Bernoulli **29**(3), 2023, 2342–2366
https://doi.org/10.3150/22-BEJ1544

Local asymptotic normality for ergodic jump-diffusion processes via transition density approximation

TEPPEI OGIHARA[1,a] and YUMA UEHARA[2,b]

[1] *Graduate School of Information Science and Technology, University of Tokyo, Tokyo, Japan,*
[a] *ogihara@mist.i.u-tokyo.ac.jp*
[2] *Department of Mathematics, Faculty of Engineering Science, Kansai University, Osaka, Japan,*
[b] *y-uehara@kansai-u.ac.jp*

We show local asymptotic normality (LAN) for a statistical model of discretely observed ergodic jump-diffusion processes, where the drift coefficient, diffusion coefficient, and jump structure are parametrized. Under the LAN property, we can discuss the asymptotic efficiency of regular estimators, and the quasi-maximum-likelihood and Bayes-type estimators proposed in Shimizu and Yoshida (*Stat. Inference Stoch. Process.* **9** (2006) 227–277) and Ogihara and Yoshida (*Stat. Inference Stoch. Process.* **14** (2011) 189–229) are shown to be asymptotically efficient in this model. Moreover, we can construct asymptotically uniformly most powerful tests for the parameters. Unlike with a model for diffusion processes, Aronson-type estimates of the transition density functions do not hold, which makes it difficult to prove LAN. Therefore, instead of Aronson-type estimates, we employ the idea of Theorem 1 in Jeganathan (*Sankhyā Ser. A* **44** (1982) 173–212) and use the L^2 regularity condition. Moreover, we show that local asymptotic mixed normality of a statistical model is implied from that for a model generated by approximated transition density functions under suitable conditions. Together with density approximation by means of thresholding techniques, the LAN property for the jump-diffusion processes is proved.

Keywords: Asymptotically efficient estimator; asymptotically uniformly most powerful test; jump-diffusion processes; local asymptotic mixed normality; L^2 regularity condition; Malliavin calculus; thresholding techniques

1. Introduction

Local asymptotic normality (LAN) is an important property in asymptotic statistical theory because it enables us to discuss asymptotic efficiency of estimators for parametric models. Hájek [14,15] showed the convolution theorem and the minimax theorem for statistical models that satisfy the LAN property. Both theorems give different concepts of asymptotic efficiency. The LAN property has mainly been studied for statistical models of independent observations. In subsequent work, this property has been extended to local asymptotic mixed normality (LAMN) so that we can address a wider class of statistical models. Jeganathan [18] showed the convolution theorem and the minimax theorem under the LAMN property. LAN and LAMN have enabled various studies of statistical methods in addition to the efficiency of estimators. Several works have studied the construction of asymptotically uniformly most powerful tests under LAN or LAMN (see e.g. Choi, Hall, and Schick [5] and Basawa and Scott [3]). Moreover, Eguchi and Masuda [9] studied the model selection problem via Schwartz-type Bayesian information criteria (BIC) and showed model selection consistency of the BIC when the statistical model is locally asymptotically quadratic (which includes the case of LAMN).

For statistical models of discrete observations of semimartingales, Gobet [12,13] showed the LAN and LAMN properties for diffusion processes in the high-frequency limit of observations on a fixed interval and on a growing observation window, respectively. Related to processes with jumps, Aït-Sahalia

1350-7265

and Jacod [1] showed the LAN property for some classes of Lévy processes, including symmetric stable processes, Kawai and Masuda [19] showed the LAN property for normal inverse Gaussian Lévy processes, and Clément and Gloter [7] proved the LAMN property for a stochastic differential equation driven by a pure jump Lévy process whose Lévy measure is an α-stable Leśvy measure near the origin. Statistical models of jump-diffusion processes have also been studied in several papers; Kohatsu, Nualart, and Tran [20] showed the LAN property for ergodic jump-diffusion processes whose drift coefficient depends on an unknown parameter, and Clément, Delattre, and Gloter [6] studied the LAMN property for stochastic differential equations with jumps when the unknown parameter determines the jump structure and the jump times are deterministic and given. Jump-diffusion processes are used for modeling various stochastic phenomena in many areas, such as econometrics, physics, and neuroscience; among the vast literature, we refer the reader to Rao [28] and Cont and Tankov [8] and the references therein. To our knowledge, no studies have shown the LAN property for jump-diffusion processes with the drift coefficient, the diffusion coefficient, and the jump structure all parametrized, and herein we focus on such a situation.

In the proofs of the LAN properties for diffusion processes by Gobet [12,13], it is crucial that transition density functions satisfy estimates from above and below by Gaussian density functions up to constants: so-called Aronson-type estimates. Unlike diffusion processes, jump-diffusion processes do not satisfy Aronson-type estimates in general, and hence we cannot apply Gobet's approach. In this paper, to show the LAN property for jump-diffusion processes, we instead employ the idea of Theorem 1 in Jeganathan [18], which uses the L^2 regularity condition. This approach is convenient in the sense that it does not require Aronson-type estimates for transition density functions. Though the original results in [18] cannot be applied to triangular array observations, Theorem 2.1 in Fukasawa and Ogihara [10] extends this result to triangular array observations, including high-frequency observations of stochastic processes. Fukasawa and Ogihara [10] showed the LAMN property for degenerate diffusion processes by using this result without Aronson-type estimates. However, because the L^2 regularity conditions in [10,18] are conditions for expectation, it is difficult to apply them to jump-diffusion processes whose tail is heavier than that of diffusion processes. To solve this problem, we weaken the L^2 regularity conditions to conditions for conditional expectation that could be applied to heavy-tailed models such as jump-diffusion processes, and we show that the LAMN property still holds under this new scheme (Theorem 3.1).

