UNIFORM LAN PROPERTY OF LOCALLY STABLE LÉVY PROCESS OBSERVED AT HIGH FREQUENCY

D. O. IVANENKO, A. M. KULIK, AND H. MASUDA

ABSTRACT. Suppose we have a high-frequency sample from the Lévy process of the form $X^\theta_t = \beta t + \gamma Z_t + U_t$, where $Z$ is a locally $\alpha$-stable symmetric Lévy process and $U$ is a nuisance Lévy process. We will derive the LAN about the explicit parameter $\theta = (\beta, \gamma)$ under very mild conditions, generalizing the LAN result of [1]. Due to the special nature of the locally $\alpha$-stable character, the asymptotic Fisher information matrix takes a clean-cut form.

1. Introduction

Ever since Le Cam’s pioneering work [14], 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 just a few. One can consult [15] 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 [7] and [8] 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 [10] and Genon-Catalot and Jacod [5]. 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 [17] 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 $\alpha$-stable, the explicit LAN result has been proved in [1] and [16]. See also [2] for the precise asymptotic behavior of the Fisher-information matrix for the same model setting as in [1].

We will consider the Lévy process $X^\theta$ described by $X^\theta_t = \beta t + \gamma Z_t + U_t$, where $Z$ is a locally $\alpha$-stable symmetric Lévy process and $U$ is a Lévy process which is independent of $Z$ and less active than $Z$, and regarded as a nuisance process. The objective of this paper is to derive the LAN about the explicit parameter $\theta = (\beta, \gamma)$ under very mild conditions, when $X^\theta$ 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 [1]. 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 w.r.t. to the nuisance process $U$, and allows us to formulate the LAN property uniformly w.r.t. 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 [11], [12] for the model where $X^\theta$ 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 [3], 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 [3] is discussed in details in Section 4 below.

It is natural to ask for extending the 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 Jegathan [13], 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 [4], [6] and [7] 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 Clément-Gloter [3] is 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 bounds are proved in Section 4 by means of specially designed version of the Malliavin calculus.

2. Main results

Let $X^\theta$ be a Lévy process of the form

$$X^\theta_t = \beta t + \gamma Z_t + U_t, \quad t \geq 0.$$  

Here $Z, U$ are independent Lévy processes defined on a probability space $(\Omega, \mathcal{F}, P)$, and $\theta = (\beta, \gamma)^T \in \mathbb{R}^2$ is an unknown parameter subject to a statistical estimation. We are focused on the following setting:

- the process $X^\theta$ is discretely observed, i.e. the $n$-th sample contains its values at the first $n$ points $\{t_{k,n} = kh_n, k = 1, \ldots, n\}$ 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 $X^\theta$ have high frequency.

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

$$\left\{X^\theta_{t_{k,n}}, k = 1, \ldots, n\right\}$$
in \((\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))\), and write
\[
(2.2) \quad \mathcal{E}_n = \left\{ \mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), (\mathbb{P}_n^\theta, \theta \in \Theta) \right\}
\]
for a statistical model based on this sample.

Under the conditions on the process \(Z\) we specify below, the law \(\mathbb{P}_n^\theta\) is absolutely continuous w.r.t. Lebesgue measure i.e. the model \(\mathcal{E}_n\) possesses the likelihood function
\[
L_n(\theta; x_1, \ldots, x_n) = \frac{\mathbb{P}_n^\theta(dx_1 \ldots dx_n)}{dx_1 \ldots dx_n}
\]
\((L_n\) is called likelihood function). Denote
\[
Z_n(\theta_0, \theta; x_1, \ldots, x_n) = \frac{L_n(\theta; x_1, \ldots, x_n)}{L_n(\theta_0; x_1, \ldots, x_n)}
\]
the likelihood ratio of \(\mathbb{P}_n^\theta\) with respect to \(\mathbb{P}_n^{\theta_0}\) with the convention (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 \(\{\tau(n) = r(n, \theta_0), \; n \geq 1\}\) and the covariance matrix \(\Sigma(\theta_0)\), if for every \(v\) the likelihood ratio
\[
Z_n(\theta_0, \theta_0 + r(n)v) = Z_n(\theta_0, \theta_0 + r(n)v); X_{11, n}^{\theta_0}, \ldots, X_{1n, n}^{\theta_0}
\]
possesses representation
\[
(2.3) \quad Z_n(\theta_0, \theta_0 + r(n)v) = \exp \left\{ v^\top \Delta_n(\theta_0) - \frac{1}{2} v^\top \Sigma(\theta_0) v + \Psi_n(v, \theta_0) \right\}
\]
with
\[
(2.4) \quad \Delta_n(\theta_0) \Rightarrow \mathcal{N}(0, \Sigma(\theta_0)), \; n \to \infty
\]
and
\[
(2.5) \quad \Psi_n(v, \theta_0) \overset{P}{\to} 0, \; n \to \infty.
\]

We will establish the LAN property under the following principal assumptions on the processes \(Z, U\) involved into the model. We assume \(Z\) to be a symmetric Lévy process with the Lévy measure \(\mu\), which satisfies the following:

**H1.** \(\mu(du) = m(u)du\), and for some \(C > 0, \alpha \in (0, 2)\)
\[
m(u) \sim C|u|^{-\alpha - 1}, \; u \to 0;
\]

**H2.** \(m \in C^1(\mathbb{R} \setminus \{0\})\), and there exists \(u_0 > 0\) 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 \(\delta > 0\).

Clearly, a symmetric \(\alpha\)-stable process satisfies **H1, H2**; in this case \(m(u) = C|u|^{-\alpha - 1}\), and \(\tau(u)\) is just a constant. 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
\[
m(u) = f(u)|u|^{-\alpha - 1},
\]
where \(f(u) \to C, |u| \to 0\).

**Example 2.1.** (Tempered \(\alpha\)-stable process). For either \(f(u) = Ce^{-\sqrt{1+u^2}}\), \(f(u) = Ce^{-u^2}\), or \(f(u) = Ce^{-|u|}\), conditions **H1, H2** hold true, although \(\tau(u)\) fails to be bounded.
Example 2.2. (Smoothly damped $\alpha$-stable process). Let $m(u) = f(u)|u|^{-\alpha-1}$, where $f$ smoothly vanishes outside some interval $[-u_1,u_1]$. If the ratio $f'/f$ is bounded at the vicinity of $\pm u_1$ by a polynomial of $|u\mp u_1|$, then conditions $H1$, $H2$ hold true, although $\tau(u)$ fails to be bounded. One particular example of such $f$ is

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

The process $Z$ is “locally stable” in the sense that

$$t^{-1/\alpha}Z_t \Rightarrow Z_1^{(\alpha)}, \quad t \to 0;$$

here and below $Z^{(\alpha)}$ denotes a symmetric $\alpha$-stable process with the Lévy measure $C|u|^{-\alpha-1} \, du$. Denote $\phi_\alpha$ the distribution density for $Z_1^{(\alpha)}$.

