LAN PROPERTY FOR DISCRETELY OBSERVED SOLUTIONS TO LÉVY DRIVEN SDE’S

D. O. IVANENKO AND A. M. KULIK

ABSTRACT. The LAN property is proved in the statistical model based on discrete-time observations of a solution to a Lévy driven SDE. The proof is based on a general sufficient condition for a statistical model based on a discrete observations of a Markov process to possess the LAN property, and involves substantially the Malliavin calculus-based integral representations for derivatives of log-likelihood of the model.

1. Introduction

Consider stochastic equation of the form

\[ dX^\theta_t = a_\theta(X^\theta_t)dt + dZ_t, \]

where \( a : \Theta \times \mathbb{R} \to \mathbb{R} \) is a measurable function, \( \Theta = (\theta_1, \theta_2) \in \mathbb{R} \) is a parametric set. For a given \( \theta \in \Theta \), assuming that the drift term \( a_\theta \) satisfies the standard local Lipschitz and linear growth conditions, Eq. (1) uniquely defines a Markov process \( X \). The aim of this paper is to establish the local asymptotic normality property (LAN in the sequel) in a model, where the process \( X \) is discretely observed with a fixed time discretization value \( h > 0 \), and a number of observation \( n \to \infty \).

The LAN property provides a convenient and powerful tool for establishing lower efficiency bounds in a statistical model, e.g. [6], [17], [18]. Such a property for statistical models, based on discrete observations of processes with Lévy noise, was studied mostly in the cases, where the likelihood function (or, at least its “main part”) is explicit, in a sense, e.g. [1], [2], [7], [12], [13]. In the above references the models are linear in the sense that the process under the observation is either a Lévy process, or a solution of a linear (Ornstein-Uhlenbeck type) SDE driven by a Lévy process. The general non-linear case remains non-studied to a great extent, and apparently the main reason for this is that the transition probability density of the observed Markov process in that case is highly implicit. In this paper we develop tools, convenient for proving the LAN property in the framework of discretely observed solutions to SDE’s with a Lévy noise. To make the exposition reasonably transparent, we confine ourselves to a particular case of one-dimensional and one-parameter model, and a fixed sample frequency \( h \). Various extensions (general state space, multiparameter model, high frequency sampling, etc.) are visible, but we postpone their detailed analysis for a further research.

Our approach consists of two principal parts. On one hand, we design a general sufficient condition for a statistical model based on a discrete observations of a Markov process to possess the LAN property, see Theorem 1 below. This result

---

*Key words and phrases.* LAN property, Likelihood function, Lévy driven SDE, Regular statistical experiment.
extends the classical LeCam’s result about the LAN property for i.i.d. samples, and it closes [5, Theorem 13], with some substantial differences in the basic assumptions, which makes our result well designed to a study of a model based on observations of a Lévy driven SDE, see Remark 1 below. On the other hand, we integral representations of derivatives of 1st and 2nd orders of the log-likelihood are available: our recent papers [11] and [10] we have derived such representations using the Malliavin calculus tools. The virtue of this approach is the same with the one developed in [4] in the diffusion setting, but with substantial changes which comes from non-diffusive structure of the noise. Combination of these two principal parts leads to a required LAN property.

The structure of the paper follows the two-stage scheme outlined above. First we formulate in Section 2.1 (and prove in Section 3) a general sufficient condition for the LAN property in a Markov model. Then we formulate in Section 2.2 (and prove in Section 4) our main result about the LAN property for a discretely observed solution to a Lévy driven SDE; here the proof involves substantially the Malliavin calculus-based integral representations of derivatives of the log-likelihood from [11] and [10].

2. The main results

2.1. LAN property for discretely observed Markov processes. Let X be a Markov process taking its values in a locally compact metric space X. The law of X is assumed to depend on a real-valued parameter θ; in what follows, we assume that the parametric set Θ is an interval (θ₁, θ₂) ∈ ℝ. We denote by P_θ x the law of X with X_0 = x, which corresponds to the parameter value θ; the expectation w.r.t. P_θ x is denoted by E_θ x. For a given h > 0, we denote by P_θ x,n the law w.r.t. P_θ x of the vector X_n = \{X_{hk}, k = 1, \ldots, n\} of discrete time observations of X with the step h. Denote by E_n the statistical experiment generated by the sample X_n with X_0 = x, i.e.

(2) E_n = (X_n, B(X^n), P_θ x,n, θ ∈ Θ);

we refer to [8] for the notation and terminology. Our aim is to establish the LAN property for the sequence of experiments \{E_n\}.

Recall that the sequence of statistical experiments \{E_n\} (or, equivalently, the family \{P_θ x,n, θ ∈ Θ\}) is said to have the LAN property at the point θ₀ ∈ Θ as n → ∞, if for some sequence r(n) > 0, n ≥ 1 and all u ∈ ℝ

Z_{n,θ₀}(u) := \frac{dP_{θ₀+μ(n)}^{X_n}}{dP_{θ₀}^{X_n}}(X^n) = \exp \left\{ Δ_n(θ₀)u - \frac{1}{2}u^2 + \Psi_n(u, θ₀) \right\},

with

(3) \mathcal{L}(Δ_n(θ₀) \mid P_θ x,n) ⇒ N(0, 1), \ n → ∞;

(4) Ψ_n(u, θ₀) \xrightarrow{P_{θ₀}} 0, \ n → ∞.