However, there remains another serious obstacle to showing the LAN property of jump-diffusion processes: The transition density functions of jump-diffusion processes are given by a mixture of different density functions depending on the jump numbers. In particular, the asymptotic behavior of the density with no jump is quite different to that with jumps, as indicated by Kohatsu-Higa, Nualart, and Tran [20]. They solved this problem by using Malliavin calculus for Wiener–Poisson space and stochastic flows and obtained an expression for the transition density functions that contains the derivative of jump-diffusion processes with respect to drift parameters. However, when the jump structure is also parametrized as in our case, not only the jump coefficient but also the associated Poisson random measure may possibly be parametrized in some way, and this makes it difficult to obtain the derivative of jump-diffusion processes with respect to jump parameters that appear in the formal expression of the transition density functions. For this reason, it is difficult to evaluate the transition density function, which is important for checking the assumptions of Theorem 3.1, in the same way as Kohatsu-Higa, Nualart, and Tran [20]. To deal with this problem, we consider the approximation of transition density functions by thresholding techniques used by Shimizu and Yoshida [30] and Ogihara and Yoshida [27] to construct quasi-maximum-likelihood estimators and Bayes-type estimators. The thresholding techniques are also used for detecting jumps in processes with jumps (see also [4,11,21,22]) and improving the estimation accuracy of continuous components. However, it is not clear whether the LAN property for the statistical model generated by approximated density functions implies the LAN property for the original model. We also provide general sufficient conditions for the property (Theorem 4.3). This result is expected to be applicable not only to jump-diffusion models but also to various models with tractable approximated transition density functions. See Remark 3.2 for the case of nonsynchronously observed diffusion processes.

From the conditional-type L^2 regularity condition and density approximation technique, we derive the LAN property of jump-diffusion processes (Theorem 2.3). The result shows that the optimal rate of convergence of the diffusion parameter is $\sqrt{n}$ and that of the drift and jump parameters is $\sqrt{nh_n}$, where n and h_n are the sample size and sampling step size, respectively, with $nh_n \to \infty$. The threshold-based estimators by Shimizu and Yoshida [30] and Ogihara and Yoshida [27] are typical estimators achieving the optimal rate. Meanwhile, the choice of threshold is important in practical finite-sample data. Shimizu [29] and Inatsugu and Yoshida [17] studied how to choose thresholds, with Shimizu [29] verifying the estimation accuracy of the estimator of Shimizu and Yoshida [30] by numerical simulation with different threshold levels. Also, Gloter, Loukianova, and Mai [11] gave an efficient estimator of the drift parameter, weakening the condition $nh_n^2 \to 0$, which was conventionally assumed for models of ergodic diffusion processes and ergodic jump-diffusion processes, to $nh_n^{3-\epsilon} \to 0$. The finite-sample performance of the estimator is presented in the supplement of [11].

The rest of this paper is organized as follows. In Section 2, we give the LAN property for discrete observations of jump-diffusion processes as the main results, and in Section 3 we state an extended result of Theorem 1 in [18]. In Section 4, we use the results in Section 3 to construct a new scheme for the LAMN property via transition density approximation. Finally, we give proofs of the LAN property of jump-diffusion processes in Section 5.

2. Main results

In this section, we introduce the LAN property of jump-diffusion processes. We start with the definition of the LAMN property.

Let $\mathbb{N}$ be the set of all positive integers. Let $\alpha_0 \in \Theta$ and $\{P_{\alpha,n}\}_{\alpha\in\Theta}$ be a family of probability measures defined on a measurable space $(\mathcal{X}_n, \mathcal{A}_n)$ for $n \in \mathbb{N}$, where Θ is an open subset of $\mathbb{R}^d$. We first consider the following slightly weaker condition than the LAMN property. We denote by $\|\cdot\|_{\mathrm{op}}$ the operator norm, by I_l the unit matrix of size $l \in \mathbb{N}$, and by $\top$ the transpose operator for a matrix or a vector.

Condition (L). The following two conditions are satisfied for $\{P_{\alpha,n}\}_{\alpha\in\Theta,n\in\mathbb{N}}$.

1. There exist a sequence $\{\epsilon_n\}_{n\in\mathbb{N}}$ of nondegenerate matrices (not necessarily symmetric), a sequence $\{V_n(\alpha_0)\}$ of $\mathcal{A}_n$-measurable d-dimensional vectors, and a sequence $\{\mathcal{T}_n(\alpha_0)\}$ of $\mathcal{A}_n$-measurable $d \times d$ symmetric matrices such that $\|\epsilon_n\|_{\mathrm{op}} \to 0$ as $n \to \infty$,

$$P_{\alpha_0,n}(\mathcal{T}_n(\alpha_0) \text{ is nonnegative definite}) = 1 \tag{2.1}$$

for any $n \in \mathbb{N}$, and

$$\log \frac{dP_{\alpha_0+\epsilon_n h,n}}{dP_{\alpha_0,n}} - h^\top V_n(\alpha_0) + \frac{1}{2} h^\top \mathcal{T}_n(\alpha_0) h \to 0 \tag{2.2}$$

as $n \to \infty$ in $P_{\alpha_0,n}$-probability for any $h \in \mathbb{R}^d$.
2. There exists an almost surely symmetric, nonnegative definite $d \times d$ random matrix $\mathcal{T}(\alpha_0)$ such that

$$\mathcal{L}(V_n(\alpha_0), \mathcal{T}_n(\alpha_0) | P_{\alpha_0,n}) \to \mathcal{L}(\mathcal{T}^{1/2}(\alpha_0) W, \mathcal{T}(\alpha_0)),$$

where W is a d-dimensional standard normal random variable independent of $\mathcal{T}(\alpha_0)$.

The following definition of the LAMN property is Definition 1 in [18].

Definition 2.1. The sequence of the families $\{P_{\alpha,n}\}_{\alpha\in\Theta,n\in\mathbb{N}}$ satisfies the LAMN condition at $\alpha=\alpha_0\in\Theta$ if Condition (L) is satisfied, ϵ_n is a symmetric, positive definite matrix and

$$P_{\alpha_0,n}(\mathcal{T}_n(\alpha_0) \text{ is positive definite}) = 1$$

for any $n\in\mathbb{N}$, and $\mathcal{T}(\alpha_0)$ is positive definite almost surely.

We say that $\{P_{\alpha,n}\}_{\alpha\in\Theta,n\in\mathbb{N}}$ satisfies LAN if the LAMN condition is satisfied with a nonrandom matrix $\mathcal{T}(\alpha_0)$.

For proving the LAMN property for diffusion processes by using a localization technique such as Lemma 4.1 in Gobet [12], Condition (L) is useful because (L) for the localized model often implies (L) for the original model. See, for example, the proofs of Theorems 2.4 and 2.5 in [10].

