Uniform LAN property of locally stable Lévy process observed at high frequency

Dmytro Ivanenko, Alexey M. Kulik and Hiroki Masuda

Kyiv National Taras Shevchenko University
Volodymyrska, 64, Kyiv, 01033, Ukraine.
E-mail address: `ida@univ.kiev.ua`

Institute of Mathematics, Ukrainian National Academy of Sciences
01601 Tereshchenkivska, 3, Kyiv, Ukraine.
E-mail address: `kulik.alex.m@gmail.com`

Kyushu University, Faculty of Mathematics
744 Motooka Nishi-ku Fukuoka 819-0395, Japan.
E-mail address: `hiroki@math.kyushu-u.ac.jp`

Abstract. Suppose we have a high-frequency sample from the Lévy process of the form $X_t^\theta = \beta t + \gamma Z_t + U_t$, where Z is a possibly asymmetric locally α-stable Lévy process, and U is a nuisance Lévy process less active than Z. We prove the LAN property about the explicit parameter $\theta = (\beta, \gamma)$ under very mild conditions without specific form of the Lévy measure of Z, thereby generalizing the LAN result of Aït-Sahalia and Jacob (2007). In particular, it is clarified that a non-diagonal norming may be necessary in the truly asymmetric case. Due to the special nature of the local α-stable property, the asymptotic Fisher information matrix takes a clean-cut form.

1. Introduction

Ever since Le Cam's pioneering work Le Cam (1960), local asymptotics of likelihood random fields has been playing a crucial role in the theory of asymptotic inference. Specifically, the celebrated local asymptotic normality property (LAN) introduced by Le Cam has been a longstanding prominent concept, based on which we can deduce, among others, asymptotic optimality criteria for estimation and testing hypothesis. Not only for the classical i.i.d. models, there are many existing LAN results for several kinds of statistical experiments of dependent data,

including ergodic times-series models, homoscedastic models, and ergodic stochastic processes, to mention juts a few. One can consult Le Cam and Yang (2000) and the references therein for a systematic account of the LAN together with many related topics.

It is a common knowledge that verification of the LAN for a stochastic processes with no closed-form likelihood is generally a difficult matter. In case of diffusions under high-frequency, Gobet (2001, 2002) successfully derived the LAN and LAMN by means of the Malliavin calculus. There the structures of the limit experiments turned out to be simple enough (normal or mixed normal). One of theoretical merits of high-frequency sampling is that it enables us to take into account a small-time approximation of the underlying model, based on which we may derive an implementable and asymptotically efficient estimator. This has been achieved for the diffusion models, see Kessler (1997) and Genon-Catalot and Jacod (1993). However, to say nothing of Lévy driven non-linear stochastic differential equations, much less has been known about the explicit LAN result for Lévy processes observed at high frequency where the transition probability is hardly available in a closed form. We refer to Masuda (2015) for several explicit case studies about LAN result and related statistical-estimation problems concerning Lévy processes observed at high frequency. Especially when the underlying Lévy process or the most active part of the process is symmetric α-stable, the explicit LAN result has been proved in Aït-Sahalia and Jacob (2007) and Masuda (2009). See also Aït-Sahalia and Jacod (2008) for the precise asymptotic behavior of the Fisher-information matrix for the same model setting as in Aït-Sahalia and Jacob (2007).

We will consider the Lévy process X^θ described by $X_t^\theta = \beta t + \gamma Z_t + U_t$, where Z is a *locally α-stable* Lévy process and where U is a Lévy process which is independent of Z and *less active* than Z, the latter being regarded as a nuisance process; we specify them below. The objective of this paper is to derive the LAN about the explicit parameter $\theta = (\beta, \gamma)$ under very mild conditions, when X^θ is observed at high-frequency. Our model setting is quite broad to cover many specific examples of infinite-activity pure-jump Lévy processes, and in particular generalizes the LAN result of Aït-Sahalia and Jacob (2007), for the locally α-stable property only requires that the Lévy measure behaves like that of the α-stable distribution only near the origin, hence is much weaker requirement than the genuine α-stable case. It turns out that the special nature of the locally α-stable character leads to a clean-cut limit experiments described in terms of the α-stable density. Owing to high-frequency sampling, the method we propose is highly non-sensitive with respect to the nuisance process U, and allows us to formulate the LAN property uniformly with respect to a class of nuisance processes; this explains the term "uniform" in the title of the paper.

Our proof of the LAN property is based on two principal ingredients. One of them is the classical L_2-regularity technique, which dates back to Le Cam. Another important ingredient is the Malliavin calculus-based integral representation for the derivative of the log-likelihood function, which we use in order to derive the L_p-bounds for this derivative. This method of proof is mainly based on the ideas developed in Ivanenko and Kulik (2014, 2015) for the model where X^θ is a solution to a Lévy-driven SDE observed with a fixed frequency, but in the high-frequency case we encounter new challenge to design the particular version of the Malliavin calculus in a way which provide asymptotically precise L_p-bounds. We mention

an independent recent paper Clément and Gloter (2015), where similar tools are developed for the same purposes. Our way to obtain the asymptotically precise L_p-bounds and its relation to that developed in Clément and Gloter (2015) is discussed in details in Section 4 below.

It is natural to ask for extending our LAN result for stochastic differential equation driven by a locally α-stable Z. This extension is far-reaching and may involve the notion of the locally asymptotically mixed normality property (LAMN) introduced by Jeganathan (1982), which covers cases of random asymptotic Fisher information matrix. This is particularly relevant to heteroscedastic processes observed at n distinct time points over a fixed time domain. In such cases it is typical that randomness of the covariance structure is not averaged out in the limit experiments. See Dohnal (1987); Genon-Catalot and Jacod (1994); Gobet (2001) for the case of diffusion processes. The LA(M)N property of a solution to a SDE driven by a locally α-stable Lévy process under high-frequency sampling is one of currently-projected topics. To the best of our current knowledge, the papers Clément and Gloter (2015) and Mai (2014) are the only existing result in this direction. This will involve more technicalities than the present Lévy-process setting, and will be investigated in a subsequent paper.

This paper is organized as follows. In Section 2 we describe the model, introduce the assumptions, and formulate the main results of the paper. Section 3 contains the main part of the proof, which is based on the Le Cam's L_2-regularity technique and relies on L_p-bounds for the derivative of the log-likelihood function. These L_p-bounds are proved in Section 4 by means of a specially designed version of the Malliavin calculus.

2. Main results

Let X^θ be a Lévy process of the form

$$X_t^\theta = \beta t + \gamma Z_t + U_t, \quad t \geq 0. \tag{2.1}$$

Here Z and U are independent Lévy processes defined on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$, and $\theta = (\beta, \gamma)^\top \in \mathbb{R}^2$ is an unknown parameter subject to a statistical estimation. We assume Z to be such that in its Lévy-Khintchine representation

$$\mathsf{E}e^{i\lambda Z_t} = e^{t\psi(\lambda)},$$

the characteristic exponent ψ has the form

$$\psi(\lambda) = \int_{\mathbb{R}} \left(e^{i\lambda u} - 1 - i\lambda u 1_{|u|\leq 1} \right) \mu(du). \tag{2.2}$$

That is, Z does not contain the diffusion term, the truncation function equals $u1_{|u|\leq 1}$, and no additional drift term is involved. Throughout this paper, the Lévy measure μ is assumed to satisfy the following conditions:

H1. $\mu(du) = m(u)du$, and for some $\alpha \in (0, 2)$,

$$m(u) \sim \begin{cases} C_+|u|^{-\alpha-1}, & u \to 0+, \\ C_-|u|^{-\alpha-1}, & u \to 0-, \end{cases} \quad C_- + C_+ > 0.$$

H2. $m \in C^1(\mathbb{R} \setminus \{0\})$, and there exists a constant $u_0 > 0$ such that the function

$$\tau(u) := \frac{|um'(u)|}{m(u)}$$

is bounded on the set $\{|u| \leq u_0\}$ and satisfies

$$\int_{|u|>u_0} \tau^{2+\delta}(u)\mu(du) < \infty$$

for some $\delta > 0$.

Recall that for an α-stable process its Lévy measure has the density

$$m_{\alpha,C_\pm}(u) := \begin{cases} C_+|u|^{-\alpha-1}, & u > 0, \\ C_-|u|^{-\alpha-1}, & u < 0. \end{cases} \qquad (2.3)$$

Hence **H1** requires that locally near the origin the Lévy measure for Z behaves similar to that for an α-stable process; that is why we call Z *locally α-stable*. The constant $(C_+ - C_-)/(C_+ + C_-) \in [-1, 1]$ determines the signed degree of skewness, see Zolotarev (1986) for details.

Note that **H2** does not require $\tau(u)$ to be bounded for "large" u; this includes into the class of admissible Z a wide range of "stable-like" Lévy processes with

$$m(u) = f(u)m_{\alpha,C_\pm}(u),$$

where $f(u) \to 1, |u| \to 0$.

Example 2.1 (Tempered α-stable process). For either $f(u) = e^{1-\sqrt{1+u^2}}$, $f(u) = e^{-u^2}$, or $f(u) = e^{-|u|}$, conditions **H1**, **H2** hold true, although $\tau(u)$ fails to be bounded.

Example 2.2 (Smoothly damped α-stable process).
Let $m(u) = f(u)|u|^{-\alpha-1}1_{[-u_1,u_1]}(u)$, $u_1 > 0$, where f is continuous in $\mathbb{R}$, $f > 0$ for $u \in [-u_1, u_1]$ with $f(u) \to 1$ as $u \to 0$, and f smoothly vanishes outside the interval $[-u_1, u_1]$ in such a way that $u \mapsto |u||f'(u)|/f(u)$ is locally bounded and moreover $u \mapsto \{|u||f'(u)|/f(u)\}^{2+\delta}m(u)$ is du-integrable over the set $\{|u| \geq u_0\}$ for some $\delta > 0$ and $u_0 > 0$. Then conditions **H1**, **H2** hold true; note that $\tau(u) \leq |u||f'(u)|/f(u) + \alpha + 1$. One particular example of such f is of the form

$$f(u) = Ce^{-1/(u+u_1)-1/(u_1-u)}1_{[-u_1,u_1]}(u).$$

We are focused on the following setting:

- the process X^θ is *discretely observed*, i.e. the n-th sample contains its values at the first n points $\{t_{k,n} = kh_n, k = 1, \dots, n\}$ of the uniform partition of the time axis with the partition interval h_n;
- $h_n \to 0$ as $n \to \infty$, i.e. the discrete observations of X^θ have *high frequency*.

We note that the terminal sampling time nh_n may or may not tend to infinity as $n \to \infty$.

In what follows an open set $\Theta \in \mathbb{R}^2$ denotes the set of possible values of the unknown parameter θ; we assume that $\Theta \subset \mathbb{R} \times (0, \infty)$, i.e. parameter γ takes only positive values. Denote P_n^θ the law of the sample

$$\left\{X_{t_{k,n}}^\theta, k = 1, \dots, n\right\}$$

in $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, and write

$$\mathcal{E}_n = \left\{\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), (\mathsf{P}_n^\theta, \theta \in \Theta)\right\}$$

for a statistical model based on this sample.

Under our conditions on the process Z, the law P_n^θ is absolutely continuous with respect to Lebesgue measure (see Section 3), i.e. the model $\mathcal{E}_n$ possesses the *likelihood function*

$$L_n\left(\theta; x_1, \ldots, x_n\right) = \frac{\mathsf{P}_n^\theta(dx_1 \ldots dx_n)}{dx_1 \ldots dx_n}$$

Denote by

$$Z_n(\theta_0, \theta; x_1, \ldots, x_n) = \frac{L_n\left(\theta; x_1, \ldots, x_n\right)}{L_n\left(\theta_0; x_1, \ldots, x_n\right)}$$

the *likelihood ratio* of P_n^θ with respect to $\mathsf{P}_n^{\theta_0}$ with the convention $(\text{anything})/0 = \infty$.

Our goal is to establish the LAN property for the sequence of statistical models $\mathcal{E}_n, n \geq 1$, specified above. Recall that the LAN property is said to hold at a point $\theta_0 \in \Theta$ with the matrix rate $\{r(n) = r(n, \theta_0), \ n \geq 1\}$ and the covariance matrix $\Sigma(\theta_0)$, if for every v the sampled likelihood ratio

$$Z_n(\theta_0, \theta_0 + r(n)v) = Z_n(\theta_0, \theta_0 + r(n)v; X_{t_{1,n}}^{\theta_0}, \ldots, X_{t_{n,n}}^{\theta_0})$$

possesses representation under

$$Z_n(\theta_0, \theta_0 + r(n)v) = \exp\left\{ v^\top \Delta_n(\theta_0) - \frac{1}{2} v^\top \Sigma(\theta_0)v + \Psi_n(v, \theta_0) \right\} \tag{2.4}$$

with

$$\Delta_n(\theta_0) \Rightarrow \mathcal{N}(0, \Sigma(\theta_0)), \quad n \to \infty \tag{2.5}$$

and

$$\Psi_n(v, \theta_0) \xrightarrow{\mathsf{P}} 0, \quad n \to \infty \tag{2.6}$$

along $\mathsf{P}_n^{\theta_0}$.