In what follows we assume that X admits a transition probability density p_h(θ; x, y) w.r.t. some σ-finite measure λ. Furthermore, we assume that the experiment E_1 is regular; that is, for every x ∈ X

(a) the function θ → p_h(θ; x, y) is continuous for λ-a.a. y ∈ X;
(b) the function \( \sqrt{p_h(\theta; x, \cdot)} \) is differentiable in \( L_2(\mathcal{X}, \lambda) \); that is, there exists \( g_h(\theta; x, \cdot) \in L_2(\mathcal{X}, \lambda) \) such that
\[
\int_{\mathcal{X}} \left( \sqrt{p_h(\theta + \delta; x, y)} - \sqrt{p_h(\theta; x, y)} - g_h(\theta; x, y) \right)^2 \lambda(dy) \to 0, \quad \delta \to 0;
\]
(c) the function \( q_h(\theta; x, \cdot) \) is continuous in \( L_2(\mathcal{X}, \lambda) \) w.r.t. \( \theta \); that is,
\[
\int_{\mathcal{X}} \left( q_h(\theta + \delta; x, y) - q_h(\theta; x, y) \right)^2 \lambda(dy) \to 0, \quad \delta \to 0.
\]
Denote
\[
g_h(\theta; x, y) = 2q_h(\theta; x, y)\sqrt{p_h(\theta; x, y)};
\]
note that by the definition of \( q_h \) the function \( g_h \) is well defined and satisfies
\[
E_\theta^x g_h(\theta; x, X_h) = 0
\]
for every \( x \in \mathbb{R}, \theta \in \Theta \). Furthermore, denote
\[
I_n(\theta) = \sum_{k=1}^{n} E_\theta^x \left( g_h(\theta; X_{h(k-1)}, X_{hk}) \right)^2 = 4E_\theta^x \sum_{k=1}^{n} \int_{\mathcal{X}} \left( q_h(\theta; X_{h(k-1)}, y) \right)^2 \lambda(dy).
\]
Assuming that the statistical experiment \( \mathcal{E}_n \) is regular, the above integral is finite and defines the Fisher information for \( \mathcal{E}_n \).

We fix \( \theta_0 \in \Theta \), and put \( r(n) = I_n^{-1/2}(\theta_0) \) for \( n \) large enough, assuming that for those \( n \) one has \( I_n(\theta_0) > 0 \).

**Theorem 1.** Suppose the following.

1. Statistical experiment \( \mathcal{E}_n \) is regular for every \( x \in \mathcal{X} \) and \( n \geq 1 \); for \( n \) large enough \( I_n(\theta_0) > 0 \).
2. The sequence
\[
r(n)\sum_{j=1}^{n} g_h(\theta_0; X_{h(j-1)}, X_{hj}), \quad n \geq 1
\]
is asymptotically normal w.r.t. \( P_{x_0}^\theta \) with parameters \((0, 1)\).
3. The sequence
\[
r^2(n)\sum_{j=1}^{n} g_h^2(\theta_0; X_{h(j-1)}, X_{hj}), \quad n \geq 1
\]
converges to 1 in \( P_{x_0}^\theta \)-probability.
4. There exists a constant \( p > 2 \) such that
\[
\lim_{n \to \infty} r^p(n) E_\theta^x \left( \sum_{j=1}^{n} g_h(\theta_0; X_{h(j-1)}, X_{hj}) \right)^p = 0.
\]
5. For every \( N > 0 \)
\[
\lim_{n \to \infty} \sup_{|v| < N} r^2(n) E_\theta^x \left( \sum_{j=1}^{n} \int_{\mathcal{X}} \left( q_h(\theta_0 + r(n)v; X_{h(j-1)}, y) - q_h(\theta_0; X_{h(j-1)}, y) \right)^2 \lambda(dy) = 0.
\]
Then \( \{P_{x,n}^\theta, \theta \in \Theta\} \) has the LAN property at the point \( \theta_0 \).
Remark 1. The above theorem is closely related to [5, Theorem 13]. One important difference is that in [5] the main conditions are formulated in the terms of the functions
\[ \sqrt{p_h(\theta + t; x, y)/p_h(\theta; x, y)} - 1, \]
while within our approach the main assumptions are imposed on the log-likelihood derivative \( g_h(\theta; x, y) \), and can be verified efficiently e.g. in a model where \( X \) is defined by an SDE with jumps; see Section 2.2 below. Another important difference is that the whole approach in [5] is developed under the assumption that the log-likelihood function smoothly depends on the parameter \( \theta \). For a model where \( X \) is defined by an SDE with jumps, such an assumption may be very restrictive, see the detailed discussion in [11]. This is the reason why we use instead the assumption of regularity of the experiments, which both is much milder and is easily verifiable, see [11].

Let us note briefly two possible extensions of the above result, which can be obtained without any essential changes in the proof. We do not expose them here in details, because they will not be used in the current paper.

Remark 2. The statement of Theorem 1 still holds true if \( h \) is allowed to depend on \( n \), with conditions 1 – 5 respectively changed.

Remark 3. The statement of Theorem 1 still holds true if, instead of one \( \theta_0 \), a sequence \( \theta_n \to \theta_0 \) is considered, with conditions 2 – 5 respectively changed. Moreover, in that case relation (3) and (4) would still hold true if instead of a fixed \( u \) a sequence \( u_n \to u \) is considered. That is, under the uniform version of conditions 2 – 5 the uniform asymptotic normality would hold true; see [8, Definition 2.2].

2.2. LAN property for families of distributions of solutions to Lévy driven SDE’s. We assume that \( Z \) in the SDE (1) is a Lévy process without a diffusion component; that is,
\[ Z_t = c t + \int_0^t \int_{|u| > 1} u \nu(ds, du) + \int_0^t \int_{|u| \leq 1} u \tilde{\nu}(ds, du), \]
where \( \nu \) is a Poisson point measure with the intensity measure \( ds \mu(du) \), and \( \tilde{\nu}(ds, du) = \nu(ds, du) - ds \mu(du) \) is respective compensated Poisson measure. In the sequel, we assume the Lévy measure \( \mu \) to satisfy the following.