Now we are able to formulate our main result.

Theorem 2.1. Let $Z$ be a symmetric Lévy process which satisfies $H1$, $H2$, and $U$ be a Lévy process such that

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

in probability. Recall that we assume $h_n \to 0$; in the case $\alpha \in (1,2)$, assume additionally

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

Then the LAN property holds true at every point $\theta_0 \in \Theta$ with

$$r(n) = \left( \begin{array}{cc} n^{-1/2}h_n^{1/\alpha-1} & 0 \\ 0 & n^{-1/2} \end{array} \right), \quad \Sigma(\theta_0) = \left( \begin{array}{cc} \Sigma_{11}(\theta_0) & 0 \\ 0 & \Sigma_{22}(\theta_0) \end{array} \right),$$

where

$$\Sigma_{11}(\theta_0) = \gamma_0^{-2} \int_{\mathbb{R}} \left( \frac{\phi_\alpha'(x)}{\phi_\alpha(x)} \right)^2 \phi_\alpha(x) \, dx, \quad \Sigma_{22}(\theta_0) = \gamma_0^{-2} \int_{\mathbb{R}} \left( 1 + \frac{x\phi_\alpha'(x)}{\phi_\alpha(x)} \right)^2 \phi_\alpha(x) \, dx.$$

The process $U$ has a natural interpretation as a “nuisance noise”, and then (2.6) is a natural requirement for a “nuisance” part of the noise to be negligible w.r.t. to the “principal” part. It is explained in [1] that for purposes of semiparametric statistical estimation, i.e. inference of $\theta$ when the law of the nuisance noise is not specified, it would be particularly useful to extend the above setup and to consider, instead of just one process (2.1), classes of processes where $U$ may vary in some “nuisance class” $\mathfrak{U}$. 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 (2.6) holds true uniformly over $U \in \mathfrak{U}$. If in addition $Z$ and $h_n$ satisfy conditions of Theorem 2.1 then for every $U \in \mathfrak{U}$ the likelihood ratio for the discretely observed process (2.1) admits a representation (2.3) with $r(n), \Sigma(\theta_0)$ specified in (2.7), and relations (2.4), (2.5) hold true uniformly over $U \in \mathfrak{U}$.

Remark 2.1. We are assuming that the activity index $\alpha$ is known. This seems disappointing, however, as was clarified in [16] and [2], if we attempt to joint maximum-likelihood estimation of $\alpha$ and the scale parameter $\gamma$ we 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.

3. Proofs

In this section we prove Theorem 2.1 and outline the proof of Theorem 2.2. The proofs are based on an $L_p$-bound for the derivative of the log-likelihood (Lemma 3.1), which is proved separately in Section 4.3 below.
3.1. **Proof of Theorem 2.1** Denote by $p_t(\theta; x, y)$ the transition probability density for $X^\theta$, considered as a Markov process; in what follows we will prove that this density exists. Denote also
\[ g_t(\theta; x, y) = \frac{\nabla_\theta p_t(\theta; x, y)}{p_t(\theta; x, y)} = \nabla_\theta \log p_t(\theta; x, y), \quad q_t(\theta; x, y) = \frac{\nabla_\theta p_t(\theta; x, y)}{2\sqrt{p_t(\theta; x, y)}} = \nabla_\theta \sqrt{p_t(\theta; x, y)}, \]
assuming the derivatives to exist for $P_t(x, \cdot)$-a.a. $y$ for every fixed $x, t$. Since $X^\theta$ has independent increments, we can write
\[ p_t(\theta; x, y) = p_t(\theta; y - x), \quad g_t(\theta; x, y) = g_t(\theta; y - x), \quad q_t(\theta; x, y) = q_t(\theta; y - x). \]

Hence the statistical model described above reduces to the one with independent observations. The LAN property for triangular arrays of independent observations is well studied, e.g. [9], 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** The function
\[ \Theta \ni \theta \to \sqrt{p_t(\theta; \cdot)} \in L_2(\mathbb{R}) \]
is continuously differentiable; that is, statistical experiment $\mathcal{E}_1$ is regular.

**A2**
\[ \lim_{n \to \infty} \mathbb{E} \left\| \sum_{k=1}^{n} \left( r(n) g_{h_n} \left( \theta_0; X^\theta_{kh_n} - X^\theta_{(k-1)h_n} \right) \right) \right\| = 0. \]

**A3** For some $\varepsilon > 0$,
\[ \lim_{n \to \infty} \sup_{\|v\| < N} n \int_{\mathbb{R}} |r(n)(g_{h_n}(\theta_0; y) + r(n)v; y) - q_{h_n}(\theta_0; y)|^2 \, dy = 0. \]

In what follows, we denote by $\tilde{\Theta}$ a subset of $\Theta$ such that
\[ (3.1) \quad \inf \{ \gamma : (\beta, \gamma) \in \tilde{\Theta} \} > 0. \]
Denote also $\tilde{r}(n) = n^{1/2}r(n)$. In Section 4 below we prove the following moment bound, which is a crucial ingredient in the proofs of A1–A4.

**Lemma 3.1.** Under conditions of Theorem 2.1, for every $\tilde{\Theta}$ and every $\varepsilon \in (0, \delta)$ (where $\delta$ comes from H2),
\[ \sup_{n \geq 1, \theta \in \tilde{\Theta}} \mathbb{E} \left| \tilde{r}(n) g_{h_n} \left( \theta; X^\theta_{h_n} \right) \right|^{2+\varepsilon} < \infty. \]

Because
\[ n \int_{\mathbb{R}} |r(n) g_{h_n} (\theta; y)|^{2+\varepsilon} p_{h_n} (\theta; y) \, dy = n^{-\varepsilon/2} \mathbb{E} \left| \tilde{r}(n) g_{h_n} (\theta; X^\theta_{h_n}) \right|^{2+\varepsilon}, \]
assertion A3 follows from Lemma 3.1 immediately.