Now, we consider a statistical model of discretely observed jump-diffusion processes. Let $\Theta_i\subset\mathbb{R}^{d_i}$ be an open set for $i\in\{1,2\}$, and $\Theta=\Theta_1\times\Theta_2$. Let $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\geq 0},P)$ be a stochastic basis. We set $\alpha=(\sigma,\theta)\in\Theta_1\times\Theta_2$ and its true value is denoted by $\alpha_0=(\sigma_0,\theta_0)$. For any $\alpha\in\Theta$, let $X^\alpha=(X_t^\alpha)_{t\geq 0}$ be an m-dimensional càdlàg $\mathbf{F}$-adapted process satisfying a stochastic differential equation:

$$dX_t^\alpha = a(X_t^\alpha,\theta)dt + b(X_t^\alpha,\sigma)dW_t + \int_E zN_\theta(dt,dz), \tag{2.3}$$

where $\mathbf{F}=(\mathcal{F}_t)_{t\geq 0}$, $E=\mathbb{R}^m\backslash\{0\}$, $W=(W_t)_{t\geq 0}$ is an m-dimensional standard $\mathbf{F}$-Wiener process, and N_θ is a Poisson random measure on $\mathbb{R}_+\times E$ relative to $\mathbf{F}$, whose mean measure is $f_\theta(z)dzdt$ with $\int_E f_\theta(z)dz<\infty$. The coefficients $a:\mathbb{R}^m\times\Theta_2\mapsto\mathbb{R}^m$ and $b:\mathbb{R}^m\times\Theta_1\mapsto\mathbb{R}^m\otimes\mathbb{R}^m$ are measurable functions and satisfy Assumption (H1) below. We assume that the distribution of X_0^α does not depend on $\alpha\in\Theta$. We denote $X_t=X_t^{\alpha_0}$ and $N(dt,dz)=N_{\theta_0}(dt,dz)$. We suppose that we observe high-frequency data $\{X_{kh_n}\}_{k=0}^n$ from the solution process $X=(X_t)_{t\geq 0}$. Here, $\{h_n\}_{n\in\mathbb{N}}$ is a positive sequence with $h_n\to 0$ and $nh_n\to\infty$. We often write $t_k=kh_n$ below. For matrices $(M_i)_{i=1}^l$, let

$$\operatorname{diag}((M_i)_{i=1}^l) = \begin{pmatrix} M_1 & O & O \\ O & \ddots & O \\ O & O & M_l \end{pmatrix}.$$

Assumption (H1). The derivatives $\partial_x^i\partial_\theta^j a(x,\theta)$ and $\partial_x^i\partial_\sigma^j b(x,\sigma)$ exist and are continuous on $\mathbb{R}^m\times\Theta_2$ and $\mathbb{R}^m\times\Theta_1$, respectively, for $i,j\in\{0,1,2,3,4\}$ such that $i+j\leq 4$. Moreover, there exist positive constants C_1 and κ such that

$$|a(x,\theta)|\leq C_1(1+|x|), \quad |\partial_x a(x,\theta)|+|b(x,\sigma)|+|\partial_x b(x,\sigma)|\leq C_1,$$

$$|\partial_x^i\partial_\theta^j a(x,\theta)|+|\partial_x^i\partial_\sigma^j b(x,\sigma)|\leq C_1(1+|x|)^\kappa$$

for all $i,j\in\{0,1,2,3,4\}$ satisfying $i+j\leq 4$, and $(\theta,\theta',\sigma,\sigma',x)$.

Assumption (H2). $b(x,\sigma)$ is symmetric, positive definite, and there exists a positive constant C_2 such that

$$C_2^{-1}I_m\leq b(x,\sigma)\leq C_2 I_m$$

for any x and σ.

Assumption (H3). X is ergodic; that is, there exists a stationary distribution π such that

$$\frac{1}{T}\int_0^T g(X_t)dt \xrightarrow{P} \int g(x)d\pi(x) \tag{2.4}$$

as $T \to \infty$ for any π-integrable function g. Moreover,

$$\sup_{\alpha\in\Theta, t\geq 0} E[|X_t^\alpha|^q] < \infty \tag{2.5}$$

for $q > 0$.

Let F_θ be a density function satisfying $f_\theta = \lambda F_\theta$ with some positive constant $\lambda = \lambda(\theta)$. Hereinafter, we write the support of any function or measure g and its boundary in E as $\mathrm{supp}(g)$ and $\partial\,\mathrm{supp}(g)$, respectively. Let $d(z, A) = \inf_{y\in A}|z-y|$ for $z \in \mathbb{R}^m$ and $A \subset \mathbb{R}^m$ ($d(z,\varnothing) = \infty$ by convention).

Assumption (H4). 1. The zero points of F_θ do not depend on θ.
2. The derivative $\partial_\theta^l \lambda$ exists and is bounded for $l \in \{0,1,2,3\}$.
3. There exist constants $\epsilon > 0$, $\rho \in (0, 1/2)$, and $N_0 \in \mathbb{N}$ satisfying

$$\int_{\{z:d(z,\partial\,\mathrm{supp}(F_\theta))\leq h_n^\rho\}} F_\theta(z)dz \leq h_n^\epsilon \tag{2.6}$$

for all $n \geq N_0$.
4. The derivative $\partial_\theta^l f_\theta(z)$ exists and is continuous with respect to $\theta \in \Theta_2$ for any $l \in \{0,1,2,3\}$ and $z \in E$. Moreover, there exist constants $\gamma \geq 0$, $C_3 > 0$, and $\epsilon' > 0$ such that

$$|F_\theta(z)|1_{\{|z|\leq\epsilon'\}} \leq C_3|z|^\gamma, \quad |\partial_\theta^l \log f_\theta(z)|1_{\{F_\theta(z)\neq 0\}} \leq C_3(1+|z|)^{C_3}, \tag{2.7}$$

$$|\partial_\theta^l \log f_\theta(z_1) - \partial_\theta^l \log f_\theta(z_2)|1_{\{F_\theta(z_1)F_\theta(z_2)\neq 0\}} \leq C_3|z_1 - z_2|(1+|z_1|+|z_2|)^{C_3} \tag{2.8}$$

for any $z, z_1, z_2 \in E$, $\theta \in \Theta_2$, and $l \in \{1,2,3\}$.
5. $\sup_\theta \int |z|^p f_\theta(z)dz < \infty$ for any $p \geq 1$, and there exists $\eta > 0$ such that

$$n^{1+\eta} h_n^{1+((m+\gamma)/2)\wedge 1} \to 0 \tag{2.9}$$

as $n \to \infty$.