**H.**
(i) for some \( \beta > 0 \),
\[ \int_{|u| \geq 1} u^{4+\beta} \mu(du) < \infty; \]
(ii) for some \( u_0 > 0 \), the restriction of \( \mu \) on \([-u_0, u_0] \) has a positive density \( m \in C^2([-u_0, 0) \cup (0, u_0]) \):
(iii) there exists \( C_0 \) such that
\[ |m'(u)| \leq C_0 |u|^{-1} m(u), \quad |m''(u)| \leq C_0 u^{-2} m(u), \quad |u| \in (0, u_0]; \]
(iv)
\[ \left( \log \frac{1}{\varepsilon} \right)^{-1} \mu(\{u : |u| \geq \varepsilon\}) \to \infty, \quad \varepsilon \to 0. \]

One particularly important class of Lévy processes satisfying **H** consists of *tempered \( \alpha \)-stable processes* (see [21]), which arise naturally in models of turbulence [20], economical models of stochastic volatility [9], etc.
Denote by \( C^{k,m}(\mathbb{R} \times \Theta), k, m \geq 0 \) the class of functions \( f : \mathbb{R} \times \Theta \rightarrow \mathbb{R} \) which has continuous derivatives
\[
\frac{\partial^i}{\partial x^i} \frac{\partial^j}{\partial \theta^j} f, \quad i \leq k, \quad j \leq m.
\]

About the coefficient \( a_\theta(x) \) in Eq. (1) we assume the following.

**A.** (i) \( a \in C^{3,2}(\mathbb{R} \times \Theta) \) have bounded derivatives \( \frac{\partial}{\partial x} a, \frac{\partial^2}{\partial x^2} a, \frac{\partial^2}{\partial x \theta} a, \frac{\partial^3}{\partial xxx} a, \frac{\partial^3}{\partial x \theta \theta} a, \frac{\partial^3}{\partial xx \theta} a, \frac{\partial^4}{\partial xxx \theta} a \) and
\[
(10) \quad |a_\theta(x)| + |\frac{\partial a_\theta}{\partial \theta}(x)| + |\frac{\partial^2 a_\theta}{\partial \theta \theta}(x)| \leq C(1 + |x|), \quad \theta \in \Theta, \quad x \in \mathbb{R}.
\]

(ii) For a given \( \theta_0 \in \Theta \), there exists a neighbourhood \((\theta_0 - \epsilon, \theta_0 + \epsilon) \subset \Theta\) of \( \theta_0 \) such that
\[
\limsup_{|x| \to \infty} \frac{a_\theta(x)}{x} < 0 \quad \text{uniformly by } \theta \in (\theta_0 - \epsilon, \theta_0 + \epsilon).
\]

It is proved in [11] that, under conditions A(i) and H, the following properties hold:

- the Markov process \( X \) given by (1) has a transition probability density \( p^\theta_t \) w.r.t. the Lebesgue measure;
- this density has a derivative \( \partial_\theta p^\theta_t(x,y) \), and the statistical experiment (2) is regular;
- the function \( g^\theta_t \), given by (5) satisfies (6).

Hence all the pre-requisites for Theorem 1 given in Section 2.1 are available with \( \lambda(dx) = dx \) (the Lebesgue measure).

Furthermore, under conditions A and H, for \( \theta = \theta_0 \) corresponding Markov process \( X \) is ergodic, i.e. there exists unique invariant probability measure \( \kappa^\theta_{inv} \) for \( X \). One can verify this easily, using conditions, sufficient for ergodicity of solutions to Lévy driven SDE’s, given in [19] and [14]. Denote by \( \{X^{st,\theta_0}_t, t \in \mathbb{R}\} \) corresponding stationary version of \( X \); that is, a Markov process, defined on whole axis \( \mathbb{R} \), which has the transition probabilities with \( X \) and one-dimensional distributions equal to \( \kappa^\theta_{inv} \). Clearly, the existence of such a process, on a proper probability space, is guaranteed by the Kolmogorov consistency theorem. Denote
\[
(11) \quad \sigma^2(\theta_0) = \mathbb{E}\left( g_h(\theta_0; X^{st,\theta_0}_0, X^{st,\theta_0}_h) \right)^2 = \int_{\mathbb{R}} \int_{\mathbb{R}} \left( g_h(\theta_0; x, y) \right)^2 p_h(\theta_0; x, y) d\kappa^\theta_{inv}(dx).
\]

The following theorem performs the main result of this paper. Its proof is given in Section 3 below.

**Theorem 2.** Let conditions A and H hold true, and
\[
\sigma^2(\theta_0) > 0.
\]

Then the family \( \{P^\theta_{x,n}, \theta \in \Theta\} \) possesses the LAN property at the point \( \theta = \theta_0 \).

### 3. Proof of Theorem 1

The method of proof goes back to LeCam’s proof of the LAN property for i.i.d. samples, see e.g. Theorem II.1.1 and Theorem II.3.1 in [8]. In the Markov setting, the dependence in the observations lead to some additional technicalities; see e.g. [19]. The possible ways to overcome these additional difficulties can be found, in a slightly different setting, in the proof of [5] Theorem 13. In order to keep
the exposition transparent and self-sufficient, we prefer to give a complete proof of Theorem 1 explicitly, rather than to give a chain of partly relevant references.

We separate the proof into several lemmas; in all the lemmas in this section we assume the conditions of Theorem 1 to be fulfilled. Values $x, \theta_0$, and $u$ are fixed; we assume that $n$ is large enough, so that $\theta_0 + r(n)u \in \Theta$. In order to simplify the notation below we write $\theta$ instead of $\theta_0$.