Put

$$c_t = t \int_{t^{1/\alpha} < |u| \leq 1} u\,\mu(du), \tag{2.7}$$

which is identically zero if μ is symmetric and denote by $Z^{\alpha,C_\pm}$ the α-stable process whose characteristic exponent has the form (2.2) with the Lévy measure (2.3), where C_+, C_- are given by the condition **H1**. Finally, denote by $\phi_{\alpha,C_\pm}$ the distribution density of $Z_1^{\alpha,C_\pm}$ (this density exists, see Zolotarev (1986) or Proposition 3.1 below).

Now we are able to formulate our main result.

Theorem 2.1. *Let X^θ be given by* (2.1) *and assume that Z satisfies **H1** and **H2**, that*

$$t^{-1/\alpha} U_t \to 0, \quad t \to 0, \tag{2.8}$$

in probability, and that

$$n^{-1/2} h_n^{1/\alpha - 1} \to 0$$

(automatic if $\alpha \in (0,1]$ since we are supposing that $h_n \to 0$). Then the LAN property holds true at every point $\theta_0 \in \Theta$ with

$$r(n) = n^{-1/2} \begin{pmatrix} h_n^{1/\alpha - 1} & c_{h_n} h_n^{-1} \\ 0 & 1 \end{pmatrix}, \quad \Sigma(\theta) = \begin{pmatrix} \Sigma_{11}(\theta) & 0 \\ 0 & \Sigma_{22}(\theta) \end{pmatrix}, \tag{2.9}$$

where

$$\Sigma_{11}(\theta_0) = \gamma_0^{-2} \int_{\mathbb{R}} \left(\frac{\phi'_{\alpha,C_\pm}(x)}{\phi_{\alpha,C_\pm}(x)} \right)^2 \phi_{\alpha,C_\pm}(x) dx,$$

$$\Sigma_{22}(\theta_0) = \gamma_0^{-2} \int_{\mathbb{R}} \left(1 + \frac{x\phi'_{\alpha,C_\pm}(x)}{\phi_{\alpha,C_\pm}(x)} \right)^2 \phi_{\alpha,C_\pm}(x) dx.$$

Remark 2.1. Recall the definition of the Blumenthal-Getoor activity index of a Lévy process Y with the Lévy measure μ_Y:

$$\alpha_Y := \inf \left\{ q \geq 0 : \int_{|u| \leq 1} |u|^q \mu_Y(du) < \infty \right\}.$$

Then it is sufficient for the condition (2.8) that $\alpha_U < \alpha$ (see, e.g., p.362 of Sato (1999)); note that $\alpha_Z = \alpha$. In this paper we are assuming that the activity index α is known. This might seem disappointing, however, as it was clarified in Aït-Sahalia and Jacod (2008) and Masuda (2009), if one attempts to make joint maximum-likelihood estimation of α and the scale parameter γ, one may confront the degeneracy of the asymptotic Fisher information matrix. This degeneracy is inevitable, and how to cope with it is beyond the scope of this paper.

Remark 2.2. In view of the standard theory Ibragimov and Has'minskiĭ (1981) concerning asymptotically efficient estimation of a LAN model, Theorem 2.1 suggests to seek an estimator $\hat{\theta} = (\hat{\beta}_n, \hat{\gamma}_n)^\top$ such that

$$r(n)^{-1}(\hat{\theta}_n - \theta_0) = \left(\begin{array}{c} \sqrt{n}h_n^{1-1/\alpha}(\hat{\beta}_n - \beta_0) - h_n^{-1/\alpha}c_{h_n} \cdot \sqrt{n}(\hat{\gamma}_n - \gamma_0) \\ \sqrt{n}(\hat{\gamma}_n - \gamma_0) \end{array} \right)$$

weakly tends to the centered normal distribution with the asymptotic covariance matrix $\Sigma(\theta_0)$. Observe that, when μ is asymmetric, the factor

$$h_n^{-1/\alpha}c_{h_n} = h_n^{1-1/\alpha} \int_{h_n^{1/\alpha} < |u| \leq 1} u\mu(du)$$

may or may not vanish, or even may diverge, implying that the asymmetry essentially and non-trivially affect estimation of the drift parameter β. As a matter of fact, the necessity of non-diagonal norming seems to be non-standard in the literature: typically, it is enough to take

$$r(n) = \text{diag}\left\{ \left(\mathsf{E} \left| \frac{\partial}{\partial \theta_j} \log L_n(\theta; X_{t_{1,n}}^{\theta_0}, \ldots, X_{t_{n,n}}^{\theta_0}) \right|^2 \right)_j \right\}$$

whenever exists; for instance, the monograph Basawa and Scott (1983) is devoted to the diagonal norming. We refer to Fahrmeir (1988) and Sweeting (1992) for some technical refinements of asymptotic inference by using a non-diagonal norming. Nevertheless, we note that since $r(n)$ of (2.9) is invertible, we have no trouble in construction of an asymptotically confidence region of an asymptotically normally distributed estimator converging at rate $r(n)$.

Under the condition (2.8) the process U is interpreted as a "nuisance noise", in the sense that U is less active than the "principal" part Z, as was mentioned in Remark 2.1. A natural question is whether or not it is possible to extend Theorem 2.1 so as to make the LAN property valid not only for each single U, but also *uniformly* over some "nuisance class" $\mathfrak{U}$ of U. Our method of proof of Theorem

2.1 is strong enough to provide the following *uniform* LAN property in such an extended setting.

Theorem 2.2. *Let $\mathfrak{U}$ be a class of Lévy processes such that condition (2.8) holds true uniformly over $U \in \mathfrak{U}$. If in addition Z and h_n satisfy conditions of Theorem 2.1, then for every $U \in \mathfrak{U}$ and $\theta_0 \in \Theta$, the likelihood ratio for the discretely observed process (2.1) admits a representation (2.4) with $r(n), \Sigma(\theta_0)$ specified in (2.9), and relations (2.5), (2.6) hold true uniformly over $U \in \mathfrak{U}$.*

As it was explained in Aït-Sahalia and Jacob (2007), the uniform negligibility of U would play an important role for purposes of *semiparametric statistical (adaptive) estimation* of θ: in p.358 of Aït-Sahalia and Jacob (2007), the authors introduce a class of possible nuisance noise distribution $\mathcal{L}(U_1)$, over which one can precisely formulate an asymptotically uniformly efficient estimation of θ; this in turn leads to the notion of asymptotically uniformly efficient estimator of θ. As a matter of fact, it would be possible to precisely state a uniform-in-U version of the Hajék-Le Cam convolution theorem, which effectively clarifies the uniform asymptotic lower bound of an expected loss of any regular estimator with $r(n)$-rate of convergence; among others, see Section 2.3 of Basawa and Scott (1983) and Section II.11 of Ibragimov and Has′minskiĭ (1981) for details. How to construct an asymptotically efficient estimator would be several things, to be reported elsewhere.

3. Proofs of Theorem 2.1 and Theorem 2.2

In this section we prove Theorem 2.1 and outline the proof of Theorem 2.2. The key ingredient in these proofs would the L_p-bound for the derivative of the log-likelihood (Proposition 3.2), which we discuss in details and prove separately in Section 4 below.

3.1. *Proof of Theorem 2.1: an outline and preliminaries.* Denote by $p_t(\theta; x, y)$ the transition probability density for X^θ, considered as a Markov process; in what follows we will prove that this density exists. Denote also

$$g_t(\theta; x, y) = \frac{\nabla_\theta p_t(\theta; x, y)}{p_t(\theta; x, y)} = \nabla_\theta \log p_t(\theta; x, y),$$

$$q_t(\theta; x, y) = \frac{\nabla_\theta p_t(\theta; x, y)}{2\sqrt{p_t(\theta; x, y)}} = \nabla_\theta \sqrt{p_t(\theta; x, y)},$$

assuming the derivatives to exist for $P_t(x, \cdot)$-a.a. y for every fixed x, t. Since X^θ has independent increments, we can write

$$p_t(\theta; x, y) = p_t(\theta; y - x), \quad g_t(\theta; x, y) = g_t(\theta; y - x), \quad q_t(\theta; x, y) = q_t(\theta; y - x).$$

Then the sampled likelihood ratio for the model can be written in the form

$$Z_n(\theta_0, \theta_0 + r(n)v) = \prod_{k=1}^{n} \frac{p_{h_n}(\theta_0 + r(n)v; X_{t_{k,n}}^{\theta_0} - X_{t_{k-1,n}}^{\theta_0})}{p_{h_n}(\theta_0; X_{t_{k,n}}^{\theta_0} - X_{t_{k-1,n}}^{\theta_0})}.$$

Denote

$$\eta_{k,n}^\theta = X_{t_{k,n}}^\theta - X_{t_{k-1,n}}^\theta, \quad 1 \le k \le n,$$

and observe that $p_{h_n}(\theta; \cdot)$ is the distribution density for η_{kn}^θ. Hence the statistical model described above, after a re-sampling

$$(X_{t_{k,n}}^\theta)_{k=1}^n \mapsto (\eta_{k,n}^\theta)_{k=1}^n,$$

actually is reduced to the one with a triangular array of independent observations. The LAN property for triangular arrays of independent observations is well studied, e.g. Ibragimov and Has′minskiĭ (1981), Theorem II.3.1′, Theorem II.6.1, and Remark II.6.2. In particular, in order to prove the required LAN property at a point $\theta_0 \in \Theta$ it would be enough for us to prove the following assertions.

A1 For every n, the function

$$\Theta \ni \theta \to \sqrt{p_{h_n}(\theta; \cdot)} \in L_2(\mathbb{R})$$

is continuously differentiable; that is, the statistical experiment is *regular*.

A2

$$\lim_{n\to\infty} \mathsf{E} \left| \sum_{k=1}^n \left(r(n)^\top g_{h_n} \left(\theta_0; X_{kh_n}^{\theta_0} - X_{(k-1)h_n}^{\theta_0} \right) \right)^{\otimes 2} - \Sigma(\theta_0) \right| = 0.$$

A3 For some $\varepsilon > 0$,

$$\lim_{n\to\infty} n \int_{\mathbb{R}} \left| r(n)^\top g_{h_n}(\theta_0; y) \right|^{2+\varepsilon} p_{h_n}(\theta_0; y) \, dy = 0.$$

A4 For every $N > 0$,

$$\lim_{n\to\infty} \sup_{|v|<N} n \int_{\mathbb{R}} \left| r(n)^\top \left(q_{h_n}(\theta_0 + r(n)v; y) - q_{h_n}(\theta_0; y) \right) \right|^2 dy = 0.$$

Before proving **A1**–**A4**, let us introduce some notation, formulate auxiliary statements, and make preliminary calculation.

Denote $\tilde{r}(n) = n^{1/2} r(n)$. Denote also by $\tilde{\Theta}$ arbitrary (but fixed) subset of Θ such that

$$\inf\{\gamma : (\beta, \gamma) \in \tilde{\Theta}\} > 0.$$

Consider the random variables

$$\zeta_{\alpha,t} = t^{-1/\alpha}(Z_t + c_t).$$

The following statement is proved in Appendix.

Proposition 3.1.   (1) $\zeta_{\alpha,t} \Rightarrow Z_1^{\alpha,C_\pm}, t \to 0+$.