Let $\Gamma = \mathrm{diag}((\Gamma_1, \Gamma_2))$, where $S(x,\sigma) = b^2(x,\sigma)$,

$$[\Gamma_1]_{ij} = \frac{1}{2}\int \mathrm{tr}(\partial_{\sigma_i} S S^{-1} \partial_{\sigma_j} S S^{-1})(x, \sigma_0)d\pi(x),$$

$$[\Gamma_2]_{ij} = \int (\partial_{\theta_i} a)^\top S^{-1}(\partial_{\theta_j} a)(x, \alpha_0)d\pi(x) + \int_E \frac{\partial_{\theta_i} f_{\theta_0} \partial_{\theta_j} f_{\theta_0}}{f_{\theta_0}} 1_{\{f_{\theta_0}\neq 0\}}(y)dy.$$

Assumption (H5). Γ is positive definite.

Regarding our technical assumptions, we make some comments below.

- Under (H1), the existence and uniqueness of the solution and its Markov property are ensured (for details, see Applebaum [2]). (H1) is also important for considering the derivatives of the flow and Malliavin calculus on the continuous part of (2.3).

- For sufficient conditions of ergodicity (2.4), we refer the reader to Masuda [24]. We need the moment condition (2.5) of X_t^α uniformly in α. This condition is somewhat stronger than the one usually assumed in studies of statistical estimation for jump-diffusion processes (estimate for only $\alpha = \alpha_0$) because evaluation of transition densities around α_0 is essential for the LAN property. However, this condition can be shown similarly to a standard procedure. See Theorem 2.2 in Masuda [23] for the details.
- Because we cannot observe fluctuation of jumps directly, we replace it by the increments of X exceeding a threshold on the estimation of θ. However, these increments may not belong to the support of F_θ typically for one-sided jumps or bounded jumps. (2.6) is useful for controlling such a ($\mathcal{F}_{t_{j-1}}$-conditional) probability; for more details, see Lemma 5.3 in the appendix.
- Suppose that $\sup_\theta \int_E F_\theta^p(z)dz < \infty$ for some $p > 1$ and $\partial\,\mathrm{supp}(F_\theta) = \{z_1, \cdots, z_k\}$ for some $k \in \mathbb{N}$, and $z_j \in E$ $(1 \le j \le k)$. Then the set $\{z; d(z, \partial\,\mathrm{supp}(F_\theta)) \le h_n^\rho\}$ is included in the union of k closed balls centered at $z_1, \cdots, z_k$ of radius h_n^ρ, and hence Hölder's inequality yields

$$\int_{\{z;d(z,\partial\,\mathrm{supp}(F_\theta))\le h_n^\rho\}} F_\theta(z)dz \le \left(\int_E F_\theta^p(z)dz\right)^{1/p}\left(\int_{\{z;d(z,\partial\,\mathrm{supp}(F_\theta))\le h_n^\rho\}} dz\right)^{1/q}$$
$$\le Ck^{1/q} h_n^{\rho m/q},$$

  where $q = p/(p-1)$. Then (2.6) is satisfied for sufficiently large n.
- When $m \ge 2$, (2.9) becomes $n^{1+\eta} h_n^2 \to 0$ for some $\eta > 0$, which is almost the same as the one usually required in the study of statistical estimation for jump-diffusion processes. We can say the same thing when $m = 1$ and $\gamma \ge 1$. This condition is weaker than the corresponding condition in Shimizu and Yoshida [30] ($\gamma > 3$ is required). We can also consider the case $m = 1$ and $\gamma \in [0, 1)$. In this case, the convergence rate of h_n becomes restrictive ($n^{2+\eta} h_n^3 \to 0$ for some $\eta > 0$ in the worst case). These things happen because we need to detect jumps by using the increment $|X_{kh_n} - X_{(k-1)h_n}|$. Roughly speaking, for $\rho \in (0, 1/2)$, we have $|X_{kh_n} - X_{(k-1)h_n}| \le h_n^\rho$ with high probability if there are no jumps in $((k-1)h_n, kh_n]$. Then we judge that jumps occur when $|X_{kh_n} - X_{(k-1)h_n}| > h_n^\rho$. If the dimension of X_t is large or γ is large, then the probability that the absolute jump size is equal to or less than h_n^ρ becomes very small, and consequently jump detection and approximation by thresholding densities work well. Otherwise, we must set h_n to be small in order to detect jumps.
- For $u^1 \in \mathbb{R}^{d_1}$ and $u^2 \in \mathbb{R}^{d_2}$, we have

$$(u^1)^\top \Gamma_1 u^1 = \frac{1}{2}\int \mathrm{tr}\left(\left(\sum_j [u^1]_j \partial_{\sigma_j} SS^{-1}\right)^2\right)(x, \sigma_0)d\pi(x),$$

  and

$$(u^2)^\top \Gamma_2 u^2 = \int \left|S^{-1/2}\left(\sum_j [u^2]_j \partial_{\theta_j} a\right)\right|^2 (x, \alpha_0)d\pi(x) + \int_E \left(\sum_j [u^2]_j \partial_{\theta_j} f_{\theta_0}\right)^2 \frac{1_{\{f_{\theta_0}\neq 0\}}}{f_{\theta_0}}(y)dy.$$

  Therefore, (H5) is satisfied if both of the following two conditions are satisfied.

  1. For $u^1 \in \mathbb{R}^{d_1}$, $\sum_j [u^1]_j \partial_{\sigma_j} S(x, \sigma_0) = 0$ for any $x \in \mathrm{supp}(\pi)$ implies $u^1 = 0$.
  2. For $u^2 \in \mathbb{R}^{d_2}$, $\sum_j [u^2]_j \partial_{\theta_j} a(x, \sigma_0) = 0$ for any $x \in \mathrm{supp}(\pi)$ and $\sum_j [u^2]_j \partial_{\theta_j} f_{\theta_0}(y) = 0$ for any $y \in E \cap \{y; f_{\theta_0}(y) \neq 0\}$ imply $u^2 = 0$.