Denote
\[
\zeta_{jn}(u) = \left( \frac{p_h(\theta + r(n)u; X_{h(j-1)}, X_{hj})}{p_h(\theta; X_{h(j-1)}, X_{hj})} \right)^{1/2} - 1 \right) I \{ p_h(\theta; X_{h(j-1)}, X_{hj}) \neq 0 \}.
\]

**Lemma 1.** One has
\[
\limsup_{n \to \infty} \sum_{j=1}^{n} E_x^\theta (\zeta_{jn}(u))^2 \leq \frac{1}{4} u^2 \tag{12}
\]
and
\[
\lim_{n \to \infty} \sum_{j=1}^{n} E_x^\theta \left( \zeta_{jn}(u) - \frac{1}{2} r(n) u g_h(\theta; X_{h(j-1)}, X_{hj}) \right)^2 = 0. \tag{13}
\]

**Proof.** By the regularity of $\mathcal{E}_1$ and the Cauchy inequality we have
\[
E_x^\theta \left( \zeta_{jn}(u) - \frac{1}{2} r(n) u g_h(\theta; X_{h(j-1)}, X_{hj}) \right)^2 \leq (r(n)u)^2 E_x^\theta \int_{\mathbb{R}} \left( \int_0^1 q_h(\theta + r(n)uv; X_{h(j-1)}, y) - q_h(\theta; X_{h(j-1)}, y) \right) \lambda(dy) dv
\]
\[
\leq (r(n)u)^2 E_x^\theta \int_{\mathbb{R}} \lambda(dy) \int_0^1 \left( q_h(\theta + r(n)uv; X_{h(j-1)}, y) - q_h(\theta; X_{h(j-1)}, y) \right) dv.
\]
This and (10) yield (13). To deduce (12) from (13), recall an elementary inequality
\[
|AB| \leq \frac{\alpha}{2} A^2 + \frac{1}{2\alpha} B^2, \quad \alpha > 0,
\]
and write
\[
\zeta_{jn}(u) = \frac{1}{2} r(n) u g_h(\theta; X_{h(j-1)}, X_{hj}) + \left( \zeta_{jn}(u) - \frac{1}{2} r(n) u g_h(\theta; X_{h(j-1)}, X_{hj}) \right) =: A + B.
\]
Then
\[
E_x^\theta (\zeta_{jn}(u))^2 \leq (1 + \alpha) \frac{1}{4} u^2 r^2(n) E_x^\theta (g_h(\theta; X_{h(j-1)}, X_{hj}))^2
\]
\[
+ \left( 1 + \frac{1}{\alpha} \right) E_x^\theta \left( \zeta_{jn}(u) - \frac{1}{2} r(n) u g_h(\theta; X_{h(j-1)}, X_{hj}) \right)^2.
\]
Because by the construction

\begin{equation}
\sum_{j=1}^{n} E_{\alpha}^{\theta} \left( g_{h}(\theta; X_{h(j-1)}, X_{h(j)}) \right)^{2} = I_{n}(\theta) = r^{2}(n),
\end{equation}

this leads to the bound

\[ \limsup_{n \to \infty} \sum_{j=1}^{n} E_{\alpha}^{\theta} \left( c_{j \alpha}^{\theta}(u) \right)^{2} \leq \frac{1 + \alpha}{4u^2}. \]

Since \( \alpha > 0 \) is arbitrary, this completes the proof. \( \square \)

**Lemma 2.** One has

\begin{equation}
\sum_{j=1}^{n} \left( c_{j \alpha}^{\theta}(u) \right)^{2} \to \frac{u^2}{4}, \quad n \to \infty
\end{equation}

in \( P_{\alpha}^{\theta} \)-probability.

**Proof.** By the Chebyshev inequality,

\[ P_{\alpha}^{\theta} \left\{ \sum_{j=1}^{n} \left| \frac{1}{4} r^{2}(n) u^{2} \sum_{j=1}^{n} \left( g_{h}(\theta; X_{h(j-1)}, X_{h(j)}) \right) \right| \geq \varepsilon \right\} \]

\[ \leq \frac{1}{\varepsilon} \sum_{j=1}^{n} E_{\alpha}^{\theta} \left| \frac{1}{4} r^{2}(n) u^{2} g_{h}(\theta; X_{h(j-1)}, X_{h(j)}) \right| \]

\[ = \frac{1}{\varepsilon} \sum_{j=1}^{n} E_{\alpha}^{\theta} \left| c_{j \alpha}^{\theta}(u) \right| \frac{1}{2} r(n) u g_{h}(\theta; X_{h(j-1)}, X_{h(j)}) \left| c_{j \alpha}^{\theta}(u) \right| \frac{1}{2} r(n) u g_{h}(\theta; X_{h(j-1), X_{h(j)}}) \]

which by (12), for a given \( \alpha > 0 \), is dominated by

\[ \frac{1}{2 \alpha \varepsilon} \sum_{j=1}^{n} E_{\alpha}^{\theta} \left( \frac{1}{2} r(n) u g_{h}(\theta; X_{h(j-1)}, X_{h(j)}) \right)^{2} \]

\[ + \frac{\alpha}{2 \varepsilon} \sum_{j=1}^{n} E_{\alpha}^{\theta} \left( \frac{1}{2} r(n) u g_{h}(\theta; X_{h(j-1), X_{h(j)}}) \right)^{2}. \]

By Cauchy inequality together with (12) and (15) imply that for the second one the upper limit does not exceed

\[ \limsup_{n \to \infty} \left( \frac{\alpha}{\varepsilon} \sum_{j=1}^{n} E_{\alpha}^{\theta} \left( c_{j \alpha}^{\theta}(u) \right)^{2} + \frac{\alpha u^{2}}{2 \varepsilon} r^{2}(n) \sum_{j=1}^{n} E_{\alpha}^{\theta} \left( g_{h}(\theta; X_{h(j-1)}, X_{h(j)}) \right)^{2} \right) \leq \frac{3 \alpha u^{2}}{2 \varepsilon}. \]

Since \( \alpha > 0 \) is arbitrary, this proves that the difference

\[ \sum_{j=1}^{n} \left( c_{j \alpha}^{\theta}(u) \right)^{2} - \frac{1}{4} r^{2}(n) u^{2} \sum_{j=1}^{n} \left( g_{h}(\theta; X_{h(j-1)}, X_{h(j)}) \right)^{2} \]

tends to 0 in \( P_{\alpha}^{\theta} \)-probability. Combined with the condition 3 of Theorem 1, this gives the required statement. \( \square \)
Lemma 3. One has
\[ \max_{1 \leq j \leq n} |\zeta_{jn}(u)| \to 0, \quad n \to \infty \]
in $P_\theta$-probability.