Next, write
\[ \sum_{k=1}^{n} \left( r(n) g_{h_n} \left( \theta; X^\theta_{kh_n} - X^\theta_{(k-1)h_n} \right) \right) \right\|^2 = \frac{1}{n} \sum_{k=1}^{n} \left( \tilde{r}(n) g_{h_n} \left( \theta; X^\theta_{kh_n} - X^\theta_{(k-1)h_n} \right) \right) \right\|^2. \]
Consider the random variables
\[ \zeta_{\alpha, t} = t^{-1/\alpha}(Z_t + \gamma^{-1}U_t), \]
and denote by $\phi_{\alpha, t}(\theta; x)$ their distribution densities, which exist by Lemma A.1. By H1 and (2.6),
\[ (3.2) \quad \zeta_{\alpha, t} \Rightarrow Z_1^{(\alpha)}. \]
This convergence combined with the bound (3.1) yields that for every $N > 0$
\begin{equation}
\sup_{|x| \leq N} |\phi_{\alpha,t}(\theta; x) - \phi_{\alpha}(x)| \to 0, \quad \sup_{|x| \leq N} |\phi'_{\alpha,t}(\theta; x) - \phi'_{\alpha}(x)| \to 0.
\end{equation}

We have
\begin{equation}
p_t(\theta; x) = \gamma^{-1} t^{-1/\alpha} \phi_{\alpha,t} \left( \theta; \gamma^{-1} t^{-1/\alpha} (x - \beta t) \right),
\end{equation}
hence the random vectors
\[
\Gamma^\theta_{k,n} = \bar{r}(n) g_{h_n} \left( \theta; X^\theta_{kh_n} - X^\theta_{(k-1)h_n} \right), \quad k = 1, \ldots, n
\]
can be written as
\[
\Gamma^\theta_{k,n} = \gamma^{-1} G_{\alpha,h_n}(\theta; \xi^\theta_{k,n}),
\]
where
\[
\xi^\theta_{k,n} = \gamma^{-1} h_n^{-1/\alpha} (X^\theta_{kh_n} - X^\theta_{(k-1)h_n} - \beta h_n), \quad k = 1, \ldots, n
\]
are independent random variables identically distributed with $\xi^\theta_{\alpha,h_n}$, and the vector-valued function $G_{\alpha,t}$ have the components
\[
G^1_{\alpha,t}(\theta; x) = - \frac{\phi'_{\alpha,t}(\theta; x)}{\phi_{\alpha,t}(\theta; x)}, \quad G^2_{\alpha,t}(\theta; x) = -1 - \frac{x \phi'_{\alpha,t}(\theta; x)}{\phi_{\alpha,t}(\theta; x)}.
\]

Because $\phi_{\alpha}$ is positive, it follows from (3.3) that $G_{\alpha,t}(\theta; \cdot) \to G_{\alpha,t} \to 0^+$ uniformly on compacts, where the vector-valued function $G_{\alpha}$ have the components
\[
G^1_{\alpha}(x) = - \frac{\phi'_{\alpha}(x)}{\phi_{\alpha}(x)}, \quad G^2_{\alpha}(x) = -1 - \frac{x \phi'_{\alpha}(x)}{\phi_{\alpha}(x)}.
\]

Combined with (3.2), this yields that the common law of $\Gamma^\theta_{k,n}, k = 1, \ldots, n$ weakly converge as $n \to \infty$ to the law of $\Gamma^\theta = \gamma^{-1} G_{\alpha}(Z_1^{(\alpha)})$. Combined with the uniform integrability property of the family $\{\Gamma^\theta_{k,n}\}$, which holds true because of Lemma 3.1 this leads to A2; observe that it is an easy calculation to show that the covariation matrix for $\Gamma^\theta$ equals $\Sigma(\theta)$.

To prove A4, we change the variable $y = \beta h_n + \gamma h_n^{1/\alpha} x$, and write for any $\theta_1, \theta_2 \in \Theta$
\[
n \int \mathbb{R} |r(n) (q_{h_n}(\theta_1; y) - q_{h_n}(\theta_2; y))|^2 \, dy = \frac{1}{4} \int \mathbb{R} |\gamma_1^{-1} G_{\alpha,h_n}(\theta_1; x) \sqrt{\phi_{\alpha,h_n}(\theta_1; x)} - \gamma_2^{-1} G_{\alpha,h_n}(\theta_2; x) \sqrt{\phi_{\alpha,h_n}(\theta_2; x)}|^2 \, dx.
\]
The argument we have used in the proof of A2 can be easily extended to show that, for any sequence $\theta_n \to \theta$,
\[
G_{\alpha,h_n}(\theta_n; \cdot) \to G_{\alpha}, \quad \phi_{\alpha,h_n}(\theta_n; \cdot) \to \phi_{\alpha}
\]
uniformly on compacts. Then, by the Cauchy inequality $(a + b)^2 \leq 2a^2 + 2b^2$, for every $N > 0$
\[
\limsup_{n \to \infty} n \int \mathbb{R} |r(n) (q_{h_n}(\theta_n; y) - q_{h_n}(\theta; y))|^2 \, dy \leq \limsup_{n \to \infty} \left( \frac{1}{2\gamma_2} \int_{|x| > N} \left( G_{\alpha,h_n}(\theta_n; x) \right)^2 \phi_{\alpha,h_n}(\theta_n; x) \, dx \right) + \limsup_{n \to \infty} \left( \frac{1}{2\gamma_2} \int_{|x| > N} \left( G_{\alpha,h_n}(\theta; x) \right)^2 \phi_{\alpha,h_n}(\theta; x) \, dx \right).
\]
We have
\[
\frac{1}{2\gamma^2} \int_{|x| > N} \left( G_{a,h_n}(\theta; x) \right)^2 \phi_{a,h_n}(\theta; x) \, dx = E(\Gamma_{1,n}^\theta)^2 1_{|\xi_{1,n}^\theta| > N} \leq \left( E(\Gamma_{1,n}^\theta)^2 + \varepsilon \right)^{2/(2+\varepsilon)} \left( P(|\xi_{1,n}^\theta| > N) \right)^{\varepsilon/(2+\varepsilon)}.
\]

Recall that \( \xi_{1,n} \) weakly converges to \( Z_1^{(\alpha)} \), hence by Lemma 3.1 we have
\[
\limsup_{n \to \infty} n \int_{\mathbb{R}} |r(n)(q_{h_n}(\theta_n; y) - q_{h_n}(\theta; y))|^2 \, dy \leq C \left( P(|Z_1^{(\alpha)}| > N) \right)^{\varepsilon/(2+\varepsilon)}
\]
with some constant \( C \). Taking \( N \to \infty \), we get finally that, for every sequence \( \theta_n \to \theta \),
\[
n \int_{\mathbb{R}} |r(n)(q_{h_n}(\theta_n; y) - q_{h_n}(\theta; y))|^2 \, dy \to 0, \quad n \to \infty.
\]
This proves A4.