Example 2.2. Condition (H4) is somewhat complicated, so we show some examples of F_θ that satisfy (H4). Let λ be a smooth function of θ satisfying $\sup_\theta |\partial_\theta^l \lambda| < \infty$ for $l \in \{0,1,2,3\}$ and $\inf_\theta \lambda > 0$.

1. (Normal distribution) Let

$$F_\theta(z) = \frac{1}{(2\pi \det \Sigma)^{m/2}} \exp\left(-\frac{1}{2}(z-\mu)^\top \Sigma^{-1}(z-\mu)\right),$$

where μ and Σ are smooth $\mathbb{R}^m$- and $\mathbb{R}^m \otimes \mathbb{R}^m$-valued functions of θ, respectively, such that $\sup_\theta(|\partial_\theta^l \mu| \vee \|\partial_\theta^l \Sigma\|_{\mathrm{op}}) < \infty$ for $l \in \{0,1,2,3\}$ and $\sup_\theta \|\Sigma^{-1}\|_{\mathrm{op}} < \infty$. Then we can easily check (2.6)–(2.8). Therefore, (H4) is satisfied if there exists $\eta > 0$ such that

$$\begin{cases} n^{2+\eta} h_n^3 \to 0, & \text{if } m = 1, \\ n^{1+\eta} h_n^2 \to 0, & \text{if } m \geq 2. \end{cases} \tag{2.10}$$

2. (Gamma distribution) Let $m = 1$ and

$$F_\theta(z) = \frac{1}{\Gamma(\alpha)\beta^\alpha} z^{\alpha-1} e^{-z/\beta} 1_{\{z>0\}},$$

where α and β are smooth $\mathbb{R}$-valued functions of θ such that $\sup_\theta(|\partial_\theta^l \alpha| \vee |\partial_\theta^l \beta|) < \infty$ for $l \in \{0,1,2,3\}$, $\inf_\theta \beta > 0$, and $\inf_\theta \alpha \geq 1$. Then $\partial\, \mathrm{supp}(F_\theta) = \emptyset$, which implies (2.6). Moreover, we have

$$\log f_\theta(z) 1_{\{z>0\}} = \left\{ (\alpha - 1)\log z - \frac{z}{\beta} - \alpha \log \beta - \log \Gamma(\alpha) + \log \lambda \right\} 1_{\{z>0\}}.$$

If $\partial_\theta \alpha \equiv 0$ (for example, the case of exponential distributions $\alpha \equiv 1$), then $\log f_\theta$ satisfies (2.7) and (2.8), and (H4) holds if there exists $\eta > 0$ such that $n^{1+\eta} h_n^{1+(\alpha/2)\wedge 1} \to 0$ as $n \to \infty$.

If $\partial_\theta \alpha \neq 0$ for some θ', then (2.7) is not satisfied because $\lim_{z\searrow 0} |\partial_\theta \log f_{\theta'}(z)| \to \infty$, and hence (H4) does not hold.

3. (Two-sided Gamma distribution) Let $m = 1$ and

$$F_\theta(z) = \frac{1}{2\Gamma(\alpha_1)\beta_1^{\alpha_1}} z^{\alpha_1 - 1} e^{-z/\beta_1} 1_{\{z>0\}} + \frac{1}{2\Gamma(\alpha_2)\beta_2^{\alpha_2}} (-z)^{\alpha_2 - 1} e^{z/\beta_2} 1_{\{z<0\}},$$

where $\alpha_1, \alpha_2, \beta_1$, and β_2 are smooth $\mathbb{R}$-valued functions of θ such that $\sup_\theta(|\partial_\theta^l \alpha_j| \vee |\partial_\theta^l \beta_j|) < \infty$, $\inf_\theta \beta_j > 0$, and $\inf_\theta \alpha_j \geq 1$ for $l \in \{0,1,2,3\}$ and $j \in \{1,2\}$.

Similarly to the above example, (H4) holds if $\partial_\theta \alpha_j \equiv 0$ for $j \in \{1,2\}$ and there exists $\eta > 0$ such that $n^{1+\eta} h_n^{1+(\alpha_1 \wedge \alpha_2 \wedge 2)/2} \to 0$. (H4) is not satisfied if $\partial_\theta \alpha_j \neq 0$ for some $j \in \{1,2\}$ and θ.

4. (Inverse Gaussian distribution) Let $m = 1$ and

$$F_\theta(z) = \sqrt{\frac{\delta}{2\pi z^3}} \exp\left(-\frac{\delta(z-\mu)^2}{2\mu^2 z} \right) 1_{\{z>0\}},$$

where δ and μ are smooth $\mathbb{R}$-valued functions of θ such that $\sup_\theta(|\partial_\theta^l \delta| \wedge |\partial_\theta^l \mu|) < \infty$, $\inf_\theta \delta > 0$, and $\inf_\theta \mu > 0$ for $l \in \{0,1,2,3\}$. From $\partial\, \mathrm{supp}(F_\theta) = \emptyset$, we have (2.6). Because $\log f_\theta(z) 1_{\{z>0\}}$ is expressed as

$$\log f_\theta(z) 1_{\{z>0\}} = \frac{1}{2}\left\{ \log \delta - \log(2\pi) - 3\log z - \delta\left(\frac{z}{\mu^2} - \frac{2}{\mu} + \frac{1}{z} \right) \right\} 1_{\{z>0\}},$$

(H4) holds if $\partial_\theta \delta \equiv 0$ and there exists $\eta > 0$ such that $n^{1+\eta} h^2 \to 0$. If $\partial_\theta \delta \neq 0$ for some θ, then (H4) does not hold because $\lim_{z \searrow 0} |\partial_\theta \log f_{\theta'}(z)| \to \infty$. We can also consider this two-sided version.

As the other examples, similarly to the first example of the normal distribution, if F_θ has no zero points, bounded and away from zero near the origin uniformly in θ, then (H4) holds under (2.10) in addition to somewhat standard conditions (the second inequality of) (2.7), (2.8) and the inequality $\sup_\theta \int |z|^p f_\theta(z)dz < \infty$ for any $p > 1$.

