|\mathbb{E}\left[\left(e^{(u+\mathrm{i}v_1)Z} - e^{(u+\mathrm{i}v_2)Z}\right) I\left\{e^{uZ} < \Xi_n\right\}\right]\Big| \\ &\leq |v_1 - v_2|\;\Xi_n\;\left[2\,L_w + \frac{1}{n}\sum_{j=1}^{n}|Z_j| + \mathbb{E}|Z|\right], \end{aligned} \quad (45)$$

where $L_\omega$ is the Lipschitz constant of $w$ and $Z$ is a random variable distributed as $Z_1$. Next, the Markov inequality implies

$$\mathrm{P}\left\{\frac{1}{n}\sum_{j=1}^{n}\Big[|Z_j| - \mathbb{E}|Z|\Big] > c\right\} \leq c^{-p}n^{-p}\,\mathbb{E}\left|\sum_{j=1}^{n}\Big[|Z_j| - \mathbb{E}|Z|\Big]\right|^p$$

for any $c > 0$. Note that

$$\mathbb{E}\left|\sum_{j=1}^{n}\Big[|Z_j| - \mathbb{E}|Z|\Big]\right|^p \leq c_p n^{p/2},$$

for some constant $c_p$ depending on $p$ and we obtain from (45)

$$\mathrm{P}\Big\{\sup_{|v_1-v_2|<\gamma} |\mathcal{W}_n^1(v_1) - \mathcal{W}_n^1(v_2)| > 2\gamma\Xi_n(L_\omega + \mathbb{E}|Z| + c)\Big\} \leq C_p\, c^{-p} n^{-p/2}.$$

Hence if $\gamma\Xi_n \geq 1$ and $\lambda \geq 4(L_\omega + \mathbb{E}|Z| + c)$ we get Now we turn to the second term on the right-hand side of (44). Applying the Bernstein inequality, we get

$$\mathrm{P}\Big\{\sup_{|v_1-v_2|<\gamma} |\mathcal{W}_n^1(v_1) - \mathcal{W}_n^1(v_2)| > \lambda/2\Big\} \leq C_p\, c^{-p} n^{-p/2}.$$

$$\mathrm{P}\left(|\mathrm{Re}\left[\mathcal{W}_n^1(v_{k,m})\right]| > \lambda/4\right) \leq \exp\left(-\frac{\lambda^2 n}{32(\Xi_n w(A_{k-1})\lambda/3 + w^2(A_{k-1})\,\mathbb{E}[e^{2uZ}])}\right).$$

Similarly,

$$\mathrm{P}\left(|\mathrm{Im}\left[\mathcal{W}_n^1(v_{k,m})\right]| > \lambda/4\right) \leq \exp\left(-\frac{\lambda^2 n}{32(\Xi_n w(A_{k-1})\lambda/3 + w^2(A_{k-1})\,\mathbb{E}[e^{2uZ}])}\right).$$

Therefore

$$\sum_{\{|v_{k,m}|>A_{k-1}\}} \mathrm{P}(|\mathcal{W}_n^1(v_{k,m})| > \lambda/2) \leq (\lfloor 2A_k/\gamma\rfloor + 1)\exp\left(-\frac{\lambda^2 n}{32(\Xi_n w(A_{k-1})\lambda/3 + w^2(A_{k-1})\,\mathbb{E}[e^{2uZ}])}\right).$$

Set now $\gamma = \sqrt{(\log n)/n}$, $\lambda = \zeta\sqrt{(\log n)/n}$ and $\Xi_n = \sqrt{n/\log(n)}$, then

$$\begin{aligned}\sum_{\{|v_{k,m}|>A_{k-1}\}} \mathrm{P}(|\mathcal{W}_n^1(v_{k,m})| > \lambda/2) &\lesssim A_k\sqrt{\frac{n}{\log(n)}}\exp\left(-\frac{\lambda^2 n}{32\big(\Xi_n w(A_{k-1})\lambda/3 + w^2(A_{k-1})\,\mathbb{E}[e^{2uZ}]\big)}\right)\\ &\lesssim \sqrt{\frac{n}{\log(n)}}\exp\left(-k + k\left[1 - \frac{\zeta^2\log(n)}{32(1+\mathbb{E}[e^{2uZ}])}\right]\right).\end{aligned}$$