Literally the same argument shows that, for every fixed \( t > 0 \), the mapping
\[
\Theta \ni \theta \mapsto q_t(\theta; \cdot) \in L_2(\mathbb{R})
\]
is continuous. Then it is easy to prove that for any \( \theta_1, \theta_2 \) such that the segment \( [\theta_1, \theta_2] \) is contained in \( \Theta \),
\[
(3.5) \quad \sqrt{p_t(\theta_1; \cdot)} - \sqrt{p_t(\theta_2; \cdot)} = \left( \int_0^1 q_t((1-s)\theta_1 + s\theta_2; \cdot) \, ds \right)^\top (\theta_1 - \theta_2),
\]
with the integral understood in the sense of convergence of the Riemann sums in \( L_2(\mathbb{R}) \). We omit the details, just notice that one possible way to justify the above equality is to approximate the function \( \sqrt{x} \) by a proper sequence of smooth functions; see e.g. the proof of Theorem 2 in [11]. It follows from (3.5) and continuity of \( q_t(\theta; \cdot) \) that the function \( \sqrt{p_t(\theta; \cdot)} \) is continuously differentiable in the \( L_2(\mathbb{R}) \) sense, with the derivative equal \( q_t(\theta; \cdot) \). This proves A1, and completes the proof of Theorem 2.1.

3.2. Outline of the proof of Theorem 2.2. To get the required LAN property uniformly in \( U \in \mathcal{U} \), it is enough to fix a sequence \( U^n \) of Lévy processes, such that
\[
h_n^{-1/\alpha} U^n \to 0
\]
in probability, and repeat the above argument for a modified statistical model, where the process \( X^\theta \) in the \( n \)-th sample is replaced by
\[
X^\theta_{h_n} = \beta t + \gamma Z_t + U^n_t.
\]
The moment bound in Lemma 3.1 is, to a very high extent, insensitive w.r.t. the process \( U \); in particular, we will show in Section 3.1 that
\[
(3.6) \quad \sup_{n \geq 1, \theta \in \Theta} \mathbb{E} \left[ \tilde{r}(n)g_{h_n}(\theta; X^\theta_{h_n}) \right]^{2+\varepsilon} < \infty.
\]
In addition, the random variables
\[
\zeta^\theta_{a,h_n} = h_n^{-1/\alpha}(Z_{h_n} + \gamma^{-1}U^n_{h_n})
\]
weakly converge to \( Z_1^{(\alpha)} \). Hence repeating, with obvious notational changes, the calculations from Section 3.1 we get properties A1 – A4 for the modified model, which proves the required LAN property.
4. Malliavin calculus-based integral representation for the derivative of the log-likelihood function and related and $L_p$-bounds

Our main aim in this section is to prove Lemma 3.1, 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. The choice of a particular design of such a calculus is a non-trivial problem, and is one of the important issues we discuss below. Our exposition in this section will consist of the following three principal parts. First, we outline the version the Malliavin calculus on a space of trajectories of a Lévy process developed in [11], which would lead to a representation for the derivative of the log-likelihood function as a conditional expectation of a certain Malliavin weight. The weight which appears in this representation does not possess a uniform $L_p$-bound, which motivates our second step, where we introduce a modified Malliavin weight, specially designed to have uniform $L_p$-bound. Finally, we prove that the derivative of the log-likelihood function still possesses the integral representation with this modified weight, which would complete then the proof of Lemma 3.1.

4.1. Preliminaries. In this section, we outline the version of the Malliavin calculus on a space of trajectories of a Lévy process $Z$ developed in [11]. Let the Lévy measure $\mu$ of $Z$ satisfy $H1, H2$ and assume additionally that there exists $\delta > 0$ such that

$$\int_{|u| \geq 1} |u|^{2+\delta} \mu(du) < \infty. \tag{4.1}$$

Fix some function $\rho \in C^2$ such that (i) $\rho$ vanishes outside $[-u_0, u_0]$ ($u_0$ comes from the condition $H2$); (ii) $\rho(u) = u^2$ in some neighbourhood 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 $Q^T_c, c \in \mathbb{R}$ of transformations of the process $Z_t, t \geq 0$ by the following convention: process $Q^T_c Z$ has the 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 $t$, respective jump of $Z$ has the amplitude equal either $Q_c(u)$ or $u$ if $t \leq T$ or $t > T$, respectively. It is proved in [11], Proposition 1, that under conditions $H1, H2$ the law of $Q^T_c Z$ in $D(0, \infty)$ is absolutely continuous w.r.t. the law of $Z$. Hence every transformation $Q^T_c$ can be naturally extended to a transformation of the space $L_0(\Omega, \sigma(Z), P)$ of the functionals of the process $Z$. Denote this transformation by the same symbol $Q^T_c$, and call a random variable $\xi \in L_2(\Omega, \sigma(Z), P)$ stochastically differentiable if there exists the mean square limit

$$\hat{D}_t \xi = \lim_{\varepsilon \to 0} \frac{Q^T_{t\varepsilon} \xi - \xi}{\varepsilon}.$$

The $L_2(\Omega, \sigma(Z), P)$-closure of the operator $\hat{D}$ is called the stochastic derivative and is denoted by $D$. The adjoint operator $\delta = D^*$ is called the divergence operator or the extended stochastic integral. The operators $D, \delta$ are well defined under conditions $H1, H2$ for every $T > 0$ and $\rho$ specified above; see [11], Remark 3.

Clearly, the above construction depends on the choice of $T$ and $\rho$; any time we need to track this dependence we use the notation $D_{T, \rho}, \delta_{T, \rho}$ instead of $D, \delta$. In a slightly larger generality, literally the same construction can be made on the space $L_0(\Omega, \sigma(Z, U), P)$ of the functionals of the pair of processes $Z, U$, with the trajectories of $U$ not being perturbed by $Q^T_c$. Then, analogously to the calculations made in [11], Sections 3.1, 3.2, we have

$$D_{T, \rho} Z_t = \int_0^{t \wedge T} \int_{\mathbb{R}} g(u) \nu(ds, du), \quad D_{T, \rho} U_t = 0.$$
\[ \delta_{T,\rho}(1) = \int_0^T \int_\mathbb{R} \chi_\rho(u) \nu(ds, du), \quad \chi_\rho(u) = -\frac{g(u) m'(u)}{m(u)} - g'(u), \]

where \( \nu(ds, du) \) and \( \nu(ds, du) = \nu(ds, du) - ds\mu(du) \) are, respectively, the Poisson point measure and the compensated Poisson measure from the Lévy-Itô representation of \( Z \):

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

Respectively,

\[ D_{T,\rho} X^\theta_t = \gamma D_{T,\rho} Z_t = \gamma \int_0^{T\wedge T} \int_\mathbb{R} g(u) \nu(ds, du). \]

Furthermore, the second order stochastic derivative of \( X^\theta_t \) is well defined:

\[ D^2_{T,\rho} X^\theta_t = \gamma \int_0^{T\wedge T} \int_\mathbb{R} \frac{\partial g(u)}{\partial u}(u) \nu(ds, du). \]