Let $\{P_{\alpha,n}\}_{\alpha,n}$ be the family of probability measures generated by $(X^\alpha_{kh_n})_{k=0}^n$.

Theorem 2.3. *Assume (H1)–(H5). Then, $\{P_{\alpha,n}\}_{\alpha,n}$ satisfies LAN at $\alpha = \alpha_0$ with $\mathcal{T}(\alpha_0) = \Gamma$ and $\epsilon_n = \mathrm{diag}(n^{-1/2} I_{d_1}, (nh_n)^{-1/2} I_{d_2})$.*

Remark 2.4. Theorem 9.1 in Chapter II of Ibragimov and Has'minskiĭ [16] yields the convolution theorem for this model. We can see that Γ^{-1} coincides with the asymptotic variances of the quasi-maximum-likelihood estimator $\hat{\alpha}_n = (\hat{\sigma}_n, \hat{\theta}_n)$ and the Bayes-type estimator $\tilde{\alpha}_n = (\tilde{\sigma}_n, \tilde{\theta}_n)$ in Shimizu and Yoshida [30] and Ogihara and Yoshida [27], respectively. Then we can show that these estimators are asymptotically efficient in the sense of the convolution theorem. The finite-sample performance of the former estimator is presented for example in [29] with tuning methods of the threshold.

Remark 2.5. Because $\hat{\alpha}_n$ is asymptotically efficient and the asymptotic covariance of $\hat{\sigma}_n$ and $\hat{\theta}_n$ is equal to zero, Theorem 2.3 allows us to construct Wald-type tests testing $H : \sigma = \sigma_0$ (resp. $\theta = \theta_0$) against $K : \sigma \neq \sigma_0$ (resp. $\theta \neq \theta_0$). These tests are asymptotically uniformly most powerful in the sense of Sections 4 and 5 in Choi, Hall, and Schick [5] (see Section 7 and Theorems 2 and 3 in [5]) (although the scaling matrix ϵ_n is assumed to be $I_d/\sqrt{n}$ in [5], their proofs remain valid for our setting).

Remark 2.6. We can generalize Theorem 2.3 when the jump part in (2.3) is given by $\int_E c(X^\alpha_{t-}, z, \theta) \times N_\theta(dt, dz)$ under similar conditions to [H6], [H7], and [G1] in Ogihara and Yoshida [27] by introducing the function $\Psi_\theta(y, x)$ in Shimizu and Yoshida [30] and Ogihara and Yoshida [27]. However, we adopt $c(x, z, \theta) = z$ in our setting to avoid excessive complexity.

In the rest of this section, we construct the thresholding quasi-likelihood function that is essential for showing Theorem 2.3. Define $(X^{\alpha,c}_t)_{t\geq 0} = (X^{\alpha,c}_{t,x})_{t\geq 0}$ by the solution of the following stochastic differential equations: $X^{\alpha,c}_0 = x$ and

$$dX^{\alpha,c}_t = a(X^{\alpha,c}_t, \theta)dt + b(X^{\alpha,c}_t, \sigma)dW_t.$$

Hereinafter, the transition density function of $(X^{\alpha,c}_t)_{t\geq 0}$ is written as $p^{c,t-s}_{x,\alpha}(y)$. Obviously, $(X^{\alpha,c}_t)_{t\geq 0}$ corresponds to the continuous part of X, and the theoretical results for its flow and $p^{c,t-s}_{x,\alpha}(y)$ are presented in the appendix. By conditioning the number of jumps on the interval $(t_{j-1}, t_j]$, $p_j(x_{j-1}, x_j, \alpha)$, the density function of $P(X^\alpha_{t_j} \in \cdot | X^\alpha_{t_{j-1}} = x_{j-1})$ can be decomposed as

$$p_j(x_{j-1}, x_j, \alpha) = p^0_j(x_{j-1}, x_j, \alpha) + p^1_j(x_{j-1}, x_j, \alpha) + \sum_{l=2}^{\infty} p^2_{l,j}(x_{j-1}, x_j, \alpha), \tag{2.11}$$

where

$$p_j^0(x_{j-1},x_j,\alpha)=e^{-\lambda h_n}p_{x_{j-1},\alpha}^{c,t_j-t_{j-1}}(x_j),$$

$$p_j^1(x_{j-1},x_j,\alpha)=\lambda e^{-\lambda h_n}\int_{t_{j-1}}^{t_j}\int\int p_{x_{j-1},\alpha}^{c,\tau-t_{j-1}}(x)F_\theta(y)p_{x+y,\alpha}^{c,t_j-\tau}(x_j)dxdyd\tau,$$

$$p_{l,j}^2(x_{j-1},x_j,\alpha)=\frac{\lambda^l e^{-\lambda h_n}}{l!}\int_{t_{j-1}}^{t_j}\cdots\int_{t_{j-1}}^{t_j}\int\cdots\int p_{x_{j-1},\alpha}^{c,\tilde{\tau}_1-t_{j-1}}(z_1)F_\theta(z_2)\cdots p_{z_{2l-1}+z_{2l},\alpha}^{c,t_j-\tilde{\tau}_l}(x_j)$$
$$\times\left(\prod_{j=1}^{2l}dz_j\right)\left(\prod_{k=1}^{l}d\tau_k\right).$$

Here, $\tau_1,\cdots,\tau_l$ are the jump times, and $\tilde{\tau}_1,\cdots,\tilde{\tau}_l$ are $\tau_1,\cdots,\tau_l$ sorted in ascending order.

From (H4), there exist $\rho\in(1/4,1/2)$ and $\eta'>0$ such that $n^{1+\eta'}h_n^{1+(m+\gamma)\rho}\to 0$. We write $L_n=\{x\in\mathbb{R}^m||x|\le h_n^\rho\}$. For $\rho\in(0,1/2)$, Shimizu and Yoshida [30] constructed a thresholding quasi-likelihood function based on the-jump-detection rule $|X_{t_j}-X_{t_{j-1}}|>h_n^\rho$ or not, and here we do the same. More specifically, we define the thresholding quasi-likelihood function by

$$\tilde{p}_j(\alpha)=\tilde{p}_j(\alpha,x_j,x_{j-1})=p_j^0(x_{j-1},x_j,\alpha)1_{L_n}(\Delta x_j)+p_j^1(x_{j-1},x_j,\alpha)1_{L_n^c}(\Delta x_j),\tag{2.12}$$

where $\Delta x_j=x_j-x_{j-1}$. The results in Section 4 show that the LAN property of $\{P_{\alpha,n}\}_{\alpha,n}$ at $\alpha=\alpha_0$ follows from that of $\{\tilde{P}_{\alpha,n}\}_{\alpha,n}$ at $\alpha=\alpha_0$ being induced by the thresholding quasi-likelihood function.