Proof. We have
\[
P_\theta^\theta \left\{ \max_{1 \leq j \leq n} |\zeta_{jn}(u)| > \varepsilon \right\} \leq \sum_{j=1}^{n} P_\theta^\theta \left\{ |\zeta_{jn}(u)| > \varepsilon \right\}
\leq \sum_{j=1}^{n} P_\theta^\theta \left\{ |\zeta_{jn}(u)| - \frac{1}{2} r(n) u g_\theta(\theta; X_{h(j-1)}, X_{hj}) > \frac{\varepsilon}{2} \right\}
+ \sum_{j=1}^{n} P_\theta^\theta \left\{ |g_\theta(\theta; X_{h(j-1)}, X_{hj})| > \frac{\varepsilon}{4 r(n) |u|} \right\}.
\]
The first sum in the r.h.s. of this inequality vanishes as $n \to \infty$ because of (13),
the second sum vanishes because of the condition 4 of Theorem 1.
\[ \square \]

Corollary 1. By Lemma 3 and Lemma 2, we have
\[ \sum_{j=1}^{n} |\zeta_{jn}(u)|^3 \to 0, \quad n \to \infty \]
in $P_\theta$-probability.

Because of the Markov structure of the sample, in addition to Lemma 2 we will need the following statement. Denote
\[ F_j = \sigma(X_{hi}, i \leq j), \quad E_{x,j}^\theta = E_x[\cdot|F_j]. \]

Lemma 4. One has
\[ \sum_{j=1}^{n} E_{x,j-1}^\theta (\zeta_{jn}(u))^2 \to \frac{u^2}{4}, \quad n \to \infty \]
in $P_\theta$-probability.

Proof. Denote
\[ \chi_{jn} = (\zeta_{jn}(u))^2 - E_{x,j-1}^\theta (\zeta_{jn}(u))^2, \quad S_n = \sum_{j=1}^{n} \chi_{jn}, \]
then by (16) it is enough to prove that $S_n \to 0$ in $P_\theta^\theta$-probability. Fix $\varepsilon > 0$, and put
\[ \chi_{jn}^\varepsilon = (\zeta_{jn}(u))^2 1_{|\zeta_{jn}(u)| \leq \varepsilon} - E_{x,j-1}^\theta ( (\zeta_{jn}(u))^2 1_{|\zeta_{jn}(u)| \leq \varepsilon} ), \quad S_n^\varepsilon = \sum_{j=1}^{n} \chi_{jn}^\varepsilon. \]
By the construction $\{ \chi_{jn}^\varepsilon, j = 1, \ldots, n \}$ is a martingale difference, hence
\[ E_x^\theta (S_n^\varepsilon)^2 = \sum_{k=1}^{n} E_x^\theta (\chi_{jn}^\varepsilon)^2 \leq \sum_{k=1}^{n} E_x^\theta ( (\zeta_{jn}(u))^2 1_{|\zeta_{jn}(u)| \leq \varepsilon} )^2 \leq \varepsilon^2 E_x^\theta \sum_{k=1}^{n} (\zeta_{jn}(u))^2, \]
Hence by (12) and the Cauchy inequality,
\begin{equation}
\limsup_{n \to \infty} E_x |S_n^\varepsilon| \leq \frac{\varepsilon |u|}{2}
\end{equation}

Now, let us estimate the difference $S_n - S_n^\varepsilon$. Note that, using the first statement in Lemma 1, one can improve the statement of Lemma 2 and show that the convergence (16) holds true in $L_1(P_\theta)$; see e.g. Theorem A.I.4 in [8]. In particular, this means that the sequence
$$\sum_{j=1}^{n} (\zeta_{jn}^\theta(u))^2, \quad n \geq 1$$
is uniformly integrable. Hence, because by Lemma 3 the probabilities of the sets
\begin{equation}
\Omega^\varepsilon_n = \left\{ \max_{j \leq n} |\zeta_{jn}| > \varepsilon \right\}
\end{equation}
tend to zero as $n \to \infty$, we have
$$E_x \left( 1_{\Omega^\varepsilon_n} \sum_{j=1}^{n} (\zeta_{jn}^\theta(u))^2 \right) \to 0.$$

One has
$$\chi_{jn} - \chi_{jn}^\varepsilon = (\zeta_{jn}^\theta(u))^2 1_{|\zeta_{jn}^\theta(u)| \geq \varepsilon} - E_{x,j}^\theta (\zeta_{jn}^\theta(u))^2 1_{|\zeta_{jn}^\theta(u)| > \varepsilon},$$

hence
$$E_x^\theta |S_n - S_n^\varepsilon| \leq 2 \sum_{j=1}^{n} E_x (\zeta_{jn}^\theta(u))^2 1_{|\zeta_{jn}^\theta(u)| \geq \varepsilon} \leq 2 E_x^\theta \left( 1_{\Omega_n^\varepsilon} \sum_{j=1}^{n} (\zeta_{jn}^\theta(u))^2 \right) \to 0.$$

Together with (20) this gives
$$\limsup_{n \to \infty} E_x |S_n| \leq \frac{\varepsilon |u|}{2},$$
which completes the proof because $\varepsilon > 0$ is arbitrary.

The final preparatory result we require is the following.

**Lemma 5.** One has
\begin{equation}
2 \sum_{j=1}^{n} \zeta_{jn}^\theta u - r(n) \sum_{j=1}^{n} g_h(\theta; X_{h(j-1)}, X_{hj}) \to - \frac{u^2}{4}, \quad n \to \infty
\end{equation}
in $P_\theta$-probability.