By Lemma \[A.3\] and formula (3.4), the variable \( X^\theta_t \) has a distribution density \( p_t(\theta; x) \) which is a \( C^2 \)-function w.r.t. \( \theta, x \). On the other hand, the Malliavin calculus developed above allows one to derive an integral representation for the ratio

\[ g_t(\theta; x, y) = \frac{\nabla g p_t(\theta; x, y)}{p_t(\theta; x, y)}. \]

Take in the above construction \( T = t \), then the assertion III of Theorem 1 in [11] gives the following representation:

\[ g_t(\theta; x) = \begin{cases} E_{0,x}^{t,\theta} \Xi^{\theta}_{t,\rho}, & p_t(\theta; x) > 0, \\ 0, & \text{otherwise}, \end{cases} \]

where \( E_{0,x}^{t,\theta} \) denotes the expectation w.r.t. the law of the bridge of the process \( X^\theta_t \), conditioned to arrive to \( x \) at time \( t \), and

\[ \Xi^\theta_{t,\rho} := \delta_{t,\rho} \left( \frac{\nabla_\theta X^\theta_t}{D_{t,\rho} X^\theta_t} \right) = \frac{(\delta_{t,\rho}(1)) \nabla_\theta X^\theta_t}{D_{t,\rho} X^\theta_t} + \frac{(D^2_{t,\rho} X^\theta_t) \nabla_\theta X^\theta_t}{(D_{t,\rho} X^\theta_t)^2} - \frac{D_{t,\rho} \nabla_\theta X^\theta_t}{D_{t,\rho} X^\theta_t}. \]

Remark 4.1. Formally, we can not apply Theorem 1 [11] directly, because now we have an additional process \( U \) which our target process \( X^\theta \) depends on. Nevertheless, because \( U \) is not perturbed under the transformations \( Q^t \) 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 [11] can be applied in the current (slightly extended) setting.

In the sequel we call vector \( \Xi^\theta_{t,\rho} \) the Malliavin weight. We have

\[ \partial_\beta X^\theta_t = t, \quad \partial_\gamma X^\theta_t = Z_t, \]

and hence

\[ D_{t,\rho}(\partial_\beta X^\theta_t) = 0, \quad D_{t,\rho}(\partial_\gamma X^\theta_t) = D_{t,\rho} Z_t = \frac{1}{\gamma} D_{t,\rho} X^\theta_t. \]

Therefore the components \( \Xi^\beta_{t,\rho}, \Xi^\gamma_{t,\rho} \) of the Malliavin weight \( \Xi^\theta_{t,\rho} \) now have the form

\[ \Xi^\beta_{t,\rho} = \frac{t \delta_{t,\rho}(1)}{D_{t,\rho} X^\theta_t} + \frac{t D^2_{t,\rho} X^\theta_t}{(D_{t,\rho} X^\theta_t)^2}, \quad \Xi^\gamma_{t,\rho} = \frac{Z_t \delta_{t,\rho}(1)}{D_{t,\rho} X^\theta_t} + \frac{Z_t D^2_{t,\rho} X^\theta_t}{(D_{t,\rho} X^\theta_t)^2} - \frac{1}{\gamma}. \]

The guideline to the proof of Lemma 3.1 now can be explained as follows. The components of the vector

\[ \tilde{r}(n) g_{h_n} \left( \theta, X^\theta_{h_n} \right) \]
can be represented in the following way:
\[ h_n^{1/\alpha - 1} E[\Xi_{h_n,\rho}^\beta | X_{h_n}^\theta], \quad E[\Xi_{h_n,\rho}^\gamma | X_{h_n}^\theta]. \]

Hence, by the Jensen inequality, to prove Lemma 3.1 it would be enough to give uniform bounds of the form
\[ (4.4) \quad E\left( h_n^{1/\alpha - 1} \Xi_{h_n,\rho}^\beta \right)^{2+\varepsilon} \leq C, \quad E\left( \Xi_{h_n,\rho}^\gamma \right)^{2+\varepsilon} \leq C. \]

Now we can observe clearly the first difficulty which we encounter when we try to prove Lemma 3.1 from the results of [11], only, and would require an extension of the above construction. Such an representation (4.2) and the uniform bound (4.4) hold true. This, however, can not be deduced in a path-wise way, i.e. as sums over finite set of jumps. To define the same part of the integral we use a notation from Section 4.1 with the subscript \( \rho \).

Everywhere below we assume \( H_1, H_2 \) to hold true. 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
\[ E(t^{-2/\alpha} \kappa_t)^{-p} \leq C_p, \quad t \in (0, 1]. \]
Proof. For any \( \varepsilon < 1 \) we have
\[
P(\kappa_t < \varepsilon^2 t^{2/\alpha}) \leq P(\nu([0, t] \times \{|u| \in [\varepsilon t^{1/\alpha}, t^{1/\alpha}]\}) = 0) = \exp\{-t\mu(|u| \in [\varepsilon t^{1/\alpha}, t^{1/\alpha}]\})
\]
Condition \( H1 \) yields that, with some positive constant \( C \),
\[
t\mu(|u| \in [\varepsilon t^{1/\alpha}, t^{1/\alpha}]) \geq Ct\int_{\varepsilon t^{1/\alpha}}^{t^{1/\alpha}} |u|^{-\alpha-1} du = C(\varepsilon^{-\alpha} - 1), \quad t \in (0, 1].
\]
Hence for the family of random variables \( t^{-2/\alpha} \kappa_t, t \in (0, 1] \) we have a uniform bound
\[
P(t^{-2/\alpha} \kappa_t < \varepsilon^2) \leq e^{C(\varepsilon^{\alpha} - 1) + C}, \quad \varepsilon < 1, \quad t \leq 1,
\]
which proves the required statement.