3. LAMN property via a conditional L^2 regularity condition

To show Theorem 2.3, we extend Theorem 1 of Jeganathan [18] and Theorem 2.1 of Fukasawa and Ogihara [10], who studied sufficient conditions for the LAMN property. For our purpose, it is sufficient to develop a theory for LAN models, but we consider general LAMN models because the results in this and the following section can be developed for LAMN models without complexity.

We provide sufficient conditions for the LAMN property that are more useful than those in Theorem 1 of [18] and Theorem 2.1 of [10] for dealing with jump-diffusion processes. Some of the assumptions in Theorem 1 of [18] and Theorem 2.1 of [10] are written with respect to expectations, whereas our new conditions are based on conditional expectations, which are convenient for heavy-tailed noise. There are many statistical models of stochastic processes that satisfy not the LAN property but the LAMN property. The results in this section are expected to be useful for showing the LAMN property not only for jump-diffusion models but also for other models of stochastic processes.

Let $(m_n)_{n=1}^\infty$ be a sequence of positive integers. For any n, let $\{\mathcal{X}_{n,j}\}_{j=1}^{m_n}$ be a sequence of complete, separable metric spaces. Let $\mathcal{X}_n=\mathcal{X}_{n,1}\times\cdots\times\mathcal{X}_{n,m_n}$ and $\mathcal{A}_n=\mathcal{B}(\mathcal{X}_n)$, where $\mathcal{B}(\mathcal{X}_n)$ denotes the Borel σ-algebra of $\mathcal{X}_n$. We consider statistical experiments $(\mathcal{X}_n,\mathcal{B}(\mathcal{X}_n),\{P_{\alpha,n}\}_{\alpha\in\Theta})$. Let $X_j=X_{n,j}:\mathcal{X}_n\to\mathcal{X}_{n,j}$ be the natural projection, $\bar{X}_j=\bar{X}_{n,j}=(X_1,\cdots,X_j)$, $\bar{\mathcal{X}}_{n,j}=\mathcal{X}_{n,1}\times\cdots\times\mathcal{X}_{n,j}$, $\mathcal{A}_{0,n}=\{\emptyset,\mathcal{X}_n\}$, and $\mathcal{A}_{j,n}$ is the minimal sub σ-algebra of $\mathcal{A}_n$ for which $\bar{X}_j$ is $\mathcal{A}_{j,n}$-measurable for $1\le j\le m_n$. Suppose that there exists a σ-finite measure $\mu_{n,j}$ on $\mathcal{X}_{n,j}$ such that $P_{\alpha,n}(X_1\in\cdot)\ll\mu_{n,1}$ and $P_{\alpha,n}(X_j\in\cdot|\bar{X}_{j-1}=\bar{x}_{j-1})\ll\mu_{n,j}$ for $2\le j\le m_n$ and $\bar{x}_{j-1}\in\bar{\mathcal{X}}_{n,j-1}$.

Let $E_\alpha=E_{\alpha,n}$ denote the expectation with respect to $P_{\alpha,n}$, and let $p_j=p_{j,n}$ be conditional density functions defined by

$$p_1(\alpha)=\frac{dP_{\alpha,n}(X_1\in\cdot)}{d\mu_{n,1}},\quad p_j(\alpha)=\frac{dP_{\alpha,n}(X_j\in\cdot\,|\bar{X}_{j-1}=\bar{x}_{j-1})}{d\mu_{n,j}}\quad(2\le j\le m_n).$$

Then, we can see that

$$\int p_j(\alpha)g(\bar{x}_{j-1},x_j)d\mu_{n,j} = E_\alpha[g(\bar{X}_{j-1},X_j)|\bar{X}_{j-1} = \bar{x}_{j-1}] \tag{3.1}$$

almost surely for any bounded Borel function $g : \bar{\mathcal{X}}_{n,j} \to \mathbb{R}$.

Next, we describe our assumptions for the LAMN property. Let ϵ_n be a $d \times d$ nondegenerate matrix, and let $\alpha_h = \alpha_0 + \epsilon_n h$ for $h \in \mathbb{R}^d$.

Assumption (A1). There are random vectors $\dot{\xi}_{nj}(\alpha_0) : \bar{\mathcal{X}}_{n,j} \to \mathbb{R}$ such that for every $h \in \mathbb{R}^d$,

$$\sum_{j=1}^{m_n} \int \left[\xi_{nj}(\alpha_0,h) - \frac{1}{2}h^\top \epsilon_n^\top \dot{\xi}_{nj}(\alpha_0)\right]^2 d\mu_{n,j} \to 0 \tag{3.2}$$

as $n \to \infty$ in $P_{\alpha_0,n}$-probability, where $\xi_{nj}(\alpha_0,h) = \sqrt{p_j(\alpha_h)} - \sqrt{p_j(\alpha_0)}$.

To show the LAMN property, we need to identify the limit distribution of $\log(dP_{\alpha',n}/dP_{\alpha,n})$ under $P_{\alpha,n}$. This involves the log-likelihood ratio of different probability measures, which is difficult to deal with for stochastic processes in general. Gobet [12] dealt with this problem for discretely observed diffusion processes by using estimates from below and above by Gaussian density functions (Aronson estimates) to show the LAMN property. Condition (A1) also involves transition density functions with different values of the parameter. However, if p_j is a positive-valued $C^2(\Theta)$ function, an estimate similar to (2.6) in [10], we can replace the left-hand side of (3.2) with a quantity in which the probability measure of expectation and p_j in the integrand have the same parameter value α_{sh} ($s \in [0,1]$), and therefore we do not need Aronson-type estimates for transition density ratios. Thus, a scheme with the L^2 regularity condition does not require Aronson-type estimates, which is one of the advantages of this scheme. Furthermore, Condition (A1) is the estimate for conditional expectation unlike (A1) in [10]. Therefore, it is much easier to show (A1) compared with (A1) in [10] under the heavy-tailed behavior of jump-diffusion processes.