(2) *The variables $\zeta_{\alpha,t}$, $t > 0$, an $Z_1^{\alpha,C_\pm}$ possess the distribution densities $\phi_{\alpha,t}$, $t > 0$, and $\phi_{\alpha,C_\pm}$, respectively. These densities are infinitely differentiable, bounded together with their derivatives, and for every $N > 0$*

$$\sup_{|x|\leq N} |\phi_{\alpha,t}(x) - \phi_{\alpha,C_\pm}(x)| \to 0, \quad \sup_{|x|\leq N} |\phi'_{\alpha,t}(x) - \phi'_{\alpha,C_\pm}(x)| \to 0, \quad t \to 0+. \quad (3.1)$$

Denote

$$Y_t = X_t - U_t = \gamma t^{1/\alpha} \zeta_{\alpha,t} + \beta t - \gamma c_t,$$

then the distribution density for Y_t under P^θ equals

$$p_t^Y(\theta; x) = \gamma^{-1} t^{-1/\alpha} \phi_{\alpha,t}\left(\gamma^{-1} t^{-1/\alpha}(x - \beta t + \gamma c_t) \right), \qquad (3.2)$$

and consequently

$$\partial_\beta p_t^Y(\theta;x) = -\gamma^{-2}t^{1-2/\alpha}\phi'_{\alpha,t}\Big(\gamma^{-1}t^{-1/\alpha}(x-\beta t+\gamma c_t)\Big),$$

$$\partial_\gamma p_t^Y(\theta;x) = -\gamma^{-2}t^{-1/\alpha}\Big[\phi_{\alpha,t}\Big(\gamma^{-1}t^{-1/\alpha}(x-\beta t+\gamma c_t)\Big)$$
$$+ \frac{x-\beta t}{\gamma t^{1/\alpha}}\phi'_{\alpha,t}\Big(\gamma^{-1}t^{-1/\alpha}(x-\beta t+\gamma c_t)\Big)\Big]$$
$$= -\gamma^{-2}t^{-1/\alpha}\Big[\phi_{\alpha,t}(z)+z\phi'_{\alpha,t}(z)-c_tt^{-1/\alpha}\phi'_{\alpha,t}(z)\Big]_{z=\gamma^{-1}t^{-1/\alpha}(x-\beta t+\gamma c_t)}.$$

Because $X_t = Y_t + U_t$ and Y, U are independent, we have

$$p_t(\theta;x) = \int_{\mathbb{R}} p_t^Y(\theta;x-t^{1/\alpha}y)\nu_t(dy), \tag{3.3}$$

where ν_t denotes the law of $t^{-1/\alpha}U_t$.

Taking (3.2) into account, we can re-arrange the above convolution formula for $p_t(\theta;x)$ in the following way:

$$p_t(\theta;x) = \gamma^{-1}t^{-1/\alpha}f_t\Big(\theta;\gamma^{-1}t^{-1/\alpha}(x-\beta t+\gamma c_t)\Big), \tag{3.4}$$

where

$$f_t(\theta;z) = \int_{\mathbb{R}}\phi_{\alpha,t}(z-\gamma^{-1}y)\nu_t(dy).$$

Note that (2.8) implies $\nu_t \Rightarrow \delta_0, t \to 0$, and γ is separated from 0 when $\theta = (\beta,\gamma) \in \tilde\Theta$. Then by the first assertion in (3.1), for every $N > 0$

$$\sup_{\theta\in\tilde\Theta}\sup_{|z|\leq N}|f_t(\theta;z)-\phi_{\alpha,C_\pm}(z)| \to 0, \quad t \to 0+. \tag{3.5}$$

It follows from (3.3) that

$$\partial_\beta p_t(\theta;x) = \int_{\mathbb{R}}\partial_\beta p_t^Y(\theta;x-t^{1/\alpha}y)\nu_t(dy), \quad \partial_\gamma p_t(\theta;x) = \int_{\mathbb{R}}\partial_\gamma p_t^Y(\theta;x-t^{1/\alpha}y)\nu_t(dy)$$

Then, similarly as above, we have

$$\partial_\beta p_t(\theta;x) = -\gamma^{-2}t^{1-2/\alpha}f_t^{(1)}\Big(\theta;\gamma^{-1}t^{-1/\alpha}(x-\beta t+\gamma c_t)\Big),$$

$$\partial_\gamma p_t(\theta;x) =$$
$$-\gamma^{-2}t^{-1/\alpha}\Big[f_t(\theta;z)+f_t^{(2)}(\theta;z)-c_tt^{-1/\alpha}f_t^{(1)}(\theta;z)\Big]_{z=\gamma^{-1}t^{-1/\alpha}(x-\beta t+\gamma c_t)} \tag{3.6}$$

with

$$f_t^{(1)}(\theta;z) = \int_{\mathbb{R}}\phi'_{\alpha,t}(z-\gamma^{-1}y)\nu_t(dy), \quad f_t^{(2)}(\theta;z) = \int_{\mathbb{R}}(z-\gamma^{-1}y)\phi'_{\alpha,t}(z-\gamma^{-1}y)\nu_t(dy),$$

and for every $N > 0$

$$\sup_{\theta\in\tilde\Theta}\sup_{|z|\leq N}|f_t^{(1)}(\theta;z)-\phi'_{\alpha,C_\pm}(z)| \to 0, \quad \sup_{\theta\in\tilde\Theta}\sup_{|z|\leq N}|f_t^{(2)}(\theta;z)-z\phi'_{\alpha,C_\pm}(z)| \to 0, \quad t \to 0+. \tag{3.7}$$

Below we use formulae (3.4) – (3.7) to control the "local" behavior of the functions g_t, q_t involved into **A1**–**A4**. To control the "global" behavior, we use the following moment bound; we evaluate this bound in Section 4 below.

Proposition 3.2. *Under conditions of Theorem 2.1, for every $\tilde{\Theta}$ and every $\delta_1 \in (0, \delta)$ (where δ comes from H2),*

$$\sup_{n \geq 1, \theta \in \tilde{\Theta}} \mathsf{E} \left| \tilde{r}(n)^\top g_{h_n} \left(\theta; X_{h_n}^\theta \right) \right|^{2+\delta_1} < \infty.$$

3.2. Proofs of *A1*–*A4*.

3.2.1. *Proof of A1.* For this it is sufficient to show that, for a fixed $t = h_n$, the mapping

$$\Theta \ni \theta \mapsto q_t(\theta; \cdot) \in L_2(\mathbb{R}) \tag{3.8}$$

is continuous, and for any θ_1, θ_2 such that the segment $[\theta_1, \theta_2]$ is contained in Θ,

$$\sqrt{p_t(\theta_1; \cdot)} - \sqrt{p_t(\theta_2; \cdot)} = \left(\int_0^1 q_t((1-s)\theta_1 + s\theta_2; \cdot) \, ds \right)^\top (\theta_1 - \theta_2) \tag{3.9}$$

with the integral understood in the sense of convergence of the Riemann sums in $L_2(\mathbb{R})$. The argument here is similar and simpler to the one used in the proof of Theorem 2 in Ivanenko and Kulik (2015), hence we just outline the main steps. Define

$$\Psi_\varepsilon(z) = \begin{cases} 0, & z < \varepsilon/2, \\ \frac{(z-\varepsilon/2)^2}{2\varepsilon^{3/2}}, & z \in [\varepsilon/2, \varepsilon], \\ \sqrt{z} - \frac{7\sqrt{\varepsilon}}{8}, & z \geq \varepsilon. \end{cases}$$

Then, by the construction, for $z > 0$

$$\Psi_\varepsilon(z) \to \Psi_0(z) := \sqrt{z}, \quad \Psi_\varepsilon'(z) \to \Psi_0'(z) = \frac{1}{2\sqrt{z}}, \quad \varepsilon \to 0.$$

Because $\Psi_\varepsilon \in C^1, \varepsilon > 0$, we have by Proposition 3.1 and (3.4), (3.6) that $\Psi_\varepsilon(p_t(\theta; x))$ depends smoothly on θ, x and

$$q_{t,\varepsilon}(\theta; x) := \nabla_\theta \Big(\Psi_\varepsilon(p_t(\theta; x)) \Big) = \Psi_\varepsilon'(p_t(\theta; x)) \nabla_\theta p_t(\theta; x).$$

Then

$$\Psi_\varepsilon \Big(p_t(\theta_1; \cdot) \Big) - \Psi_\varepsilon \Big(p_t(\theta_2; \cdot) \Big) = \left(\int_0^1 q_{t,\varepsilon}((1-s)\theta_1 + s\theta_2; \cdot) \, ds \right)^\top (\theta_1 - \theta_2),$$

and to prove (3.9) it is sufficient to prove that

$$\Psi_\varepsilon \Big(p_t(\theta; \cdot) \Big) \to \Psi_0 \Big(p_t(\theta; \cdot) \Big) = \sqrt{p_t(\theta; \cdot)}, \quad \varepsilon \to 0$$

in $L_2(\mathbb{R})$ for every $\theta \in \Theta$, and that

$$q_{t,\varepsilon}(\theta; \cdot) \to q_t(\theta; \cdot), \quad \varepsilon \to 0$$

in $L_2(\mathbb{R})$ uniformly in $\theta \in [\theta_1, \theta_2]$. The latter would also provide that the function (3.8) is continuous as a uniform limit of continuous functions. Let us prove the second convergence, since the proof of the first one is similar and simpler. By the construction, we have $0 \leq \Psi_\varepsilon'(z) \leq \Psi_0'(z) = (2\sqrt{z})^{-1}$, therefore

$$q_{t,\varepsilon}(\theta; x) = \Psi_\varepsilon'(p_t(\theta; x)) \nabla_\theta p_t(\theta; x) = \Upsilon_\varepsilon(p_t(\theta; x)) g_t(\theta; x),$$

where

$$\Upsilon_\varepsilon(z) = \Psi_\varepsilon'(z) z \quad \begin{cases} \leq (1/2)\sqrt{z}, & z > 0; \\ = (1/2)\sqrt{z}, & z \geq \varepsilon. \end{cases}$$

Hence

$$\int_{\mathbb{R}} (q_{t,\varepsilon}(\theta; x) - q_t(\theta; x))^2 \, dx \leq \frac{1}{4} \int_{\{x : p_t(\theta; x) \leq \varepsilon\}} \left(g_t(\theta; x) \right)^2 p_t(\theta; x) \, dx.$$

Recall that $t = h_n$. Take $\tilde{\Theta}$ in Proposition 3.2 equal to the segment $[\theta_1, \theta_2]$, then

$$\sup_{\theta \in \tilde{\Theta}} \int_{\mathbb{R}} \left| g_t(\theta; x) \right|^{2 + \delta_1} p_t(\theta; x) \, dx = \sup_{\theta \in \tilde{\Theta}} \mathsf{E} \left| g_t(\theta; X_t) \right|^{2 + \delta_1} < \infty.$$

Hence by the Hölder inequality

$$\int_{\mathbb{R}} (q_{t,\varepsilon}(\theta; x) - q_t(\theta; x))^2 \, dx \leq C \left(\int_{\{x : p_t(\theta; x) \leq \varepsilon\}} p_t(\theta; x) \, dx \right)^{\delta_1 / (2 + \delta_1)}.$$

The density $p_t(\theta; x)$ is given explicitly by (3.4). Using this representation and changing the variables $z = \gamma^{-1} t^{-1/\alpha}(x - \beta t + \gamma c_t)$, we get

$$\int_{\{x : p_t(\theta; x) \leq \varepsilon\}} p_t(\theta; x) \, dx = \int_{\{z : f(\theta; z) \leq \gamma t^{1/\alpha} \varepsilon\}} f(\theta; z) \, dz.$$

We have $\phi_{\alpha, t} \in L_1(\mathbb{R})$, and therefore the mapping

$$[\theta_1, \theta_2] \ni \theta = (\beta, \gamma) \mapsto \phi_{\alpha, t}(\cdot - \gamma^{-1} y) \in L_1(\mathbb{R})$$

is continuous. Hence the mapping

$$[\theta_1, \theta_2] \ni \theta \mapsto f(\theta; \cdot) = \int_{\mathbb{R}} \phi_{\alpha, t}(\cdot - \gamma^{-1} y) \nu_t(dy)$$

is continuous, as well. This finally implies that

$$\int_{\{x : p_t(\theta; x) \leq \varepsilon\}} p_t(\theta; x) \, dx = \int_{\{z : f(\theta; z) \leq \gamma t^{1/\alpha} \varepsilon\}} f(\theta; z) \, dz \to 0, \quad \varepsilon \to 0$$

uniformly in $\theta \in [\theta_1, \theta_2]$, which completes the proof of the required convergence and provides **A1**.

3.2.2. *Proof of* **A2**. Denote

$$\Gamma_{k,n}^{\theta} = \tilde{r}(n)^{\top} g_{h_n} \left(\theta; X_{kh_n}^{\theta} - X_{(k-1)h_n}^{\theta} \right), \quad k = 1, \ldots, n,$$

then

$$\sum_{k=1}^{n} \left(r(n)^{\top} g_{h_n} \left(\theta; X_{kh_n}^{\theta} - X_{(k-1)h_n}^{\theta} \right) \right)^{\otimes 2} = \frac{1}{n} \sum_{k=1}^{n} (\Gamma_{k,n}^{\theta})^{\otimes 2}.$$

Since X is a Lévy process, $\{\Gamma_{k,n}^{\theta}\}_{1 \leq k \leq n}$ is a triangular array of random vectors, which are row-wise independent and identically distributed. Let us analyze the common law of $\Gamma_{k,n}^{\theta}$ at an n-th row.