Because
\[
\mathbf{D}X^\theta_t = \gamma \int_0^t \int_{\mathbb{R}} u^2 \nu(ds, du) \geq \gamma \kappa_t,
\]
this immediately gives the following: for every \( p \geq 1 \),
\[
(4.6) \quad \sup_{\theta \in \tilde{\Theta}, t \in (0, 1]} E \left( \frac{t^{1/\alpha}}{\sqrt{\mathbf{D}X^\theta_t}} \right)^p < \infty.
\]
Next, observe that both \( \mathbf{D}X^\theta_t \) and \( \mathbf{D}^2X^\theta_t \) are represented as sums over the set of jumps of the process \( Z \). Because
\[
\left( \sum_i a_i \right)^{3/2} \geq \sum_i a_i^{3/2}, \quad \{a_i\} \subset \mathbb{R},
\]
we have
\[
(4.7) \quad \left| \frac{\mathbf{D}^2X^\theta_t}{(\mathbf{D}X^\theta_t)^{3/2}} \right| \leq \frac{2}{\gamma}.
\]

**Lemma 4.2.** For \( \tilde{\Theta} \) same as above,
\[
(4.8) \quad \sup_{\theta \in \tilde{\Theta}, t \in (0, 1]} E \left( \frac{\delta(t)}{\sqrt{\mathbf{D}X^\theta_t}} \right)^{2+\delta} < \infty,
\]
where \( \delta \) comes from \( H2 \), and, for every \( p \geq 1 \),
\[
(4.9) \quad \sup_{\theta \in \tilde{\Theta}, t \in (0, 1]} E \left( \frac{Z_t}{\sqrt{\mathbf{D}X^\theta_t}} \right)^p < \infty.
\]

**Proof.** First, we consider (4.9). Taking into account that \( Z_t \) is symmetric, we decompose it in the following way:
\[
Z_t = \int_0^t \int_{|u| \leq t^{1/\alpha}} u\nu(ds, du) + \int_0^t \int_{|u| > t^{1/\alpha}} u\nu(ds, du) =: \xi_t + \zeta_t.
\]
Similarly, write
\[
\mathbf{D}X^\theta_t = \gamma \int_0^t \int_{\mathbb{R}} u^2 \nu(ds, du) = \gamma (\kappa_t + \eta_t), \quad \eta_t = \int_0^t \int_{|u| > t^{1/\alpha}} u^2 \nu(ds, du).
\]
Then
\[
(4.10) \quad \left| \frac{Z_t}{\sqrt{\mathbf{D}X^\theta_t}} \right| \leq \frac{1}{\gamma} \left( \frac{|\xi_t|}{\sqrt{\kappa_t + \eta_t}} + \frac{|\zeta_t|}{\sqrt{\kappa_t + \eta_t}} \right) \leq \frac{1}{\gamma} \left( \frac{|\xi_t|}{\sqrt{\kappa_t}} + \frac{|\zeta_t|}{\sqrt{\eta_t}} \right).
\]
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 
\[ \left\{ \frac{t^{-1/\alpha} \xi_t}{\sqrt{\kappa t}} \right\}_{t \in (0,1]} \]
also has bounded \( L_p \)-norms for any \( p \geq 1 \). To see this, observe that \( \xi_t \) is an integral of a deterministic function over a compensated Poisson point measure, and therefore its exponential moments can be expressed explicitly:
\[ \mathbb{E} \exp \left( c \xi_t \right) = \exp \left[ t \int_{|u| \leq t^{1/\alpha}} (e^{cu} - 1 - cu) \mu(du) \right]. \]
Taking \( c = \pm t^{-1/\alpha} \) and using H1, we get
\[ \mathbb{E} \exp \left( \pm t^{-1/\alpha} \xi_t \right) \leq \exp \left[ C_1 t \int_{|u| \leq t^{1/\alpha}} (t^{-1/\alpha} u)^2 \mu(du) \right] \leq C_2, \]
which yields the required \( L_p \)-bounds. Applying the Cauchy inequality, we get finally that the family 
\[ \left\{ \frac{|\xi_t|}{\sqrt{\kappa t}} \right\}_{t \in [0,1]} \]
also has bounded \( L_p \)-norms, which appears in the right hand side of (4.10), has bounded \( L_p \)-norms.

For the second summand in the right hand side of (4.10), we write the Cauchy inequality:
\[ |\zeta_t| \sqrt{\eta_t} \leq \sqrt{N_t}, \quad N_t = \nu([0, t] \times \{|u| > t^{1/\alpha}\}). \]
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 by 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 (4.9).

The proof of (4.8) is similar, with some additional technicalities which arise because now we have \( \chi(u) \) instead of \( u \) under the integral w.r.t. \( \tilde{\nu} \) in the numerator. Write
\[ \delta_t(1) = \int_0^t \int_{|u| \leq t^{1/\alpha}} \chi(u) \tilde{\nu}(ds, du) + \int_0^t \int_{|u| > t^{1/\alpha}} \chi(u) \nu(ds, du) =: \hat{\xi}_t + \hat{\zeta}_t. \]
Denote \( t_0 = (u_0)^\alpha \), where \( u_0 \) comes from H2, then the ratio
\[ \nu(u) = \frac{\chi(u)}{u} = \frac{\chi(u)}{m'(u)} - 2 \]
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 
\[ \left\{ \frac{|\hat{\zeta}_t|}{\sqrt{\kappa t}} \right\}_{t \in (0,t_0]} \]
has bounded \( L_p \)-norms for every \( p \geq 1 \).

Next, by the Cauchy inequality we have
\[ |\hat{\xi}_t| \leq \sqrt{\kappa t} \sqrt{J_t}, \]
where
\[ J_t = \int_0^t \int_{|u| > t^{1/\alpha}} v^2(u) \nu(ds, du) \]
\[ = \int_0^t \int_{t^{1/\alpha} < u \leq u_0} v^2(u) \nu(ds, du) + \int_0^t \int_{|u| > u_0} v^2(u) \nu(ds, du) = J_t^1 + J_t^2. \]
The function \( v(u) \) is bounded on \( \{|u| \leq u_0\} \), hence
\[ J_t^1 \leq C N_t, \quad t \in (0, t_0), \]
which has bounded \( L_p \)-norms for every \( p \geq 1 \). The process \( J_t^2 \) is a compound Poisson process with the total jump intensity \( \mu(|u| > u_0) \) and the law of a jump equal to the image under \( \nu \) of the measure \( \mu \) conditioned to \( \{|u| > u_0\} \). By condition \( H_2 \), this law have a finite moment of the order \( 2 + \delta \), therefore the values \( J_t^2, t \in (0, t_0) \) have bounded \( L_{2+\delta} \)-norms. This proves
\[ \sup_{\theta \in \tilde{\Theta}, t \in (0, t_0]} E \left( \frac{\delta_t(1)}{\sqrt{D_t X_t^p}} \right)^{2+\delta} < \infty. \]
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 w.r.t. \( \{|u| \leq t^{1/\alpha}\}, \{|u| > t^{1/\alpha}\} \), one should consider, both in the numerator and the denominator, the integrals w.r.t. \( \{|u| \leq u_0\}, \{|u| > u_0\} \).