Assuming that $\zeta^2 \geq 32\theta(1+\mathbb{E}[e^{2uZ}])$ for some $\theta > 1$, we arrive at

$$\sum_{k=2}^{\infty}\sum_{\{|v_{k,m}|>A_{k-1}\}} \mathrm{P}(|\mathcal{W}_n(v_{k,m})| > \lambda/2) \lesssim e^{-k}\frac{n^{1/2-\theta}}{\sqrt{\log(n)}}, \quad n\to\infty$$

**Step 2**. Now we turn to (43). Consider the sequence

$$R_n(v) := \frac{1}{n}\sum_{j=1}^{n} e^{(u+\mathrm{i}v)Z_j}\mathbb{I}\left\{e^{uZ_j} \geq \Xi_n\right\}.$$

By the Markov inequality we get for any $p > 1$

$$|\mathbb{E}\left[R_n(u)\right]| \leq \mathbb{E}\left[e^{uZ_j}\right]\mathrm{P}\left\{e^{uZ_j} \geq \Xi_n\right\} \leq \Xi_n^{-p}\ \mathbb{E}\left[e^{uZ_j}\right]\ \mathbb{E}\left[e^{upZ_j}\right] = o\Big(\sqrt{(\log n)/n}\Big)$$

Set $\eta_k = 2^k, k = 1, 2, \ldots$, then it holds for any $p > 2$

$$\sum_{k=1}^{\infty}\mathrm{P}\Big\{\max_{j=1,\ldots,\eta_{k+1}} e^{uZ_j} \geq \Xi_{\eta_k}\Big\} \leq \sum_{k=1}^{\infty}\eta_{k+1}\,\mathrm{P}\{e^{uZ} \geq \Xi_{\eta_k}\} \leq \mathbb{E}[e^{puZ}]\sum_{k=1}^{\infty}\eta_{k+1}\Xi_{\eta_k}^{-p} < \infty.$$

By the Borel-Cantelli lemma,

$$\mathrm{P}\Big\{\max_{j=1,\ldots,\eta_{k+1}} e^{uZ_j} \geq \Xi_{\eta_k} \quad \text{for infinitely many } k\Big\} = 0.$$

From here it follows that $R_n(u) - \mathbb{E}R_n(u) = o\Big(\sqrt{(\log n)/n}\Big)$. This completes the proof. $\square$

**Lemma 7.2.** *Let $(L_t, t \geq 0)$ be a Lévy process with the triplet $(\mu, \sigma^2, \nu)$. Suppose that $\int_{\{|x|>1\}} |x|\nu(dx) < \infty$, and that $\sigma$ and $\nu$ are not both zero. It then holds for $\psi(u) = -\log(\mathbb{E}(\exp(\mathrm{i}uL_t)))$*

$$(i): \ |\psi(u)| \lesssim u^2 \quad \text{and} \quad (ii): \ |\psi'(u)| \lesssim u, \quad u \to \infty. \tag{46}$$

*Further, if*

$$d = \mu + \int_{\{|x|>1\}} x\nu(dx) \neq 0 \tag{47}$$

*we have*

$$(i): \ |\psi(u)| \gtrsim u \quad \text{and} \ (ii): \ |\psi'(u)| \lesssim 1, \quad u \downarrow 0. \tag{48}$$

*If $d = 0$ we have in the case $\nu(\{|x| > 1\} \cap dx) \equiv 0$,*

$$(i): \ |\psi(u)| \gtrsim u^2, \quad \text{and} \quad (ii): \ |\psi'(u)| \lesssim u, \quad u \downarrow 0, \tag{49}$$

*and in the case $\nu(\{|x| > 1\} \cap dx) \neq 0$,*

$$(i): \ |\psi(u)| \gtrsim u, \quad \text{and} \quad (ii): \ |\psi'(u)| = o(1), \quad u \downarrow 0. \tag{50}$$

*Proof.* In general we have

$$\psi(u) = -\mathrm{i}u\mu + \frac{u^2\sigma^2}{2} + \int_{\mathbb{R}} (1 - e^{\mathrm{i}ux} + \mathrm{i}ux1_{|x|\leq 1})\nu(dx), \tag{51}$$

where

$$\int_{\mathbb{R}} (1 - e^{\mathrm{i}ux} + \mathrm{i}ux1_{|x|\leq 1})\nu(dx) = u^2 \int_{\{|x|\leq 1\}} \frac{1 - e^{\mathrm{i}ux} + \mathrm{i}ux}{(ux)^2} x^2\nu(dx) + \int_{\{|x|>1\}} \left(1 - e^{\mathrm{i}ux}\right)\nu(dx). \tag{52}$$