Denote

$$\xi_{k,n}^{\theta} = \gamma^{-1} h_n^{-1/\alpha}(X_{kh_n}^{\theta} - X_{(k-1)h_n}^{\theta} - \beta h_n + \gamma c_{h_n}), \quad k = 1, \ldots, n,$$

which are i.i.d. random variables with

$$\xi_{1,n}^{\theta} \overset{d}{=} \zeta_{\alpha, h_n} + \gamma^{-1} h_n^{-1/\alpha} U_{h_n}.$$

By statement (1) of Proposition 3.1 and (2.8), we have then

$$\xi_{1,n}^{\theta} \Rightarrow Z_1^{\alpha, C_{\pm}}. \tag{3.10}$$

Next, by (3.4), (3.6) the components of $g_{h_n}(\theta; X^\theta_{kh_n} - X^\theta_{(k-1)h_n})$ are given by

$$g^1_{h_n}\left(\theta; X^\theta_{kh_n} - X^\theta_{(k-1)h_n}\right) = -\gamma^{-1} h_n^{1-1/\alpha} \frac{f^{(1)}_{h_n}(\theta; \xi^\theta_{k,n})}{f_{h_n}(\theta; \xi^\theta_{k,n})},$$

$$g^2_{h_n}\left(\theta; X^\theta_{kh_n} - X^\theta_{(k-1)h_n}\right) = -\gamma^{-1}\left[1 + \frac{f^{(2)}_{h_n}(\theta; \xi^\theta_{k,n})}{f_{h_n}(\theta; \xi^\theta_{k,n})}\right] + \gamma^{-1} c_{h_n} h_n^{-1/\alpha} \frac{f^{(1)}_{h_n}(\theta; \xi^\theta_{k,n})}{f_{h_n}(\theta; \xi^\theta_{k,n})}.$$

Recall that

$$\tilde{r}(n)^\top = \begin{pmatrix} h_n^{1/\alpha - 1} & 0 \\ c_{h_n} h_n^{-1} & 1 \end{pmatrix},$$

hence we can write finally

$$\Gamma^\theta_{k,n} = \gamma^{-1} G_{\alpha, h_n}(\theta; \xi^\theta_{k,n}),$$

where the vector-valued functions G_{α, h_n} have the components

$$G^1_{\alpha, h_n}(\theta; x) = -\frac{f^{(1)}_{h_n}(\theta; x)}{f_{h_n}(\theta; x)}, \quad G^2_{\alpha, h_n}(\theta; x) = -1 - \frac{f^{(2)}_{h_n}(\theta; x)}{f_{h_n}(\theta; x)}.$$

Denote by $G_{\alpha, C_\pm}$ the vector-valued function with the components

$$G^1_{\alpha, C_\pm}(x) = -\frac{\phi'_{\alpha, C_\pm}(x)}{\phi_{\alpha, C_\pm}(x)}, \quad G^2_{\alpha, C_\pm}(x) = -1 - \frac{x \phi'_{\alpha, C_\pm}(x)}{\phi_{\alpha, C_\pm}(x)},$$

and denote for $\varepsilon > 0, N > 0$

$$K_{\varepsilon, N} = \{x : |x| \le N, \phi_{\alpha, C_\pm}(x) \ge \varepsilon\},$$

which is a compact set in $\mathbb{R}$. It follows from (3.5), (3.7) that, for every fixed $\theta \in \Theta, \varepsilon > 0, N > 0$,

$$G_{\alpha, t}(\theta; x) \to G_{\alpha, C_\pm}(x), \quad t \to 0+$$

uniformly with respect to $x \in K_{\varepsilon, N}$. By (3.10) and continuity of the limit distribution,

$$\limsup_{n \to \infty} \mathsf{P}(\xi^\theta_{1,n} \notin K_{\varepsilon, N}) \le \mathsf{P}(Z^{\alpha, C_\pm}_1 \notin K_{\varepsilon, N}).$$

We also have

$$\mathsf{P}(Z^{\alpha, C_\pm}_1 \notin K_{\varepsilon, N}) \to 0, \quad \varepsilon \to 0, \quad N \to \infty.$$

Let us summarize: the random vectors $\Gamma^\theta_{k,n}$ are represented as images of the i.i.d. random variables $\xi^\theta_{k,n}$ under the functions $G_{\alpha, h_n}(\theta; \cdot)$, and

- the common law of $\xi^\theta_{k,n}$ weakly converge to the law of $Z^{\alpha, C_\pm}_1$;
- on every compact set $K_{\varepsilon, N}$, the functions $G_{\alpha, h_n}(\theta; \cdot)$ converge uniformly to the function $G_{\alpha, C_\pm}$ which is continuous on this compact;
- by choosing $\varepsilon > 0$ small and $N > 0$ large, the probability for $\xi^\theta_{k,n} \in K_{\varepsilon, N}$ can be made arbitrarily close to 1.

Because the weak convergence is preserved by continuous mappings, we deduce from the above that the common law of $\Gamma^\theta_{k,n}, k = 1, \ldots, n$ weakly converge as $n \to \infty$ to the law of $\Gamma^\theta = \gamma^{-1} G_{\alpha, C_\pm}(Z^{\alpha, C_\pm}_1)$. On the other hand Proposition 3.2 yields that the family $\{(\Gamma^\theta_{k,n})^{\otimes 2}\}$ is uniformly integrable, hence by the Law of Large Numbers for independent random variables

$$\mathsf{E}\left|\frac{1}{n}\sum_{k=1}^n (\Gamma^\theta_{k,n})^{\otimes 2} - \mathsf{E}(\Gamma^\theta)^{\otimes 2}\right| \to 0, \quad n \to \infty.$$

Observe that it is an easy calculation to show that the covariation matrix for Γ^θ equals $\Sigma(\theta)$ given by the second identity in (2.9). Taking $\theta = \theta_0$, we complete the proof of **A2**.

3.2.3. *Proof of **A3***. Because

$$n \int_{\mathbb{R}} |r(n)g_{h_n}(\theta; y)|^{2+\delta_1} p_{h_n}(\theta; y) \, dy = n^{-\delta_1/2} \mathsf{E} \left| \tilde{r}(n) g_{h_n}\left(\theta; X_{h_n}^\theta\right) \right|^{2+\delta_1},$$

assertion **A3** follows from Proposition 3.2 immediately.

3.2.4. *Proof of **A4***. We have

$$q_t(\theta; x) = \frac{1}{2} g_t(\theta; x) \sqrt{p_t(\theta; x)},$$

hence for any $\theta_1, \theta_2 \in \Theta$

$$n \int_{\mathbb{R}} \left| r(n)^\top \left(q_{h_n}(\theta_2; x) - q_{h_n}(\theta_1; x) \right) \right|^2 \, dx$$

$$= \frac{1}{4} \int_{\mathbb{R}} \left| \tilde{r}(n)^\top g_{h_n}(\theta_2; x) \sqrt{\frac{p_{h_n}(\theta_2; x)}{p_{h_n}(\theta_1; x)}} - \tilde{r}(n)^\top g_{h_n}(\theta_1; x) \right|^2 p_{h_n}(\theta_1; x) \, dx$$

$$= \frac{1}{4} \mathsf{E} \left| \frac{1}{\gamma_2} G_{\alpha,h_n}(\theta_2, \xi_{1,n}^{\theta_1}) \sqrt{\frac{f_{h_n}(\theta_2; \xi_{1,n}^{\theta_1})}{f_{h_n}(\theta_1; \xi_{1,n}^{\theta_1})}} - \frac{1}{\gamma_1} G_{\alpha,h_n}(\theta_1, \xi_{1,n}^{\theta_1}) \right|^2 ;$$

we keep using the notation introduced in the proof of **A2**. Take $\theta_1 = \theta_0$, and let $\theta_n = (\beta_n, \gamma_n) \to \theta_0$ be arbitrary sequence. It follows from (3.5), (3.7) that for every fixed $\varepsilon > 0, N > 0$

$$\sup_{x \in K_{\varepsilon,N}} \left| \frac{1}{\gamma_n} G_{\alpha,h_n}(\theta_n, x) \sqrt{\frac{f_{h_n}(\theta_n; x)}{f_{h_n}(\theta_0; x)}} - \frac{1}{\gamma_0} G_{\alpha,h_n}(\theta_0, x) \right| \to 0, \quad n \to \infty.$$

Then, by the Cauchy inequality $(a+b)^2 \le 2a^2 + 2b^2$,

$$\limsup_{n \to \infty} n \int_{\mathbb{R}} \left| r(n)^\top \left(q_{h_n}(\theta_n; x) - q_{h_n}(\theta_0; x) \right) \right|^2 \, dx$$

$$\le \limsup_{n \to \infty} \left(\frac{1}{2\gamma_n^2} \mathsf{E} \left| G_{\alpha,h_n}(\theta_n, \xi_{1,n}^{\theta_0}) \right|^2 \frac{f_{h_n}(\theta_n; \xi_{1,n}^{\theta_0})}{f_{h_n}(\theta_0; \xi_{1,n}^{\theta_0})} 1_{\xi_{1,n}^{\theta_0} \notin K_{\varepsilon,N}} \right.$$

$$\left. + \frac{1}{2\gamma_0^2} \mathsf{E} \left| G_{\alpha,h_n}(\theta_0, \xi_{1,n}^{\theta_0}) \right|^2 1_{\xi_{1,n}^{\theta_0} \notin K_{\varepsilon,N}} \right)$$

$$= \limsup_{n \to \infty} \left(\frac{1}{2\gamma_n^2} \mathsf{E} \left| G_{\alpha,h_n}(\theta_n, \xi_{1,n}^{\theta_n}) \right|^2 1_{\xi_{1,n}^{\theta_n} \notin K_{\varepsilon,N}} \right.$$

$$\left. + \frac{1}{2\gamma_0^2} \mathsf{E} \left| G_{\alpha,h_n}(\theta_0, \xi_{1,n}^{\theta_0}) \right|^2 1_{\xi_{1,n}^{\theta_0} \notin K_{\varepsilon,N}} \right).$$

Since

$$\gamma^{-1} G_{\alpha,h_n}(\theta; \xi_{1,n}^\theta) = \Gamma_{1,n}^\theta = \tilde{r}(n)^\top g_{h_n}\left(\theta; X_{kh_n}^\theta - X_{(k-1)h_n}^\theta\right),$$

we deduce using Proposition 3.2 and the Hölder inequality that

$$\limsup_{n\to\infty} n \int_{\mathbb{R}} \left| r(n)^\top \left(q_{h_n}\left(\theta_n; x\right) - q_{h_n}\left(\theta_0; x\right)\right)\right|^2 \, dx$$

$$\leq C \limsup_{n\to\infty} \left(\mathsf{P}(\xi_{1,n}^{\theta_n} \notin K_{\varepsilon,N})^{\delta_1/(2+\delta_1)} + \mathsf{P}(\xi_{1,n}^{\theta_0} \notin K_{\varepsilon,N})^{\delta_1/(2+\delta_1)}\right)$$

with some constant C. We have

$$\xi_{1,n}^{\theta_n} \Rightarrow Z_1^{\alpha,C\pm}, \quad \xi_{1,n}^{\theta_0} \Rightarrow Z_1^{\alpha,C\pm}, \quad n \to \infty,$$

hence we get

$$\limsup_{n\to\infty} n \int_{\mathbb{R}} \left| r(n)\left(q_{h_n}\left(\theta_n; x\right) - q_{h_n}\left(\theta_0; x\right)\right)\right|^2 \, dx \leq 2C\mathsf{P}(Z_1^{\alpha,C\pm} \notin K_{\varepsilon,N})^{\delta_1/(2+\delta_1)}.$$