Now, using the Hölder inequality and (4.6) – (4.9), we obtain finally the required moment bounds for the components of the modified Malliavin weight defined by explicit formula (4.5):
\[ \sup_{\theta \in \tilde{\Theta}, t \in (0, 1]} E \left( t^{1/\alpha - 1} \Xi_t^1 \right)^{2+\epsilon} < \infty, \quad \sup_{\theta \in \tilde{\Theta}, t \in [0, 1]} E \left( \Xi_t^1 \right)^{2+\epsilon} < \infty. \]

**Remark 4.2.** Now it is easy to explain the main idea of the above extension of the construction from [11]. If \( \rho \) is compactly supported, the “large jumps” are excluded from the formula for \( D_t \rho X_t^p \). On the other hand, “large jumps” are involved e.g. into \( Z_t \), which appears in the numerator in one term in (4.3). Because \( \rho(u) = 0 \) when \( |u| > u_0 \), the integrals
\[ \int_0^t \int_{|u| > u_0} u \tilde{\nu}(ds, du), \quad \int_0^t \int_{\mathbb{R}} \rho(u) \nu(ds, du) \]
are independent. In addition, we know that
\[ E \left( \int_0^t \int_{|u| > u_0} u \tilde{\nu}(ds, du) \right)^2 \geq t \int_{|u| > u_0} u^2 \mu(du), \]
\[ t^{-2/\alpha} \int_0^t \int_{\mathbb{R}} \rho(u) \nu(ds, du) \Rightarrow \zeta, \quad t \to 0, \]
where \( \zeta \) is a positive \((\alpha/2)\)-stable variable. Using this, it is easy to deduce a lower bound
\[ E \left( \int_0^t \int_{|u| > u_0} u \tilde{\nu}(ds, du) \right)^2 \geq C t^{-2/\alpha + 1}, \]
which is unbounded for small \( t \) because \( \alpha < 2 \). This indicates that the construction from [11] would hardly provide Malliavin weights which satisfy (4.4).

In the modified construction we “extend the support” of \( \rho \); 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 weights satisfy the required uniform moment bounds.
Remark 4.3. Another natural possibility to design the Malliavin weight would be to take into account the local scale for the process $Z$ and to make the function $\rho$ depend on $t$ in the following way:

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

with $\zeta \in C^1$ such that $\zeta(u) = 1, |u| \leq 1$ and $\zeta(u) = 0, |u| \geq 2$. 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 [3]. It is easy to see that, under such a choice, an analogue of (4.12) would hold true; the reason for that is that now $\delta_t$ would contain only “small jumps part”, which is well balanced with $\sqrt{D X^\theta_t}$ in completely the same way we have seen that in the proof of Lemma 4.2. This would give the first bound in (4.11). This means that, when the unknown parameter is involved into the drift term, only, the “scaled” choice of $\rho = \rho_t$ is another possibility to prove the required $L_p$-bounds. However, this setting does not seem to be appropriate to deal with the parameter involved into the jump term: we have seen in Remark 4.2 that the “large jump part” of $Z_t$ is not well balanced by the “small jump part” of $\sqrt{D X^\theta_t}$, hence neither (4.11) nor the second bound in (4.11) have no means to hold true.

4.3. Integral representation for $g_t(\theta, x)$ and the proof of Lemma 3.1. Our aim in this section is to prove that the modified Malliavin weight $\Xi^\theta_t$, defined by the formula (4.5), provides an integral representation for $g_t(\theta, x)$, similar to (4.2):

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

Because we already have proved the moment bound (4.11), this will complete the proof of Lemma 3.1.

Now it would be difficult to prove (4.12) using the arguments from Section 4.1 because we do not assume (4.1), the terms such as $D_t X^\theta_t$ may fail to be square integrable, and hence it is unclear how to apply the integration-by-parts arguments, usual for a Malliavin calculus framework.

Hence we use another way to prove (4.12), which is based on (4.2) and on a two-step approximation procedure. On the first step, we assume (4.1) and prove (4.12), approximating $\rho(u) = u^2$ by a sequence $\rho_N(u), N \geq 1$. On the second step, we remove the assumption (4.1), approximating $Z$ by a sequence $Z^L, L \geq 1$.

Since we already know that $p_t(\theta; x)$ exists and is smooth w.r.t. $\theta, x$, the required formula (4.12) now is equivalent to the following: for every compactly supported $f \in C^1(\mathbb{R})$,

$$\nabla_\theta \mathcal{E} f (X^\theta_t) = \mathcal{E} f (X^\theta_t \Xi^\theta_t);$$

see [11], proof of Theorem 1, for details. Write the above identity in an integral form, which would be convenient for approximation purposes:

$$\mathcal{E} f \left( X^\theta_t + x \right) - \mathcal{E} f \left( X^\theta_t \right) = \int_0^1 \mathbb{E} \left( f (X^\theta_{t + s v}) (\Xi^\theta_{t + s v}, v) \right) ds.$$

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

$$D_t, \rho_N X^\theta_t \geq \gamma \kappa_t.$$
Next, there exist a constant $C$ such that

$$\rho_N(u) \leq C u^2, \quad |\rho'_N(u)| \leq C|u|,$$

and therefore

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

Then, repeating literally the calculations from Section 4.2, we can obtain a bound similar to (4.11), where $t$ is fixed, but a family of weights $\Xi_{t,\rho_N}, N \geq 1$ is considered instead:

$$\sup_{N \geq 1, \theta \in \hat{\Theta}} E \left| \Xi_{t,\rho_N}^{\theta} \right|^{2+\epsilon} < \infty.$$

Hence the family $\{\Xi_{t,\rho_N}^{\theta}, N \geq 1, \theta \in \hat{\Theta}\}$ is uniformly integrable. It is straightforward to see that

$$\Xi_{t,\rho_N}^{\theta} \rightarrow \Xi_t^{\theta}, \quad n \rightarrow \infty$$

with probability 1 for any sequence $\theta_N \rightarrow \theta \in \Theta$. Combined with the above uniform integrability, this shows that

$$\Xi_{t,\rho_N}^{\theta} \rightarrow \Xi_t^{\theta}, \quad \rightarrow \infty$$

in $L_1(\Omega, \mathbb{P})$ uniformly w.r.t. $\theta \in \hat{\Theta}$. If we additionally assume (4.1), for any $N$ the representation (4.2) with $\rho = \rho_N$ holds true; cf. Section 4.1. Writing this representation in the integral form, and then passing to the limit as $N \rightarrow \infty$, we get (4.13) under the additional assumption (4.1).