**Proof.** We have the equality
$$(\zeta_{jn}^\theta(u))^2 = \frac{p_h(\theta + r(n)u; X_{h(j-1)}, X_{hj})}{p_h(\theta; X_{h(j-1)}, X_{hj})} - 1 - 2 \zeta_{jn}^\theta(u)$$
valid $P_\theta$-a.s. Note that by the Markov property of $X$ one has
$$E_{x,j-1}^\theta \frac{p_h(\theta + r(n)u; X_{h(j-1)}, X_{hj})}{p_h(\theta; X_{h(j-1)}, X_{hj})} \frac{p_h(\theta + r(n)u; X_{h(j-1)}, y)}{p_h(\theta; X_{h(j-1)}, y)} p_h(\theta; X_{h(j-1)}, y) \lambda(dy) = 1;$$
hence by Lemma 4 one has that
\[ \sum_{j=1}^{n} E_{x,j-1}^\theta \sigma_{jn}^\theta (u) \to \frac{u^2}{8} \]
in $P_{x}^\theta$-probability. Therefore, what we have to prove in fact is that
\[ V_n := 2 \sum_{j=1}^{n} (\sigma_{jn}^\theta (u) - E_{x,j-1}^\theta \sigma_{jn}^\theta (u)) - r(n)u \sum_{j=1}^{n} g_h(\theta; X_{h(j-1)}, X_{hj}) \to 0 \]
in $P_{x}^\theta$-probability. By (4) the sequence
\[ \zeta_{jn}^\theta (u) - E_{x,j-1}^\theta \zeta_{jn}^\theta (u) - r(n)ug_h(\theta; X_{h(j-1)}, X_{hj}), \quad j = 1, \ldots, n \]
is a martingale difference, hence
\[ E_{x}^\theta V_n^2 \leq 4 \sum_{j=1}^{n} E_{x}^\theta \left( \zeta_{jn}^\theta (u) - \frac{1}{2} r(n)ug_h(\theta; X_{h(j-1)}, X_{hj}) \right)^2, \]
which tends to zero as $n \to \infty$ by (13).

Now, we can finalize the proof of Theorem 1. Fix $\varepsilon \in (0,1)$ and consider the sets $\Omega_n^\varepsilon$ defined by (24); by Lemma 3 we have $P_{x}^\theta (\Omega_n^\varepsilon) \to 0$. Using the Taylor expansion for the function $\log(1+x)$, we obtain that there exist a constant $C_\varepsilon$ and random variables $\alpha_{jn}$ such that $|\alpha_{jn}| < C_\varepsilon$, for which the following identity holds true outside of the set $\Omega_n^\varepsilon$:
\[ \sum_{j=1}^{n} \log \frac{p_h(\theta + r(n)u; X_{h(j-1)}, X_{hj})}{p_h(\theta; X_{h(j-1)}, X_{hj})} = 2 \sum_{j=1}^{n} \zeta_{jn}^\theta (u) - \sum_{j=1}^{n} (\zeta_{jn}^\theta (u))^2 + \sum_{j=1}^{n} \alpha_{jn}|\zeta_{jn}^\theta (u)|^3. \]

Then by Lemma 2, Lemma 5, and Corollary 11 we have
\[ \log Z_{n,\theta}(u) = \sum_{j=1}^{n} \log \frac{p_h(\theta + r(n)u; X_{h(j-1)}, X_{hj})}{p_h(\theta; X_{h(j-1)}, X_{hj})} \]
\[ = r(n)u \sum_{j=1}^{n} g_h(\theta; X_{h(j-1)}, X_{hj}) - \frac{u^2}{4} - \frac{u^2}{4} + \Psi_n, \]
where $\Psi_n \to 0$ in $P_{x}^\theta$-probability. By the asymptotic normality condition 2, this completes the proof.

**4. Proof of Theorem 2**

To prove Theorem 2, we verify the conditions of Theorem 1. First, let us give an auxiliary result, which will be used repeatedly in the proof.

**Lemma 6.** Under conditions A and H for every $p \in (2, 4 + \beta)$ there exists a constant $C$ such that for all $x \in \mathbb{R}$, $\theta \in (\theta_-, \theta_+)$, and $t \geq 0$
\[ E_{x}^\theta \left| g_h(\theta; x, X_t) \right|^p \leq C(1 + |x|^p), \quad E_{x}^\theta |X_t|^p \leq C(1 + |x|^p). \]

**Proof.** The first inequality is proved in Lemma 1 [11]. One can prove the second inequality, using a standard argument based on the Lyapunov condition for the function $V(x) = |x|^p$; e.g. Proposition 4.1 [14]. \[ \square \]
Recall (e.g. [14], Section 3.2) that one standard way to construct the invariant measure \( \kappa_{inv}^{\theta_0} \) is to take a weak limit point (as \( T \to \infty \)) for the family of Khas’minskii’s averages

\[
\kappa_T^{\theta_0}(dy) = \frac{1}{T} \int_0^T P_0^{\theta_0}(X_t \in dy) dt.
\]

Then, by the Fatou lemma, the second relation in (23) implies the following moment bound for \( \kappa_{inv}^{\theta_0} \).

**Corollary 2.** For every \( p \in (2, 4 + \beta) \),

\[
\int_{\mathbb{R}} |y|^p \kappa_{inv}^{\theta_0}(dy) < \infty.
\]

Everywhere below we assume conditions of Theorem 2 to hold true.

**Lemma 7.** The sequence

\[
\frac{1}{\sqrt{n}} \sum_{j=1}^n g_h (\theta_0; X_{h(j-1)}, X_{hj}), \quad n \geq 1
\]

is asymptotically normal w.r.t. \( P_0^{\theta_0} \) with parameters \( (0, \sigma^2(\theta_0)) \), where \( \sigma^2(\theta_0) \) is defined in (11).