Recall that $\varepsilon > 0, N > 0$ here are arbitrary. Taking in the above inequality $\varepsilon \to 0, N \to \infty$ we get finally

$$\limsup_{n\to\infty} n \int_{\mathbb{R}} \left| r(n)\left(q_{h_n}\left(\theta_n; x\right) - q_{h_n}\left(\theta_0; x\right)\right)\right|^2 \, dx = 0,$$

which completes the proof of **A4**. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3.3. *Outline of the proof of Theorem 2.2.* To get the required LAN property uniformly in $U \in \mathfrak{U}$, it is enough to fix a sequence U^n of Lévy processes, such that

$$h_n^{-1/\alpha} U_{h_n}^n \to 0$$

in probability, and repeat the above argument for a modified statistical model, where the process X^θ in the n-th sample is replaced by

$$X_t^{\theta,n} = \beta t + \gamma Z_t + U_t^n.$$

The moment bound in Proposition 3.2 is, to a very high extent, insensitive with respect to the process U; in particular, we will show in Section 4 that

$$\sup_{n\geq 1, \theta\in\tilde{\Theta}} \mathsf{E} \left| \tilde{r}(n) g_{h_n} \left(\theta; X_{h_n}^{\theta,n}\right)\right|^{2+\varepsilon} < \infty. \qquad\qquad (3.11)$$

The law of the process U is involved in the definition of the functions $f, f^{(1)}, f^{(2)}$, but it is straightforward to see that the relations (3.5), (3.7) in fact hold true uniformly with respect to $U \in \mathfrak{U}$. Finally, the random variables

$$\xi_{k,n}^{\theta,n} = \gamma^{-1} h_n^{-1/\alpha}(X_{kh_n}^{\theta,n} - X_{(k-1)h_n}^{\theta,n} - \beta h_n + \gamma c_{h_n}) \overset{d}{=} \zeta_{\alpha,h_n} + \gamma^{-1} h_n^{-1/\alpha} U_{h_n}^n$$

weakly converge to $Z_1^{\alpha,C\pm}$. Hence repeating, with obvious notational changes, the calculations from Section 3.2, we get properties **A1** – **A4** for the modified model, which proves the required LAN property, uniform in $U \in \mathfrak{U}$. $\qquad\qquad\square$

4. Malliavin calculus-based integral representation for the derivative of the log-likelihood function and related L_p-bounds

Our main aim in this section is to prove Proposition 3.2, which is the cornerstone of the proof of Theorem 2.1. With this purpose in mind, we give an integral representation for the derivative of the log-likelihood function by means of a certain version of the Malliavin calculus. In the diffusive case, such a representation was developed by Gobet (2001, 2002); see also Corcuera and Kohatsu-Higa (2011). In

the Lévy setting, the choice of a particular design of such a calculus is a non-trivial problem, and we discuss it in details below.

4.1. *The main statement: formulation, discussion, and an outline of the proof.* In what follows, $\nu(ds, du)$ and $\tilde{\nu}(ds, du) = \nu(ds, du) - ds\mu(du)$ are, respectively, the Poisson point measure and the compensated Poisson measure from the Lévy-Itô representation of Z:

$$Z_t = \int_0^t \int_{|u|>1} u\nu(ds, du) + \int_0^t \int_{|u|\leq 1} u\tilde{\nu}(ds, du).$$

We denote

$$D_t X_t^\theta = \gamma D_t Z_t = \gamma \int_0^t \int_{\mathbb{R}} u^2 \nu(ds, du), \quad D_t^2 X_t^\theta = 2\gamma \int_0^t \int_{\mathbb{R}} u^3 \nu(ds, du),$$

the genealogy of the notation will become clear later. To define the stochastic integrals with respect to ν properly, we decompose them in two integrals (which is a usual trick). The "small jump" parts, which correspond to values $u \in [-1, 1]$, are well defined because the functions u^2, u^3 are integrable with respect to μ on $[-1, 1]$; respective integrals with respect to ν are understood in L_1 sense. The "large jump" parts of the integrals with respect to ν are understood in the path-wise way, i.e. as sums over finite set of jumps. Next, we denote

$$\chi(u) = -u^2 \frac{m'(u)}{m(u)} - 2u$$

and put

$$\delta_t(1) = \int_0^t \int_{|u|\leq u_0} \chi(u)\tilde{\nu}(ds, du) + \int_0^t \int_{|u|>u_0} \chi(u)\nu(ds, du) + tu_0^2 \Big[m(u_0) - m(-u_0) \Big],$$

where u_0 comes from the condition **H2**. By **H2**, $|\chi(u)| \leq C|u|$ for $|u| \leq u_0$, hence the "small jump" integral above is well defined in L_2 sense; the "large jump" integral is understood in the path-wise sense.

We define the *modified Malliavin weight* (we postpone for a while the explanation of the terminology) as the vector $\Xi_t^\theta = (\Xi_t^\beta, \Xi_t^\gamma)^\top$ with

$$\Xi_t^\beta = \frac{t\delta_t(1)}{D_t X_t^\theta} + \frac{tD_t^2 X_t^\theta}{(D_t X_t^\theta)^2}, \quad \Xi_t^\gamma = \frac{Z_t \delta_t(1)}{D_t X_t^\theta} + \frac{Z_t D_t^2 X_t^\theta}{(D_t X_t^\theta)^2} - \frac{1}{\gamma}. \tag{4.1}$$

Denote by $\mathsf{E}_{x,y}^{t,\theta}$ the expectation with respect to the law of the *bridge* of the process X^θ conditioned by $X_0^\theta = x, X_t^\theta = y$. Note that because the process X^θ possesses a continuous transition probability density $p_t(\theta; x, y) = p_t(\theta; y - x)$, the law of the bridge is well defined for any t, x, y such that $p_t(\theta; x, y) > 0$ (cf. Chaumont and Uribe Bravo (2011)).

The main statement in this section is the following.

Theorem 4.1.   (1) *Let δ be the same as in **H2**. Then for every $\delta_1 \in (0, \delta)$ and every $\tilde{\Theta}$,*

$$\sup_{\theta \in \tilde{\Theta}} \sup_{n \geq 1} \mathsf{E} \Big| \tilde{r}(n)^\top \Xi_{h_n}^\theta \Big|^{2+\delta_1} < \infty. \tag{4.2}$$

(2) *The following intergal representation formula holds true:*

$$g_t(\theta; x) = \begin{cases} \mathsf{E}_{0,x}^{t,\theta} \Xi_t^\theta, & p_t(\theta; x) > 0, \\ 0, & otherwise. \end{cases} \tag{4.3}$$

It follows from (4.3) that

$$\tilde{r}(n)^\top g_t(\theta; X_{h_n}^\theta) = \mathsf{E}\left[\tilde{r}(n)^\top \Xi_{h_n}^\theta \,\Big|\, X_{h_n}^\theta\right].$$

Note that the explicit formula for the weight Ξ_t^θ does not involve the "nuisance noise" U at all, and the dependence of $p_t(\theta; X_{h_n}^\theta)$ on U is contained in the operation of the conditional expectation, only. Then by the Jensen inequality

$$\mathsf{E}\left|\tilde{r}(n)^\top g_t(\theta; X_{h_n}^\theta)\right|^{2+\delta_1} \le \mathsf{E}\left|\tilde{r}(n)^\top \Xi_{h_n}^\theta\right|^{2+\delta_1},$$

where U is not involved in the right hand side term. Hence Theorem 4.1 immediately yields both Proposition 3.2 and the moment bound (3.11).

Let us explain the main idea which Theorem 4.1 is based on. By analogy to Gobet's results in the diffusive case (Gobet (2001, 2002)), one can naturally expect that an integral representation of the form (4.3) could be obtained by means of a proper version of the Malliavin calculus for jump processes. Two possible ways to do that were developed in Clément and Gloter (2015) and Ivanenko and Kulik (2015), being in fact close to each other and, heuristically, being based on "infinitesimal perturbation of the jump configuration" with an intensity function ρ which is a "compactly" supported smooth function (see a more detailed exposition in Section 4.2 below). Using either of these two approaches it is possible to prove an analogue of (4.3) with Ξ_t^θ being replaced by some $\Xi_{t,\rho}^\theta$ which, in full analogy to Gobet's approach, has the meaning of the *Malliavin weight* (see formula (4.12) below). However, we then encounter following two difficulties, both being related with moment bounds for the corresponding terms.

- In order to provide that $\Xi_{t,\rho}^\gamma$ is square integrable (which is necessary for $\Xi_{t,\rho}^\theta$ to have representation (4.12) with the Skorokhod integral in the right hand side), we need an additional moment bound for Z: for some $\delta' > 0$,

$$\int_{|u| \ge 1} |u|^{2+\delta'} \mu(du) < \infty. \tag{4.4}$$

  This excludes from the consideration "heavy tailed" Lévy processes, e.g. the particularly important α-stable process.
- Even if we confine ourselves by the class of "light tailed" Lévy processes satisfying (4.4), we can not obtain analogue of the moment bound (4.2) for $\Xi_{t,\rho}^\theta$: namely, respective upper bound for $L_{2+\delta_1}$-norm of $\Xi_{t,\rho}^\theta$ would explode as $t \to 0+$: see Remark 4.1 below for detail.

Both of these "moment" difficulties are resolved when we put formally

$$\rho(u) = u^2, \quad u \in \mathbb{R},$$

in the formula for $\Xi_{t,\rho}^\theta$; see Section 4.2 below, and especially Remark 4.1 which explains the heuristics behind the particular choice $\rho(u) = u^2$. This explains both the name "the modified Malliavin weight" we have used for the term defined by (4.1), and the background for the notation D_t, δ_t: we take the explicit formulae for the Malliavin derivative $\mathrm{D}_{t,\rho}$ and respective Skorokhod integral $\delta_{t,\rho}$, and put therein

$\rho(u) = u^2$. Note that because $\rho(u) = u^2$ is not compactly supported, $D_t X_t^\theta$ may fail to be square integrable, which means that now $D_t X_t^\theta$ can not be interpreted as a Malliavin derivative. As a consequence, now one can not apply the Malliavin calculus tools to prove (4.3) directly. Hence we will use the following three step procedure to prove (4.3):

(1) first, we apply Malliavin calculus tools to prove analogue of (4.3) for $\Xi_{t,\rho}^\theta$ with compactly supported ρ under the additional moment condition (4.4);

(2) second, we approximate $\rho(u) = u^2$ by a sequence of compactly supported ρ's;

(3) finally, we approximate general Z by a sequence of "light tailed" Lévy processes $Z^L, L \geq 1$, each of them satisfying (4.4).

4.2. *Proof of (4.2): Moment bound.* First, we give an explicit expression for $\tilde{r}(n)^\top \Xi_t^\theta$. Denote

$$\tilde{Z}_t = Z_t + c_t = \int_0^t \int_{|u|>t^{1/\alpha}} u\nu(ds, du) + \int_0^t \int_{|u|\leq t^{1/\alpha}} u\tilde{\nu}(ds, du),$$

then

$$\tilde{r}(n)^\top \Xi_t^\theta = \left(\frac{t^{1/\alpha}\delta_t(1)}{D_t X_t^\theta} + \frac{t^{1/\alpha}D_t^2 X_t^\theta}{(D_t X_t^\theta)^2}, \frac{\tilde{Z}_t \delta_t(1)}{D_t X_t^\theta} + \frac{\tilde{Z}_t D_t^2 X_t^\theta}{(D_t X_t^\theta)^2} - \frac{1}{\gamma} \right)^\top. \tag{4.5}$$

We will conclude the required bound (4.2) from a sequence of auxiliary estimates for the terms involved in the explicit expression (4.5).