Note that

$$0 < c_1 < \frac{\left|1 - e^{\mathrm{i}y} + \mathrm{i}y\right|}{y^2} < c_2 \quad \text{for} \quad y \in \mathbb{R},$$

with $0 < c_1 < c_2$, and that

$$\int_{\{|x|>1\}} \left(1 - e^{\mathrm{i}ux}\right) xv(dx) \longrightarrow \int_{\{|x|>1\}} xv(dx) \quad \text{for } u \to \infty$$

by Riemann-Lebesgue. This yields (46)-$(i)$. It is not difficult to show by standard arguments that due to the integrability condition we have

$$\psi'(u) = -\mathrm{i}\mu + u\sigma^2 - \mathrm{i}\int_{\mathbb{R}} (e^{\mathrm{i}ux} - 1_{|x|\le 1})x\nu(dx).$$

Next, (46)-$(ii)$ follows by observing that

$$\int_{\{|x|\le 1\}} (e^{\mathrm{i}ux} - 1)xv(dx) = u\int_{\{|x|\le 1\}} \frac{e^{\mathrm{i}ux}-1}{ux}x^2\nu(dx),$$

where $\left(e^{\mathrm{i}y}-1\right)/y$ is bounded for $y\in\mathbb{R}$. Suppose $d\neq 0$. By (47), $\psi'(0) = -\mathrm{i}d \neq 0$, and since $\psi(0)=0$ we have (48)-$(i)$, and (48)-$(ii)$ is obvious. Next suppose $d=0$, i.e. $\psi'(0)=0$. We then have,

$$\begin{aligned}
\psi(u) &= \psi(u) - u\psi'(0) = \psi(u) + \mathrm{i}ud \\
&= \frac{u^2\sigma^2}{2} + \int_{\mathbb{R}} (1 - e^{\mathrm{i}ux} + \mathrm{i}ux1_{|x|\le 1})\nu(dx) + \mathrm{i}u\int_{\{|x|>1\}} x\nu(dx) \\
&= \frac{u^2\sigma^2}{2} + \int_{\{|x|\le 1\}} (1 - e^{\mathrm{i}ux} + \mathrm{i}ux)\nu(dx) \\
&+ \int_{\{|x|>1\}} (1 - e^{\mathrm{i}ux})\nu(dx) + \mathrm{i}u\int_{\{|x|>1\}} x\nu(dx)
\end{aligned}$$

and

$$\begin{aligned}
\psi'(u) &= \psi'(u) - \psi'(0) \\
&= u\sigma^2 - \mathrm{i}\int_{\mathbb{R}} (e^{\mathrm{i}ux} - 1_{|x|\le 1})x\nu(dx) + \mathrm{i}\int_{\{|x|>1\}} x\nu(dx).
\end{aligned}$$

If $\nu(\{|x|>1\}\cap dx)\equiv 0$ we thus have

$$\begin{aligned}
\psi(u) &= \frac{u^2\sigma^2}{2} + \int_{\{|x|\le 1\}} (1 - e^{\mathrm{i}ux} + \mathrm{i}ux)\nu(dx) \\
&= \frac{u^2\sigma^2}{2} + u^2\int_{\{|x|\le 1\}} \frac{1 - e^{\mathrm{i}ux} + \mathrm{i}ux}{(ux)^2}x^2\nu(dx)
\end{aligned}$$

and we observe that

$$\operatorname{Re}\left(1 - e^{\mathrm{i}ux} + \mathrm{i}ux\right) = 1 - \cos(ux) \ge 0$$

so in particular $\operatorname{Re}\psi(u) \gtrsim u^2$ while $|\psi(u)| \lesssim u^2$. Hence (49)-$(i)$ is shown. Then,

$$\begin{aligned}
\psi'(u) &= u\sigma^2 - \mathrm{i}\int_{\{|x|\le 1\}} (e^{\mathrm{i}ux} - 1)x\nu(dx) \\
&= u\sigma^2 - \mathrm{i}u\int_{\{|x|\le 1\}} \frac{e^{\mathrm{i}ux}-1}{ux}x^2\nu(dx)
\end{aligned}$$

and note again that $\left(e^{\mathrm{i}y}-1\right)/y$ is bounded, hence we have (49)-$(ii)$. Finally, if $d=0$ and $\nu(\{|x|>1\}\cap dx)\neq 0$, let us write