**Proof.** The idea of the proof is similar to the one of the proof of Theorem 3.3 [16]. Denote

\[
Q_n(\theta_0, X) = \frac{1}{\sqrt{n}} \sum_{j=1}^n g_h (\theta_0; X_{h(j-1)}, X_{hj}).
\]

By Theorem 2.2 [19] (see also Theorem 1.2 [14]), the \( \alpha \)-mixing coefficient \( \alpha(t) \) for the stationary version of the process \( X \) does not exceed \( C_3 e^{-C_4 t} \), where \( C_3, C_4 \) are some positive constants. Then by CLT for stationary sequences (Theorem 18.5.3 [9]) and (23) we have

\[
Q_n(\theta_0, X^{st, \theta_0}) \Rightarrow \mathcal{N}(0, \tilde{\sigma}^2(\theta_0)), \quad n \to \infty
\]

with

\[
\tilde{\sigma}^2(\theta_0) = \sum_{k=-\infty}^{+\infty} \mathbb{E} \left( g_h \left( \theta_0; X_0^{st, \theta_0}, X_k^{st, \theta_0} \right) g_h \left( \theta; X_{h(k-1)}^{st, \theta_0}, X_h^{st, \theta_0} \right) \right).
\]

Furthermore, under conditions of Theorem 2 there exists an exponential coupling for the process \( X \); that is, a two-component process \( Y = (Y^1, Y^2) \), possibly defined on another probability space, such that \( Y^1 \) has the distribution \( P_0^{\theta_0} \), \( Y^2 \) has the same distribution with \( X^{st, \theta_0} \), and for all \( t > 0 \)

\[
P(Y^1_t \neq Y^2_t) \leq C_1 e^{-C_2 t}
\]

(24)
with some constants $C_1, C_2$. The proof of this fact can be found in [13] (Theorem 2.2). Then for any Lipschitz continuous function $f : \mathbb{R} \to \mathbb{R}$ we have

$$|E^\theta_x f(Q_n(\theta_0, X)) - Ef(Q_n(\theta_0, X^{st, \theta_0}))| = |Ef(Q_n(\theta_0, Y^1)) - Ef(Q_n(\theta_0, Y^2))|$$

$$= \text{Lip}(f) |Ef(Q_n(\theta_0, Y^1) - Q_n(\theta_0, Y^2)|$$

$$\leq \frac{\text{Lip}(f)}{\sqrt{n}} \sum_{k=1}^n E \left|g_h(\theta_0; Y_{h(k-1)}^1, Y_{h(k)}^k) - g_h(\theta_0; Y_{h(k-1)}^2, Y_{h(k)}^k)\right| 1_{(Y_{h(k-1)}^1, Y_{h(k)}^k) \neq (Y_{h(k-1)}^2, Y_{h(k)}^2)}$$

$$\leq \frac{2\text{Lip}(f)}{\sqrt{n}} \sum_{k=1}^n \left(E \left|g_h\left(\theta_0; Y_{h(k-1)}^1, Y_{h(k)}^k\right)\right|^p + E \left|g_h\left(\theta_0; Y_{h(k-1)}^2, Y_{h(k)}^k\right)\right|^p\right)^{1/p}$$

$$\times \left(P(Y_{h(k-1)}^1 \neq Y_{h(k-1)}^2) + P(Y_{h(k)}^1 \neq Y_{h(k)}^2)\right)^{1/q},$$

where $p, q > 1$ are such that $1/p + 1/q = 1$. Since $Y^1$ has the distribution $P^\theta_x$, by [23] we have for $p \in (2, 4 + \beta)$

$$\left|E^\theta_x g_h\left(\theta_0; Y_{h(k-1)}^1, Y_{h(k)}^k\right)^p\right| = E^\theta_x \left|g_h\left(\theta_0; X_{h(k-1)}, X_{h(k)}\right)^p\right|$$

$$\leq CE^\theta_x \left(1 + \left|X_{h(k-1)}\right|^p\right) \leq C + C^2(1 + \left|x\right|^p).$$

Similarly,

$$\left|E^\theta_x g_h\left(\theta_0; Y_{h(k-1)}^2, Y_{h(k)}^k\right)^p\right| = E \left|g_h\left(\theta_0; X_{h(k-1)}^{st, \theta_0}, X_{h(k)}^{st, \theta_0}\right)^p\right|$$

$$\leq CE \left(1 + \left|X_{h(k-1)}^{st, \theta_0}\right|^p\right) = C + C \int_{\mathbb{R}} |y|^p x_{inv}^{\theta_0}(dy),$$

and the constant in the right hand side is finite by Corollary [2]. Hence [24] and [25] yield that

$$E^\theta_x f(Q_n(\theta_0, X)) \to Ef(\xi), \quad n \to \infty, \quad \xi \sim \mathcal{N}(0, \sigma^2(\theta_0))$$

for every Lipschitz continuous function $f : \mathbb{R} \to \mathbb{R}$. This means that the sequence $Q_n(\theta_0, X), n \geq 1$ is asymptotically normal w.r.t. $P^\theta_x$ with parameters $(0, \sigma^2(\theta_0))$.

To conclude the proof, it remains to show that $\sigma^2(\theta_0) = \sigma^2(\theta_0)$. This follows easily from [6] because, by the Markov property of $X_{h(k-1)}^{st, \theta_0}$,

$$\sigma^2(\theta_0) = \sigma^2(\theta_0) + 2 \sum_{k=1}^\infty E \left(g_h\left(\theta_0; X_{0}^{st, \theta_0}, X_{h(k-1)}^{st, \theta_0}\right) g_h\left(\theta; X_{h(k-1)}^{st, \theta_0}, X_{h(k)}^{st, \theta_0}\right)\right)$$

$$= \sigma^2(\theta_0) + 2 \sum_{k=1}^\infty E \left[g_h(\theta; X_{0}^{st, \theta_0}, X_{h(k)}^{st, \theta_0}) \left(E^\theta_x g_h(\theta_0; x, X_{h(k)}^{st, \theta_0})\right)_{x=X_{h(k-1)}^{st, \theta_0}}\right].$$

Similarly, one can prove that

$$\frac{1}{n} \sum_{j=1}^n \left(g_h\left(\theta_0; X_{h(j-1)}, X_{h(j)}\right)\right)^2 \to \sigma^2(\theta_0), \quad n \to \infty.$$
in $L_1(\mathbb{P}^n_\theta)$: the argument is completely the same, with the CLT for a stationary sequence replaced by the Birkhoff-Khinchin ergodic theorem (we omit the details). Hence

\begin{equation}
I_n(\theta_0) \sim n \sigma^2(\theta_0), \quad r(n) \sim \frac{1}{\sqrt{n} \sigma(\theta_0)}, \quad n \to \infty.
\end{equation}

Therefore conditions 2 – 4 of Theorem 1 are verified. Condition 1 of Theorem 1 also holds true: regularity property is proved in [11], and positivity of $I_n(\theta)$ follows from (28).