Denote

$$\kappa_t = \int_0^t \int_{|u|\leq t^{1/\alpha}} u^2 \nu(ds, du).$$

Lemma 4.1. *For every $p \geq 1$ there exists $C_p < \infty$ such that*

$$\mathsf{E}(t^{-2/\alpha}\kappa_t)^{-p} \leq C_p, \quad t \in (0, 1].$$

Proof: For any $\varepsilon < 1$ we have

$$\mathsf{P}(\kappa_t < \varepsilon^2 t^{2/\alpha}) \leq P\Big(\nu([0,t] \times \{|u| \in [\varepsilon t^{1/\alpha}, t^{1/\alpha}]\}) = 0 \Big)$$

$$= \exp\Big\{ -t\mu(|u| \in [\varepsilon t^{1/\alpha}, t^{1/\alpha}]) \Big\}.$$

Condition **H1** yields that, with some positive constant C,

$$t\mu(|u| \in [\varepsilon t^{1/\alpha}, t^{1/\alpha}]) \geq Ct \int_{\varepsilon t^{1/\alpha}}^{t^{1/\alpha}} \alpha|u|^{-\alpha-1} du = C(\varepsilon^{-\alpha} - 1), \quad t \in (0, 1].$$

Hence for the family of random variables $t^{-2/\alpha}\kappa_t, t \in (0, 1]$ we have the uniform bound

$$\mathsf{P}(t^{-2/\alpha}\kappa_t < \varepsilon^2) \leq e^{-C\varepsilon^{-\alpha}+C}, \quad \varepsilon < 1, \quad t \leq 1,$$

which proves the required statement. $\qquad\qquad\square$

Because

$$D X_t^\theta = \gamma \int_0^t \int_{\mathbb{R}} u^2 \nu(ds, du) \geq \gamma\kappa_t,$$

Lemma 4.1 immediately gives the following: for every $p \geq 1$,

$$\sup_{\theta \in \tilde{\Theta}, t \in (0,1]} \mathsf{E} \left(\frac{t^{1/\alpha}}{\sqrt{\mathrm{D}_t X_t^\theta}} \right)^p < \infty. \tag{4.6}$$

Next, observe that both $\mathrm{D}_t X_t^\theta$ and $\mathrm{D}_t^2 X_t^\theta$ are represented as sums over the set of jumps of the process Z. Because

$$\left(\sum_i a_i \right)^{3/2} \geq \sum_i a_i^{3/2}, \quad \{a_i\} \subset [0, \infty),$$

we have

$$\left| \frac{\mathrm{D}_t^2 X_t^\theta}{(\mathrm{D}_t X_t^\theta)^{3/2}} \right| \leq \frac{2}{\sqrt{\gamma}}. \tag{4.7}$$

Lemma 4.2. *For every $p \geq 1$,*

$$\sup_{\theta \in \tilde{\Theta}, t \in (0,1]} \mathsf{E} \left(\frac{\tilde{Z}_t}{\sqrt{\mathrm{D}_t X_t^\theta}} \right)^p < \infty. \tag{4.8}$$

Proof: We have

$$\tilde{Z}_t = \int_0^t \int_{|u| \leq t^{1/\alpha}} u \tilde{\nu}(ds, du) + \int_0^t \int_{|u| > t^{1/\alpha}} u \nu(ds, du) =: \xi_t + \zeta_t,$$

$$\mathrm{D} X_t^\theta = \gamma \int_0^t \int_{\mathbb{R}} u^2 \, \nu(ds, du) = \gamma \left(\kappa_t + \eta_t \right), \quad \eta_t = \int_0^t \int_{|u| > t^{1/\alpha}} u^2 \, \nu(ds, du)$$

(κ_t is already defined above). Then

$$\left| \frac{\tilde{Z}_t}{\sqrt{\mathrm{D} X_t^\theta}} \right| \leq \frac{1}{\gamma} \left(\frac{|\xi_t|}{\sqrt{\kappa_t + \eta_t}} + \frac{|\zeta_t|}{\sqrt{\kappa_t + \eta_t}} \right) \leq \frac{1}{\gamma} \left(\frac{|\xi_t|}{\sqrt{\kappa_t}} + \frac{|\zeta_t|}{\sqrt{\eta_t}} \right). \tag{4.9}$$

By Lemma 4.1, the family

$$\left\{ \frac{t^{1/\alpha}}{\sqrt{\kappa_t}} \right\}_{t \in (0,1]}$$

has bounded L_p-norms for any $p \geq 1$. In addition, the family

$$\left\{ t^{-1/\alpha} \xi_t \right\}_{t \in [0,1]}$$

also has bounded L_p-norms for any $p \geq 1$. To see this, observe that ξ_t is an integral of a deterministic function over a compensated Poisson point measure, and therefore its exponential moments can be expressed explicitly:

$$\mathsf{E} \exp(c \xi_t) = \exp \left[t \int_{|u| \leq t^{1/\alpha}} (e^{cu} - 1 - cu) \, \mu(du) \right].$$

Taking $c = \pm t^{-1/\alpha}$ and using **H1**, we get

$$\mathsf{E} \exp \left(\pm t^{-1/\alpha} \xi_t \right) \leq \exp \left[C_1 t \int_{|u| \leq t^{1/\alpha}} (t^{-1/\alpha} u)^2 \mu(du) \right] \leq C_2,$$

which yields the required L_p-bounds. Applying the Cauchy inequality, we get finally that the family

$$\left\{ \frac{|\xi_t|}{\sqrt{\kappa_t}} \right\}_{t \in (0,1]}$$

has bounded L_p-norms.

For the second summand in the right hand side of (4.9), we write the Cauchy inequality:

$$\frac{|\zeta_t|}{\sqrt{\eta_t}} \leq \sqrt{N_t}, \quad N_t = \nu([0,t] \times \{|u| > t^{1/\alpha}\}).$$

Observe that N_t has a Poisson law with the intensity

$$t\mu(|u| > t^{1/\alpha}), \quad t \in (0,1],$$

which is bounded because of **H1.** Hence the family

$$\left\{ \frac{|\zeta_t|}{\sqrt{\eta_t}} \right\}_{t \in [0,1]}$$

also has bounded L_p-norms, which completes the proof of (4.8). □

Lemma 4.3. *Let δ be the same as in **H2**. Then*

$$\sup_{\theta \in \tilde{\Theta}, t \in (0,1]} \mathsf{E} \left(\frac{\delta_t(1)}{\sqrt{\mathrm{D}_t X_t^\theta}} \right)^{2+\delta} < \infty. \tag{4.10}$$

Proof: The proof is similar to the previous one, but some additional technicalities arise because now we have $\chi(u)$ instead of u under the integrals in the numerator. We have

$$\delta_t(1) = \int_0^t \int_{|u| \leq t^{1/\alpha}} \chi(u)\tilde{\nu}(ds, du) + \int_0^t \int_{|u| > t^{1/\alpha}} \chi(u)\nu(ds, du)$$
$$+ t^{1+2/\alpha}\left[m(t^{1/\alpha}) - m(-t^{1/\alpha}) \right]$$
$$=: \hat{\xi}_t + \hat{\zeta}_t + \varpi_t.$$

Denote $t_0 = (u_0)^\alpha$ with u_0 coming from **H2**, then the ratio

$$\upsilon(u) := \frac{\chi(u)}{u} = -\frac{um'(u)}{m(u)} - 2$$

is bounded on the set $\{|u| \leq t^{1/\alpha}\} \subset \{|u| \leq u_0\}$. Then the same argument as we have used before shows that the family

$$\left\{ \frac{|\hat{\xi}_t|}{\sqrt{\kappa_t}} \right\}_{t \in (0,t_0]}$$

has bounded L_p-norms for every $p \geq 1$.

Next, we have by **H1** that

$$\varpi_t \sim (C_+ - C_-)t^{1/\alpha}, \quad t \to 0+,$$

and thus

$$|\varpi_t| \leq Ct^{1/\alpha}, \quad t \in (0,1].$$

Then by Lemma 4.1 the family

$$\left\{ \frac{\varpi_t}{\sqrt{\kappa_t}} \right\}_{t \in (0,t_0]}$$

has bounded L_p-norms for every $p \geq 1$.

Finally, for $t \leq t_0$ by the Cauchy inequality we have

$$|\hat{\zeta}_t| \leq \sqrt{\eta_t}\sqrt{J_t},$$

where

$$J_t = \int_0^t \int_{|u|>t^{1/\alpha}} \upsilon^2(u)\nu(ds,du)$$

$$= \int_0^t \int_{t^{1/\alpha}<t\le u_0} \upsilon^2(u)\nu(ds,du) + \int_0^t \int_{|u|>u_0} \upsilon^2(u)\nu(ds,du) =: J_t^1 + J_t^2.$$

The function $\upsilon(u)$ is bounded on $\{|u| \le u_0\}$, hence

$$J_t^1 \le CN_t, \quad t \in (0,t_0],$$

where N_t is the same as in the proof of Lemma 4.2. Hence $J_t^1, t \le t_0$ have bounded L_p-norms for every $p \ge 1$. The random variable J_t^2 has a compound Poisson distribution with the intensity of the Poisson random variable equal $t\mu(|u| > u_0)$, and the law of a single jump equal to the image under υ of the measure μ conditioned to $\{|u| > u_0\}$. By condition **H2**, this law have a finite moment of the order $2 + \delta$, therefore the variables $J_t^2, t \le t_0$ have bounded $L_{2+\delta}$-norms. Summarizing all the above, we get

$$\sup_{\theta\in\tilde\Theta, t\in(0,t_0]} \mathsf{E}\left(\frac{\delta_t(1)}{\sqrt{D_t X_t^\theta}}\right)^{2+\delta} < \infty.$$

The same bound for $t \in [t_0, 1]$ can be proved in a similar and simpler way; in that case instead of taking the integrals with respect to $\{|u| \le t^{1/\alpha}\}, \{|u| > t^{1/\alpha}\}$, one should consider, both in the numerator and the denominator, the integrals with respect to $\{|u| \le u_0\}, \{|u| > u_0\}$. $\qquad\square$

Now we deduce (4.2) by simply applying the Hölder inequality to (4.5) with the estimates (4.6)–(4.8) and (4.10).

Remark 4.1. Now we can explain the main idea, which the choice of the intensity function $\rho(u) = u^2$ is based on. When ρ is compactly supported as was in Ivanenko and Kulik (2014), the "large jumps" are excluded from the formula for $D_{t,\rho}X_t^\theta$. On the other hand, "large jumps" are involved e.g. into $\tilde{Z}_t$, which will appear in the numerator in one term in (4.12). We have $\rho(u) = 0, |u| > u_*$ for some $u_* > 0$, and hence the integrals

$$\int_0^t \int_{|u|>u_*} u\nu(ds,du), \quad \int_0^t \int_{\mathbb{R}} \rho(u)\nu(ds,du)$$

are independent. In addition, we know that

$$\mathsf{E}\left(\int_0^t \int_{|u|>u_*} u\nu(ds,du)\right)^2 = t\int_{|u|>u_*} u^2\mu(du) + t^2\left(\int_{|u|>u_*} u^2\mu(du)\right)^2,$$

$$t^{-2/\alpha}\int_0^t \int_{\mathbb{R}} \rho(u)\nu(ds,du) \Rightarrow \zeta, \quad t \to 0,$$

where ζ is a positive $(\alpha/2)$-stable variable. Using this, it is easy to deduce a lower bound

$$\mathsf{E}\left(\frac{\int_0^t \int_{|u|>u_0} u\nu(ds,du)}{\sqrt{D_{t,\rho}X_t^\theta}}\right)^2 \ge Ct^{-2/\alpha+1},$$

which is unbounded for small t because $\alpha < 2$. This indicates that, in the present high-frequency sampling setting, one can hardly expect to get the *uniform* moment

bound of the type (4.2) for a Malliavin weight which corresponds to a compactly supported ρ. Nevertheless, in the modified construction we "extend the support" of ρ; this brings a "large jumps" part to the denominator, which provides a good balance to respective parts which appear in the numerator, and this is the reason why the modified weight satisfies the required uniform moment bounds.

Remark 4.2. Another natural possibility to design the Malliavin weight is to take into account the local scale for the process Z and to make the function ρ depend on t in the following way:

$$\rho(u) = \rho_t(u) = u^2 \varsigma(t^{-1/\alpha} u)$$

with $\varsigma \in C^1$ such that $\varsigma(u) = 1, |u| \leq 1$ and $\varsigma(u) = 0, |u| \geq 2$. Actually, the "scaled" choice of $\rho = \rho_t$ with the size of its support $\asymp t^{1/\alpha}$ is essentially the one used in the Malliavin calculus construction developed in Clément and Gloter (2015). It is easy to see that, under such a choice, an analogue of (4.10) would hold true; the reason is that now δ_t would contain only "small jumps part", which is well balanced with $\sqrt{DX_t^\theta}$ in completely the same way we have seen in the proof of Lemma 4.2. This would give the uniform moment bound for the *first* component of the respective Malliavin weight, and hence the "scaled" choice of $\rho = \rho_t$ is appropriate when the unknown parameter is involved into the drift term, only. However, the model with the parameter involved into the jump term this choice does not seem to be appropriate by the reason stated in Remark 4.1: the "large jump part" of $\tilde{Z}_t$ is not well balanced by the "small jump part" $\sqrt{\kappa_t}$ of $\sqrt{DX_t^\theta}$. In regard to this point, for the sake of reference let us discuss Clément and Gloter (2015) in a little bit more detail. Therein the authors proved the drift-parameter LAMN property when observing a sample $(X_{j/n})_{j=0}^n$ from the process X of the form

$$X_t = x_0 + \int_0^t b(X_s, \beta) ds + Z_t.$$

Their technical assumptions are: (i) the boundedness of the Lévy measure of Z, which is assumed to be locally α-stable in a neighborhood of the origin; (ii) the smoothness and boundedness of $(x, \beta) \mapsto b(x, \beta)$; and that (iii) the stable-like index $\alpha \in (1, 2)$, imposed just for ignoring the presence of the drift term in looking at the "scaled" increment $n^{-1/\alpha}(X_{j/n} - X_{(j-1)/n})$, which is to be close to $L_{j/n} - L_{(j-1)/n}$. In particular, the boundedness (ii) seems essential in their proof, as well as the fact that only the drift parameter is the subject to statistical estimation. We emphasize that because of absence of the jump-related parameter γ, the result of Clément and Gloter (2015) is not comparable neither with our current results nor even with those from the aforementioned paper Aït-Sahalia and Jacob (2007). We expect that our proof technique based on the modified Malliavin weight combined with the approximation of the intensity function ρ will be workable for the general case of state-dependent coefficients including the jump coefficient which contains a parameter γ; this is a subject of a further research.