$$\psi(u)=\frac{u^2\sigma^2}{2}+u^2\int_{\{|x|\leq 1\}}\frac{1-e^{\mathrm{i}ux}+\mathrm{i}ux}{(ux)^2}x^2\nu(dx)$$
$$+\int_{\{|x|>1\}}(1-\cos(ux))\nu(dx)+\mathrm{i}\int_{\{|x|>1\}}(ux-\sin(ux))\,\nu(dx)$$

where

$$0\leq\int_{\{|x|>1\}}(ux-\sin(ux))\,\nu(dx)\leq u\int_{\{|x|>1\}}x\nu(dx)\lesssim u,$$

but due to dominated convergence also

$$\int_{\{|x|>1\}}(ux-\sin(ux))\,\nu(dx)=u\int_{\{|x|>1\}}x\nu(dx)+o(1).$$

Hence,

$$\int_{\{|x|>1\}}(ux-\sin(ux))\,\nu(dx)\asymp u,\quad u\downarrow 0,$$

and from this (50)-$(i)$. For the derivative we have,

$$\begin{aligned}\psi'(u)&=u\sigma^2-\mathrm{i}\int_{\mathbb{R}}(e^{\mathrm{i}ux}-1_{|x|\leq 1})xv(dx)+\mathrm{i}\int_{\{|x|>1\}}x\nu(dx)\\&=u\sigma^2-\mathrm{i}u\int_{\{|x|\leq 1\}}\frac{e^{\mathrm{i}ux}-1}{ux}x^2\nu(dx)-\mathrm{i}\int_{\{|x|>1\}}\left(e^{\mathrm{i}ux}-1\right)x\nu(dx)\\&=o(1),\qquad u\downarrow 0,\end{aligned}$$

by similar arguments, i.e. (50)-$(ii)$. □

**Lemma 7.3.** *For any $\alpha\geq -2$, there exist positive constants $C_1$ and $C_2(\alpha)$ such that uniformly for $|\beta|\geq 2$,*

$$C|\beta|^{\alpha-1/2}e^{-|\beta|\pi/2}\leq|\Gamma(\alpha+\mathrm{i}\beta)|\leq C_\alpha|\beta|^{\alpha-1/2}e^{-|\beta|\pi/2}. \tag{53}$$

**Corollary 7.4.** *For all $0<\alpha<1/2$ and all $U>2$, it holds*

$$\int_{-U}^{U}\frac{d\beta}{|\Gamma(\alpha+\mathrm{i}\beta)|}\leq CU^{1/2-\alpha}e^{U\pi/2} \tag{54}$$

*for a constant $C>0$. For $\alpha>1/2$, we have*

$$\int_{-U}^{U}\frac{d\beta}{|\Gamma(\alpha+\mathrm{i}\beta)|}\leq C_1(\alpha)+C_2e^{U\pi/2} \tag{55}$$

*where $C_2$ does not depend on $\alpha$.*

# References


[1] O Barndorff-Nielsen, John Kent, and Michael Sørensen. Normal variance-mean mixtures and z distributions. *International Statistical Review*, 50(2):145–159, 1982.

[2] Denis Belomestny. Statistical inference for time-changed Lévy processes via composite characteristic function estimation. *The Annals of Statistics*, 39(4):2205–2242, 2011.

[3] Nicholas H Bingham and Rüdiger Kiesel. Semi-parametric modelling in finance: theoretical foundations. *Quantitative Finance*, 2(4):241–250, 2002.

[4] Adam D Bull. Estimating time-changes in noisy Lévy models. *arXiv preprint arXiv:1312.5911*, 2013.

[5] Fabienne Comte and Valentine Genon-Catalot. Adaptive laguerre density estimation for mixed poisson models. *Archives-Ouvertes*, 2014.

[6] Lars Korsholm. The semiparametric normal variance-mean mixture model. *Scandinavian journal of statistics*, 27(2):227–261, 2000.

[7] Alexander Meister. *Deconvolution problems in nonparametric statistics*, volume 193. Springer, 2009.

[8] Jan Obłój et al. The Skorokhod embedding problem and its offspring. *Probability Surveys*, 1:321–392, 2004.

[9] Viktor Todorov and George Tauchen. Realized Laplace transforms for pure-jump semimartingales. *The Annals of Statistics*, 40(2):1233–1262, 2012.

[10] Alexandre B Tsybakov. *Introduction to nonparametric estimation*, volume 11. Springer, 2009.

[11] Cun-Hui Zhang. Fourier methods for estimating mixing densities and distributions. *The Annals of Statistics*, pages 806–831, 1990.