Let us prove (9), which then would allow us to apply Theorem 1. It is proved in [10] that, under the conditions of Theorem 2, the function $q_h(\theta, x, y)$ is $L_2$-differentiable w.r.t. $\theta$, and

$$\partial_\theta q_h = \frac{1}{2} (\partial_\theta g_h) \sqrt{p_h} + \frac{1}{4} (g_h)^2 \sqrt{p_h}.$$ 

In addition, it is proved therein that for every $\gamma \in [1, 2 + \beta/2)$

\begin{equation}
E^\theta_x \left| \partial_\theta g_h(\theta; x, X_h) \right|^\gamma \leq C(1 + |x|)^\gamma.
\end{equation}

Then

$$E^\theta_x \int_R \left( q_h(\theta + r(n)v, X_{h(j-1)}), y) - q_h(\theta, X_{h(j-1)}), y) \right)^2 dy 
\leq r(n)v E^\theta_x \int_0^{r(n)v} \left( \partial_\theta g_h(\theta + s, X_{h(j-1)}), y) \right)^2 ds 
\leq \frac{r(n)v}{4} E^\theta_x \int_0^{r(n)v} ds \int_R \left( \partial_\theta g_h(\theta + s; X_{h(j-1)}), y) + \frac{1}{2} g_h(\theta + s; X_{h(j-1)}), y) \right)^2 
\times p_h(X_{h(j-1)}), y)dy 
\leq Cr(n)^2 v^2 E^\theta_x \left( 1 + (X_{h(j-1)})^4 \right) ;$$

in the last inequality we have used (29) and the first relation in (23). Using the second relation in (23), we get then

$$\sup_{|v| < N} r(n)^2 E^\theta_x \sum_{j=1}^n E^\theta_x \int_R \left( q_h(\theta + r(n)v, X_{h(j-1)}), y) - q_h(\theta, X_{h(j-1)}), y) \right)^2 dy \leq CN^2 nr(n)^4$$

with a constant $C$ that depends only on $x$. This relation together with (28) completes the proof.

**Acknowledgements**

The authors are deeply grateful to H. Masuda for a valuable bibliographic help and useful discussion.

**References**

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

[2] M. G. Akritas and R. A. Johnson. Asymptotic inference in Lévy processes of the discontinuous type. Ann. Statist., 9:604 – 614, 1981

[3] P. Carr, H. Geman, D.B. Madan, M. Yor. Stochastic volatility for Lévy processes. Math. Finance, 13:345-382, 2003.

[4] E. Gobet. Local asymptotic mixed normality property for elliptic diffusion: a Malliavin calculus approach. Bernoulli, 7(6):899-912, 2001.
[5] P.E. Greenwood, A.N. Shiryaev. *Contiguity and the statistical invariance principle*. London, Th. and Appl. of Stoch., Proc., 1985.

[6] J. Hajek, Local asymptotic minimax admissibility in estimation, in: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability. *Berkeley and Los Angeles, Univ. of California Press*, 175 - 194, 1971.

[7] R. Höpfner. Two comments on parameter estimation in stable processes. Mathematical methods of statistics , 6:125 – 134, 1997.

[8] I.A. Ibragimov and R. Z. Hasminskii. *Statistical estimation: asymptotic theory*. New York, Springer-Verlag, 1981.

[9] I.A. Ibragimov and Yu.V. Linnik, *Independents and stationary associated variables*. Moscow, Nauka, 1965 (In Russian).

[10] D.O. Ivanenko. Stochastic derivatives of solution to Lévy driven SDE’s. *To appear in Visnyk Kyiv Nat. Univ.*

[11] D.O. Ivanenko and A.M. Kulik. Malliavin calculus approach to statistical inference for Lévy driven SDE’s. *To appear in Meth. and Comp. in Appl. Prob., preprint available at arXiv:1301.5141*

[12] R. Kawai and H. Masuda. Local asymptotic normality for normal inverse Gaussian Lévy processes with high-frequency sampling. *ESAIM: Probability and Statistics* 17, 2013.

[13] A. Kohatsu-Higa, E. Nualart, and Ngoc Khue Tran. LAN property for a linear model with jumps *arXiv:1402.4956*

[14] A.M. Kulik. Exponential ergodicity of the solutions to SDE’s with a jump noise. *Stoch. Proc. and their Appl.*, 119:602-632, 2009.

[15] A.M. Kulik. Asymptotic and spectral properties of exponentially φ-ergodic Markov processes. *Stoch. Proc. and their Appl.*, 121:1044-1075, 2011.

[16] A.M. Kulik and N.N. Leonenko. Ergodicity and mixing bounds for the Fisher-Snedecor diffusion. *Bernoulli*, In Press, 2011.

[17] L. Le Cam, Limits of experiments, in: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability. *Berkeley and Los Angeles, Univ. of California Press*, 245261, 1971.

[18] L. Le Cam and G.L. Yang, Asymptotics in Statistics. *Springer*, 1990.

[19] H. Masuda. Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps. *Stoch. Proc. Appl.*, 117:35-56, 2007.

[20] E.A. Novikov. Infinitely divisible distributions in turbulence. *Phys. Rev. E*, 50:R3303-R3305, 1994.

[21] J. Rosiński. Tempering stable processes. *Stoch. Proc. and their Appl.*, 117(6):677-707, 2007.