4.3. *Proof of (4.3): Integral representation.* The proof consists of three steps outlined at the end of Section 4.1. The first step is based on a version of the Malliavin calculus on a space of trajectories of a Lévy process, which we outline below and which is essentially developed in Ivanenko and Kulik (2015). Note that the Malliavin calculus for Lévy noises is a classical and well developed tool, which dates back

to Bichteler et al. (1987) and Bismut (1983). However, we found it difficult to apply existing technique directly for our purposes: the reason is that unlike in the classical approach developed in Bichteler et al. (1987) and Bismut (1983), we are interested not in the distribution density $p_t(\theta; x, y)$ itself (which is typically treated by means of the inverse Fourier transform), but in the ratio $\partial_\theta p_t(\theta; x, y)/p_t(\theta; x, y)$. To get the L_p bounds for this ratio, and especially to make approximating procedures outlined at the end of Section 4.1, we need to have an integral representation for this ratio in a most simple possible form. For that purpose mainly, and also to make the exposition self-consistent, we introduce a specially designed simple version of the Malliavin calculus. This version of course is neither a substantial novelty nor is unique possible one; see e.g. the construction in Clément and Gloter (2015) aimed at similar purposes.

Let the Lévy measure μ of Z satisfy **H1**, **H2** and assume additionally that Z is "light tailed" in the sense that (4.4) holds true for some $\delta' > 0$. Fix some function $\rho \in C^2$ such that $\rho(u) = u^2$ in a neighborhood of the point $u = 0$. Consider a flow $Q_c, c \in \mathbb{R}$ of transformations of $\mathbb{R}$, which satisfies

$$\frac{d}{dc}Q_c(u) = \rho(Q_c(u)), \quad Q_0(u) = u,$$

and for a fixed $t > 0$ define respective family $\mathcal{Q}_c^t, c \in \mathbb{R}$ of transformations of the process $Z_s, s \geq 0$ by the following convention: the process $\mathcal{Q}_c^t Z$ has jumps at the same time instants with the initial process Z; if the process has a jump with the amplitude u then at the time moment s, respective jump of Z has the amplitude equal either $Q_c(u)$ or u if $s \leq t$ or $s > t$, respectively. It is proved in Ivanenko and Kulik (2015), Proposition 1, that under conditions **H1**, **H2** the law of $\mathcal{Q}_c^t Z$ in $\mathbb{D}(0, \infty)$ is absolutely continuous with respect to the law of Z. Hence every transformation $\mathcal{Q}_c^t$ can be naturally extended to a transformation of the space $L_0(\Omega, \sigma(Z), \mathsf{P})$ of the functionals of the process Z. Denote this transformation by the same symbol $\mathcal{Q}_c^t$, and call a random variable $\xi \in L_2(\Omega, \sigma(Z), \mathsf{P})$ stochastically differentiable if there exists the mean square limit

$$\hat{\mathrm{D}}\xi = \lim_{\varepsilon \to 0} \frac{\mathcal{Q}_\varepsilon^t \xi - \xi}{\varepsilon}.$$

The $L_2(\Omega, \sigma(Z), \mathsf{P})$-closure of the operator $\hat{\mathrm{D}}$ is called the *stochastic derivative* and is denoted by D. The adjoint operator $\delta = \mathrm{D}^*$ is called the *divergence operator* or the *extended stochastic integral*. The operators D, δ are well defined under conditions **H1**, **H2** for every $t > 0$ and ρ specified above; see Ivanenko and Kulik (2015), Remark 3.

Clearly, the above construction depends on the choice of t and ρ: to track this dependence we use the notation $\mathrm{D}_{t,\rho}, \delta_{t,\rho}$ instead of D, δ. In a slightly larger generality, literally the same construction can be made on the space $L_0(\Omega, \sigma(Z, U), \mathsf{P})$ of the functionals of the pair of processes Z and U, with the trajectories of U not being perturbed by $\mathcal{Q}_c^T$. Then, analogously to the calculations made in Ivanenko and Kulik (2015), Sections 3.1, 3.2, we have

$$\mathrm{D}_{t,\rho}Z_t = \int_0^t \int_{\mathbb{R}} \varrho(u)\nu(ds, du), \quad \mathrm{D}_{t,\rho}U_t = 0,$$

$$\delta_{t,\rho}(1) = \int_0^t \int_{\mathbb{R}} \chi_\rho(u)\tilde{\nu}(ds, du), \quad \chi_\rho(u) = -\varrho(u)\frac{m'(u)}{m(u)} - \varrho'(u);$$

recall that ν and $\tilde{\nu}$ are, respectively, the Poisson point measure and the compensated Poisson measure from the Lévy-Itô representation of Z. Respectively,

$$\mathrm{D}_{t,\rho}X_t^\theta = \gamma\mathrm{D}_{t,\rho}Z_t = \gamma\int_0^t\int_\mathbb{R}\varrho(u)\nu(ds,du).$$

Furthermore, the second order stochastic derivative of X_t^θ is well defined:

$$\mathrm{D}_{t,\rho}^2X_t^\theta = \gamma\int_0^t\int_\mathbb{R}\varrho(u)\varrho'(u)\,\nu(ds,du).$$

By Proposition 3.1 and formula (3.4), the variable X_t^θ has a distribution density $p_t(\theta;x)$ which is a C^2-function with respect to θ, x. On the other hand, the Malliavin calculus developed above allows one to derive an integral representation for the ratio

$$g_t(\theta;x,y) = \frac{\nabla_\theta p_t(\theta;x,y)}{p_t(\theta;x,y)}.$$

Namely, repeating literally the proof of the assertion III of Theorem 1 in Ivanenko and Kulik (2015), we obtain the following representation:

$$g_t(\theta;x) = \begin{cases} \mathsf{E}_{0,x}^{t,\theta}\Xi_{t,\rho}^\theta, & p_t(\theta;x) > 0, \\ 0, & \text{otherwise,} \end{cases} \tag{4.11}$$

where

$$\Xi_{t,\rho}^\theta := \delta_{t,\rho}\left(\frac{\nabla_\theta X_t^\theta}{\mathrm{D}_{t,\rho}X_t^\theta}\right) = \frac{(\delta_{t,\rho}(1))(\nabla_\theta X_t^\theta)}{\mathrm{D}_{t,\rho}X_t^\theta} + \frac{(\mathrm{D}_{t,\rho}^2X_t^\theta)(\nabla_\theta X_t^\theta)}{(\mathrm{D}_{t,\rho}X_t^\theta)^2} - \frac{\mathrm{D}_{t,\rho}(\nabla_\theta X_t^\theta)}{\mathrm{D}_{t,\rho}X_t^\theta}. \tag{4.12}$$

Note that, formally, we can not apply Theorem 1 of Ivanenko and Kulik (2015) directly, because now we have an additional process U which our target process X^θ depends on. Nevertheless, because U is not perturbed under the transformations $\mathcal{Q}_c^t$ which give the rise for the Malliavin calculus construction, it is easy to check that literally the same argument as the one used in the proof of Theorem 1 Ivanenko and Kulik (2015) can be applied in the current (slightly extended) setting.

Recall that we already know $p_t(\theta;x,y)$ exists and is smooth with respect to θ, x, y. We have

$$\int_\mathbb{R} f(y)\nabla_\theta p_t(\theta;x,y)\,dy = \nabla_\theta\mathsf{E}_x^\theta f(X_t^\theta) = \mathsf{E}_x^\theta f'(X_t^\theta)(\nabla_\theta X_t^\theta)$$

$$= \mathsf{E}_x^\theta\mathrm{D}_{t,\rho}f(X_t^\theta)\left(\frac{\partial_\theta X_t}{\mathrm{D}_{t,\rho}X_t}\right) = \mathsf{E}_x^\theta f(X_t^\theta)\Xi_{t,\rho}^1$$

$$= \mathsf{E}_x^\theta f(X_t^\theta)g_t^\theta(x,X_t^\theta) = \int_\mathbb{R} f(y)g_t(\theta;x,y)p_t(\theta;x,y)\,dy.$$

Hence the formula (4.11) is actually equivalent to the following: for every compactly supported $f \in C^1(\mathbb{R})$,

$$\nabla_\theta\mathsf{E}f(X_t^\theta) = \mathsf{E}f(X_t^\theta)\Xi_{t,\rho}^\theta, \tag{4.13}$$

and to prove (4.3) it is sufficient to prove (4.13) with the modified Malliavin weight Ξ_t^θ instead of $\Xi_{t,\rho}^\theta$. To do that, we exploit an approximation procedure, hence we rewrite (4.13) in an integral form, which is convenient for approximation purposes:

$$\mathsf{E}f\left(X_t^{\theta+v}\right) - \mathsf{E}f\left(X_t^\theta\right) = \int_0^1\mathsf{E}\left\{f\left(X_t^{\theta+sv}\right)\left(\Xi_{t,\rho}^{\theta+sv},v\right)\right\}ds. \tag{4.14}$$

Because
$$\partial_\beta X^\theta = t, \quad \partial_\gamma X^\theta = Z_t,$$

we have
$$D_{t,\rho}(\partial_\beta X^\theta) = 0, \quad D_{t,\rho}(\partial_\gamma X^\theta) = D_{t,\rho} Z_t = \frac{1}{\gamma} D_{t,\rho} X_t^\theta,$$

therefore
$$\Xi_{t,\rho}^\theta = \left(\frac{t\delta_{t,\rho}(1)}{D_{t,\rho} X_t^\theta} + \frac{t D_{t,\rho}^2 X_t^\theta}{(D_{t,\rho} X_t^\theta)^2}, \frac{Z_t \delta_{t,\rho}(1)}{D_{t,\rho} X_t^\theta} + \frac{Z_t D_{t,\rho}^2 X_t^\theta}{(D_{t,\rho} X_t^\theta)^2} - \frac{1}{\gamma} \right)^\top.$$

Now we proceed with the first approximation step as follows. Fix some $\rho_1 \in C^2(\mathbb{R}), \rho_1(u) \geq 0$ such that
$$\rho_1(u) = \begin{cases} u^2, & |u| \leq 1; \\ 0, & |u| \geq 2, \end{cases}$$

and define
$$\rho_N(u) = N^2 \rho_1(u/N).$$

Observe that, for t fixed and N large enough, $\rho_N(u) = u^2$ for $|u| \leq t^{1/\alpha}$, hence
$$D_{t,\rho_N} X_t^\theta \geq \gamma \kappa_t.$$

Next, there exists a constant C such that
$$\rho_N(u) \leq Cu^2, \quad |\rho'_N(u)| \leq C|u|,$$

and therefore
$$\left| \frac{\chi_{\rho_N}(u)}{u} \right| \leq C(\tau(u) + 1).$$

Then, repeating literally the calculations from Section 4.2, we can obtain a bound similar to (4.2) for t fixed, but a family of weights $\Xi_{t,\rho_N}^\theta, N \geq 1$ is considered instead:
$$\sup_{N \geq 1, \theta \in \tilde{\Theta}} \mathsf{E} \left| \Xi_{t,\rho_N}^\theta \right|^{2+\delta_1} < \infty, \quad \delta_1 < \delta \wedge \delta' \tag{4.15}$$

(here δ comes from H2, and δ' comes from (4.4)). Hence the family $\{\Xi_{t,\rho_N}^\theta, N \geq 1, \theta \in \tilde{\Theta}\}$ is uniformly integrable. It is straightforward to see that
$$\Xi_{t,\rho_N}^{\theta_N} \to \Xi_t^\theta, \quad N \to \infty$$

with probability 1 for any sequence $\theta_N \to \theta \in \Theta$. Combined with the above uniform integrability, this shows that
$$\Xi_{t,\rho_N}^\theta \to \Xi_t^\theta, \quad N \to \infty$$

in $L_1(\Omega, \mathsf{P})$ uniformly with respect to $\theta \in \tilde{\Theta}$. Hence we can pass to the limit in (4.14) as $N \to \infty$ and get the required identity (4.13) with the modified Malliavin weight Ξ_t^θ. This proves the representation (4.3) under the additional moment assumption (4.4).