Similar argument can be applied to remove the assumption (4.1). Consider a family of processes $Z^L, L \geq 1$ with Lévy measures

$$\mu_L(du) = m_L(u)du, \quad m_L(u) = m(u)e^{-u^2/L}.$$

Because $|u|^{2+\delta}e^{-u^2/L} \leq C$, every $\mu_L$ satisfies (4.1). In addition, it is an easy calculation to show that conditions $H1, H2$ are satisfied for $\mu_L$ uniformly w.r.t. $L \geq 1$. Hence we have the following:

(a) for every $L \geq 1$, (4.13) holds true with $X_t^\theta, \Xi_t^\theta$ replaced by respective $X_t^\theta, \Xi_t^\theta; (b)$ the family $\Xi_t^{\theta,L}, \theta \in \hat{\Theta}, t \in (0,1], L \geq 1$ satisfies an analogue of (4.11) uniformly w.r.t. $L \geq 1$ (to prove this, one should repeat the argument from Section 4.2 uniformly w.r.t. $L \geq 1$). It is straightforward to see that the pairs $X_t^{\theta,L}, \Xi_t^{\theta,L}$ weakly converge to the pair $X_t^\theta, \Xi_t^\theta$ as $L \rightarrow \infty$. Since the family $\{\Xi_t^{\theta,L}\}$ is uniformly integrable by the above property (b), we can pass to the limit in the relation (4.13) for $X_t^{\theta,L}, \Xi_t^{\theta,L}$, and get finally (4.13) for $X_t^\theta, \Xi_t^\theta$.

This proves (4.12) and completes the proof of Lemma 3.1.

4.4. Proof of (3.6). Observe that, because

$$D_tX_t^\theta = \gamma D_tZ_t, \quad D_t^2X_t^\theta = \gamma D_t^2Z_t,$$

the weight $\Xi_t^\theta$ is the functional of the process $Z$, only, and the process $U$ is not involved therein. Hence the right hand side term in the identity

$$g_\epsilon(\theta; X_t^\theta) = E[\Xi_t^\theta | X_t^\theta]$$

depends on $U$ only implicitly, i.e. through the conditional expectation w.r.t. $X_t^\theta$. Recall that, in the proof of Lemma 3.1 we derive an $L_{2+\epsilon}$ bound for $g_\epsilon(\theta; X_t^\theta)$ from the above identity, applying Jensen’s inequality and removing the conditional expectation. Hence literally the same argument, applied to a sequence of processes $X_{\theta,n}$ instead of $X_t^\theta$, derives (3.6) from the bound (4.11) where $n$ is not involved to.
Lemma A.1. Let $Z$ satisfy H1. Then the random variables

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

possesses distribution densities $\phi_{\alpha,t}$, and

$$\sup_{x \in \mathbb{R}, t \in (0,1]} \left( |\phi_{\alpha,t}(x)| + |\phi'_{\alpha,t}(x)| + |\phi''_{\alpha,t}(x)| \right) < \infty. \quad (A.1)$$

**Proof.** Because the law of $\zeta_{\alpha,t}$ is a convolution, it is enough to prove the same statement for $t^{-1/\alpha}Z_t$. Respective distribution densities and their derivatives would have the following representations, provided that the functions in the right hand sides are absolutely integrable:

$$\phi_{\alpha,t}(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i\lambda x} \psi_{\alpha,t}(\lambda) d\lambda,$$
$$\phi'_{\alpha,t}(x) = \frac{1}{2\pi} \int_{\mathbb{R}} (-i\lambda)e^{-i\lambda x} \psi_{\alpha,t}(\lambda) d\lambda,$$
$$\phi''_{\alpha,t}(x) = \frac{1}{2\pi} \int_{\mathbb{R}} (-i\lambda)^2 e^{-i\lambda x} \psi_{\alpha,t}(\lambda) d\lambda.$$

Here $\psi_{\alpha,t}$ is the characteristic function of $t^{-1/\alpha}Z_t$, and

$$|\psi_{\alpha,t}(\lambda)| = \exp \left( t \int_{\mathbb{R}} (\cos(t^{-1/\alpha} \lambda u) - 1)\mu(du) \right) \leq \exp \left( t \int_{t^{-1/\alpha} |\lambda u| < 1} (\cos(t^{-1/\alpha} \lambda u) - 1)\mu(du) \right).$$

Now it is easy to show that, by H1, there exist $c_1, c_2 > 0$ such that

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

which proves both existence of $\phi_{\alpha,t}$ and bound (A.1) for them. □

**References**

[1] Aït-Sahalia, Y. and Jacod, J. (2007), Volatility estimators for discretely sampled Lévy processes. Ann. Statist. 35 (2007), no. 1, 355–392.
[2] Aït-Sahalia, Y. and Jacod, J. (2008), Fisher’s information for discretely sampled Lévy processes. Econometrica 76, 727–761.
[3] Clément, E. and Gloter, A. (2013), Local Asymptotic Mixed Normality property for discretely observed stochastic differential equations driven by stable Lévy processes. hal-00914138. Available at https://hal-upec-upem.archives-ouvertes.fr/hal-00914138
[4] Dohnal, G (1987), On estimating the diffusion coefficient. J. Appl. Probab. 24, 105–114.
[5] Genon-Catalot, V. and Jacod, J. (1993), On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. Ann. Inst. H. Poincaré Probab. Statist. 29, 119–151.
[6] Genon-Catalot, V. and Jacod, J. (1994), Estimation of the diffusion coefficient for diffusion processes: random sampling. Scand. J. Statist. 21, 193–221.
[7] Gobet, E. (2001), Local asymptotic mixed normality property for elliptic diffusion: a Malliavin calculus approach. Bernoulli 7, 899 – 912.
[8] Gobet, E. (2001), LAN property for ergodic diffusions with discrete observations. Ann. Inst. H. Poincaré Probab. Statist. 38, 711–737.
[9] Ibragimov, I. A. and Z. Hasminskii, R. Z. (1981), Statistical estimation: asymptotic theory. New York, Springer-Verlag.
[10] Kessler, M. (1997), Estimation of an ergodic diffusion from discrete observations. Scand. J. Statist. 24, 211–229.
[11] Ivanenko, D. O. and Kulik, A. M. Malliavin calculus approach to statistical inference for Levy driven SDEs. To appear in Methodology and Computing in Applied Probability. arXiv:1301.5142
[12] Ivanenko, D. O. and Kulik, A. M. LAN property for families of distributions of solutions to Levy driven SDEs. Modern Stochastics: Theory and Applications, v. 1, n. 1, p. 33 – 47. arXiv:1308.3089
[13] Jeganathan, P. (1982), On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal. Sankhyà Ser. A 44, 173–212.
[14] Le Cam, L. (1960), 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.

[15] Le Cam, L. and Yang, G. L. (2000), Asymptotics in statistics. Some basic concepts. Second edition. Springer Series in Statistics. Springer-Verlag, New York.

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

[17] Masuda, H. Parametric estimation of Lévy processes. To appear in Lévy Matters IV, Lecture Notes in Mathematics, Springer. arXiv:1409.0292

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

INSTITUTE OF MATHEMATICS FOR INDUSTRY, KYUSHU UNIVERSITY, 744 MOTOOKA, NISHI-KU, FUKUOKA 819-0395, JAPAN
E-mail address: hiroki@imi.kyushu-u.ac.jp