The second approximation step is aimed to remove the assumption (4.4), and is similar to the above one. Consider a family of processes $Z^L, L \geq 1$ with Lévy measures
$$\mu_L(du) = m_L(u)du, \quad m_L(u) = m(u)e^{-u^2/L};$$

the Z^L-driven versions of the processes X_t^θ and Ξ_t^θ are also specified by the superscript L: $X^{\theta,L}, \Xi^{\theta,L}$. Because $|u|^{2+\delta} e^{-u^2/L} \leq C$, every μ_L satisfies (4.4). In

addition, it is an easy calculation to show that conditions **H1**, **H2** are satisfied for μ_L uniformly with respect to $L \geq 1$. Hence we have the following:

(a) for every $L \geq 1$, (4.14) holds true with X_t^θ and Ξ_t^θ replaced by $X_t^{\theta,L}$ and $\Xi_t^{\theta,L}$, respectively;

(b) for every $t \in (0,1]$, the family $\Xi_t^{\theta,L}, \theta \in \tilde\Theta, L \geq 1$ satisfies an analogue of (4.15) uniformly with respect to $L \geq 1$ (to prove this, one should repeat literally the calculations from Section 4.2).

Finally, it is straightforward to see that $(X_t^{\theta,L}, \Xi_t^{\theta,L})$ weakly converge to $(X_t^\theta, \Xi_t^\theta)$ as $L \to \infty$. Since the family $\{\Xi_t^{\theta,L}\}$ is uniformly integrable by the above property (b), we can pass to the limit in the relation (4.14) for $X_t^{\theta,L}, \Xi_t^{\theta,L}$, and get finally (4.14) for X_t^θ, Ξ_t^θ. This proves (4.3) and completes the proof of Theorem 4.1. $\quad\square$

Appendix A. Proof of Proposition 3.1

(1) Because U is negligible (see (2.8)), we can restrict our considerations to the variables

$$\zeta_{\alpha,t} = t^{-1/\alpha}(Z_t + c_t).$$

Their characteristic functions have the form $\mathsf{E}e^{i\lambda\zeta_{\alpha,t}} = e^{\psi_{\alpha,t}(\lambda)}$, where

$$\psi_{\alpha,t}(\lambda) = t \int_{\mathbb{R}} \left(e^{i\lambda t^{-1/\alpha}u} - 1 - i\lambda t^{-1/\alpha}u 1_{|u|\leq 1} \right) \mu(du) + i\lambda c_t t^{-1/\alpha}$$

$$= t \int_{\mathbb{R}} \left(e^{i\lambda t^{-1/\alpha}u} - 1 - i\lambda t^{-1/\alpha}u 1_{|u|\leq t^{1/\alpha}} \right) \mu(du);$$

in the last identity we have used the formula (2.7) for c_t. Changing the variable $v = ut^{-1/\alpha}$, we get

$$\psi_{\alpha,t}(\lambda) = \int_{\mathbb{R}} \left(e^{i\lambda v} - 1 - i\lambda v 1_{|v|\leq 1} \right) \mu_t(dv),$$

where $\mu_t(dv)$ has the density

$$m_t(v) = t^{1+1/\alpha} m(t^{1/\alpha}v).$$

By **H1**, for every $\varepsilon > 0$ there exists $u_\varepsilon > 0$ such that

$$(1-\varepsilon)m_{\alpha,C_\pm}(v) \leq m_t(v) \leq (1+\varepsilon)m_{\alpha,C_\pm}(v), \quad |v| \leq t^{-1/\alpha}u_\varepsilon.$$

On the other hand, the term $\left(e^{i\lambda v} - 1 - i\lambda v 1_{|v|\leq 1} \right)$ is bounded, and

$$\mu_t\left(\{v : |v| > t^{-1/\alpha}u_\varepsilon\} \right) = t\mu\left(\{u : |u| > u_\varepsilon\} \right) \to 0, \quad t \to 0.$$

Using that, one can easily derive

$$\psi_{\alpha,t}(\lambda) \to \psi_{\alpha,C_\pm}(\lambda) := \int_{\mathbb{R}} \left(e^{i\lambda v} - 1 - i\lambda v 1_{|v|\leq 1} \right) m_{\alpha,C_\pm}(v)\, dv, \quad t \to 0.$$

Because the characteristic function of $Z_1^{\alpha,C_\pm}$ equals $e^{\psi_{\alpha,C_\pm}(\lambda)}$, this completes the proof.

(2) Consider first the case $U \equiv 0$; now $\phi_{\alpha,t}$ does not depend on θ, and we omit θ in the notation. We would like to apply the inverse Fourier transform representation for $\phi_{\alpha,t}$ and its derivatives:

$$\phi_{\alpha,t}(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\lambda x + \psi_{\alpha,t}(\lambda)}\, d\lambda,$$

$$(\partial_x)^k \phi_{\alpha,t}(x) = \frac{1}{2\pi} \int_{\mathbb{R}} (-i\lambda)^k e^{-i\lambda x + \psi_{\alpha,t}(\lambda)} \, d\lambda,$$

To do that, we have to verify that the functions under the integrals are absolutely integrable. We have

$$|e^{-i\lambda x + \psi_{\alpha,t}(\lambda)}| \le e^{\operatorname{Re}\psi_{\alpha,t}(\lambda)},$$

$$\operatorname{Re}\psi_{\alpha,t}(\lambda) = t \int_{\mathbb{R}} (\cos(t^{-1/\alpha}\lambda u) - 1)\mu(du) \le t \int_{t^{-1/\alpha}|\lambda u| < 1} (\cos(t^{-1/\alpha}\lambda u) - 1)\mu(du).$$

Then by **H1** there exist $c_1, c_2 > 0$ such that

$$|e^{-ix\lambda + \psi_{\alpha,t}(\lambda)}| \le c_1 e^{-c_2|\lambda|^\alpha}, \quad x, \lambda \in \mathbb{R}, \quad t \in (0, 1].$$

This proves existence of $\phi_{\alpha,t}(x)$ and all its derivatives. Moreover, we have

$$\sup_{x\in\mathbb{R}, t\in(0,1]} |\phi'_{\alpha,t}(x)| < \infty, \qquad \sup_{x\in\mathbb{R}, t\in(0,1]} |\phi''_{\alpha,t}(x)| < \infty.$$

Now we come back to the case of non-zero U. Because the law of $\zeta^\theta_{\alpha,t}$ is a convolution of the laws of $\zeta_{\alpha,t}$ and $\gamma^{-1}U_t$, the above bound can be extended:

$$\sup_{\theta\in\tilde{\Theta}} \sup_{x\in\mathbb{R}, t\in(0,1]} |\phi'_{\alpha,t}(\theta;x)| < \infty, \qquad \sup_{\theta\in\tilde{\Theta}} \sup_{x\in\mathbb{R}, t\in(0,1]} |\phi''_{\alpha,t}(\theta;x)| < \infty. \qquad (A.1)$$

In addition, $\zeta^\theta_{\alpha,t} \Rightarrow Z_1^{\alpha,C_\pm}$ uniformly in $\theta \in \tilde{\Theta}, U \in \mathfrak{U}$: this follows from the statement (1) and the fact that $\gamma^{-1}t^{-1/\alpha}U_t$ is uniformly negligible. This convergence and the first (resp. second) bound in (A.1) provide the first (resp. second) convergence in (3.1). $\qquad\square$

References

Y. Aït-Sahalia and J. Jacob. Volatility estimators for discretely sampled Lévy processes. *Ann. Statist.* **35** (1), 355–392 (2007). MR2332279.

Y. Aït-Sahalia and J. Jacod. Fisher's information for discretely sampled Lévy processes. *Econometrica* **76** (4), 727–761 (2008). MR2433480.

I.V. Basawa and D.J. Scott. *Asymptotic optimal inference for nonergodic models*, volume 17 of *Lecture Notes in Statistics*. Springer-Verlag, New York-Berlin (1983). ISBN 0-387-90810-2. MR688650.

K. Bichteler, J.-B. Gravereaux and J. Jacod. *Malliavin calculus for processes with jumps*, volume 2 of *Stochastics Monographs*. Gordon and Breach Science Publishers, New York (1987). ISBN 2-88124-185-9. MR1008471.

J.-M. Bismut. Calcul des variations stochastique et processus de sauts. *Z. Wahrsch. Verw. Gebiete* **63** (2), 147–235 (1983). MR701527.

L. Chaumont and G. Uribe Bravo. Markovian bridges: weak continuity and pathwise constructions. *Ann. Probab.* **39** (2), 609–647 (2011). MR2789508.

E. Clément and A. Gloter. Local asymptotic mixed normality property for discretely observed stochastic differential equations driven by stable Lévy processes. *Stochastic Process. Appl.* **125** (6), 2316–2352 (2015). MR3322866.

J.M. Corcuera and A. Kohatsu-Higa. Statistical inference and Malliavin calculus. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progr. Probab.*, pages 59–82. Birkhäuser/Springer Basel AG, Basel (2011). MR2857018.

G. Dohnal. On estimating the diffusion coefficient. *J. Appl. Probab.* **24** (1), 105–114 (1987). MR876173.

L. Fahrmeir. A note on asymptotic testing theory for nonhomogeneous observations. *Stochastic Process. Appl.* **28** (2), 267–273 (1988). MR952833.

V. Genon-Catalot and J. Jacod. On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Ann. Inst. H. Poincaré Probab. Statist.* **29** (1), 119–151 (1993). MR1204521.

V. Genon-Catalot and J. Jacod. Estimation of the diffusion coefficient for diffusion processes: random sampling. *Scand. J. Statist.* **21** (3), 193–221 (1994). MR1292636.

E. Gobet. Local asymptotic mixed normality property for elliptic diffusion: a Malliavin calculus approach. *Bernoulli* **7** (6), 899–912 (2001). MR1873834.

E. Gobet. LAN property for ergodic diffusions with discrete observations. *Ann. Inst. H. Poincaré Probab. Statist.* **38** (5), 711–737 (2002). MR1931584.

I.A. Ibragimov and R.Z. Has′minskiĭ. *Statistical estimation*, volume 16 of *Applications of Mathematics*. Springer-Verlag, New York-Berlin (1981). ISBN 0-387-90523-5. Asymptotic theory, Translated from the Russian by Samuel Kotz. MR620321.

D. Ivanenko and A. Kulik. LAN property for discretely observed solutions to Lévy driven SDE's. *Mod. Stoch. Theory Appl.* **1** (1), 33–47 (2014). MR3314792.

D.O. Ivanenko and A.M. Kulik. Malliavin calculus approach to statistical inference for Lévy driven SDE's. *Methodol. Comput. Appl. Probab.* **17** (1), 107–123 (2015). MR3306674.

P. Jeganathan. On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal. *Sankhyā Ser. A* **44** (2), 173–212 (1982). MR688800.

M. Kessler. Estimation of an ergodic diffusion from discrete observations. *Scand. J. Statist.* **24** (2), 211–229 (1997). MR1455868.

L. Le Cam. Locally asymptotically normal families of distributions. Certain approximations to families of distributions and their use in the theory of estimation and testing hypotheses. *Univ. california Publ. Statist.* **3**, 37–98 (1960). MR0126903.

L. Le Cam and G.L. Yang. *Asymptotics in statistics*. Springer Series in Statistics. Springer-Verlag, New York, second edition (2000). ISBN 0-387-95036-2. Some basic concepts. MR1784901.

H. Mai. Efficient maximum likelihood estimation for Lévy-driven Ornstein-Uhlenbeck processes. *Bernoulli* **20** (2), 919–957 (2014). MR3178522.

H. Masuda. Joint estimation of discretely observed stable Lévy processes with symmetric Lévy density. *J. Japan Statist. Soc.* **39** (1), 49–75 (2009). MR2571802.

H. Masuda. Parametric estimation of Lévy processes. In *Lévy matters. IV*, volume 2128 of *Lecture Notes in Math.*, pages 179–286. Springer, Cham (2015). MR3364258.

K.-i. Sato. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge (1999). ISBN 0-521-55302-4. Translated from the 1990 Japanese original,

Revised by the author. MR1739520.

T.J. Sweeting. Asymptotic ancillarity and conditional inference for stochastic processes. *Ann. Statist.* **20** (1), 580–589 (1992). MR1150364.

V.M. Zolotarev. *One-dimensional stable distributions*, volume 65 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI (1986). ISBN 0-8218-4519-5. Translated from the Russian by H. H. McFaden, Translation edited by Ben Silver. MR854867.