Define

$$\eta_j(\bar{x}_{j-1},x_j) = \begin{cases} \dot{\xi}_{nj}(\alpha_0)/\sqrt{p_j(\alpha_0)}, & \text{if } p_j(\alpha_0) \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

We use abbreviation η_j both for the random variable $\eta_j(\bar{X}_{j-1},X_j)$ and for the function $\eta_j(\bar{x}_{j-1},x_j)$ when there is no confusion; the same is true for other functions of $(\bar{x}_{j-1},x_j)$. Moreover, let

$$\mathcal{T}_n = \epsilon_n^\top \sum_{j=1}^{m_n} E_{\alpha_0}[\eta_j\eta_j^\top|\mathcal{A}_{j-1,n}]\epsilon_n, \quad \text{and} \quad V_n = \epsilon_n^\top \sum_{j=1}^{m_n} \eta_j. \tag{3.3}$$

We consider the following conditions in addition to (A1).

Assumption (A2). There exists $n_0 \in \mathbb{N}$ such that $E_{\alpha_0}[|\eta_j|^2|\mathcal{A}_{j-1,n}] < \infty$ and $E_{\alpha_0}[\eta_j|\mathcal{A}_{j-1,n}] = 0$, $P_{\alpha_0,n}$-almost surely for every $1 \le j \le m_n$ and $n \ge n_0$.

Assumption (A3). For every $\epsilon > 0$ and $h \in \mathbb{R}^d$,

$$\sum_{j=1}^{m_n} E_{\alpha_0}[|h^\top \epsilon_n^\top \eta_j|^2 1_{\{|h^\top \epsilon_n^\top \eta_j| > \epsilon\}}|\mathcal{A}_{j-1,n}] \to 0$$

as $n \to \infty$ in $P_{\alpha_0,n}$-probability.

Assumption (A4). There exists a random $d \times d$ symmetric matrix $\mathcal{T}$ such that

$$P(\mathcal{T} \text{ is nonnegative definite}) = 1$$

and

$$\mathcal{L}((V_n, \mathcal{T}_n) | P_{\alpha_0,n}) \to \mathcal{L}(\mathcal{T}^{1/2} W, \mathcal{T}),$$

where $W \sim N(0, I_d)$ independent of $\mathcal{T}$.

Conditions (A2)–(A4) are similar to Conditions (A2)–(A5) in [10]. However, (A3) and (A4) in [10] are replaced by estimates for conditional expectations and a tightness property that is trivially satisfied under (A4). It is natural to weaken such conditions of expectations to conditions of conditional expectations and tightness because the LAMN property is the result of convergences in probability and in distribution. Under these assumptions, we can show the following extension of Theorem 1 in [18]. The proof is left to the appendix.

Theorem 3.1. *Assume (A1)–(A4). Then the family $\{P_{\alpha,n}\}_{\alpha,n}$ satisfies Condition (L) with $\mathcal{T}_n$ and V_n in (3.3). If further $\mathcal{T}$ in (A4) is positive definite almost surely and ϵ_n is a symmetric, positive definite matrix for any $n \in \mathbb{N}$, then $\{P_{\alpha,n}\}_{\alpha,n}$ satisfies the LAMN property at $\theta = \theta_0$.*

Remark 3.2. As in Remark 2.1 of [10], if Condition (L) is satisfied, ϵ_n is symmetric and positive definite for any $n \in \mathbb{N}$, and $\mathcal{T}$ is positive definite almost surely, then we can easily show the LAMN property by replacing $\mathcal{T}_n$ with $\dot{\mathcal{T}}_n = \mathcal{T}_n 1_{\{\mathcal{T}_n \text{ is positive definite}\}} + I_d 1_{\{\mathcal{T}_n \text{ is not positive definite}\}}$.

4. LAMN property via transition density approximation

Theorem 3.1 is an important tool for showing Theorem 2.3 because this result requires neither Aronson-type estimates nor an expectation-type L^2 regularity condition. The another important issue in showing Theorem 2.3 is handling the mixture of density functions that behave quite differently depending on the jump numbers. To deal with this issue, we use the thresholding techniques developed by Shimizu and Yoshida [30] and Ogihara and Yoshida [27]. We approximate the transition density functions of jump-diffusion processes with thresholding density functions whose asymptotic behaviors are much easier to deal with. We will show that the LAN property of the original model is proved under some conditions on the approximating density functions.

Let $\tilde{p}_1(\alpha) = \tilde{p}_1(\alpha, x_1)$ and $\tilde{p}_j(\alpha) = \tilde{p}_j(\alpha, x_j, \bar{x}_{j-1})$ be nonnegative-valued functions such that $\tilde{p}_1(\alpha, \cdot)$ is measurable and the mapping $(\bar{x}_{j-1}, A) \mapsto \int_A \tilde{p}_j(\alpha, x_j, \bar{x}_{j-1}) \mu_{n,j}(dx_j)$ is a transition kernel for $2 \le j \le m_n$. We emphasize that $\tilde{p}_1(\alpha, \cdot)$ and $\tilde{p}_j(\alpha, \cdot, \bar{x}_{j-1})$ are not supposed to be probability measures. This is important in the sense that we can consider normalized probability measures on sets that do not contain original rare events. We introduce associated normalizing constants $d_1(\alpha) = \int \tilde{p}_1(\alpha) \mu_{n,1}(dx_1)$ and $d_j(\bar{x}_{j-1}, \alpha) = \int \tilde{p}_j(\alpha) \mu_{n,j}(dx_j)$ for $2 \le j \le m_n$. Assume $d_j(\bar{x}_{j-1}, \alpha)$ is nonzero and finite for any $(\bar{x}_{j-1}, \alpha)$, and let $\tilde{P}_{\alpha,n}$ be a probability measure defined by $\tilde{P}_{\alpha,n} = \prod_{j=1}^{m_n} (\tilde{p}_j(\alpha)/d_j(\alpha))(\bigotimes_{j=1}^{m_n} \mu_{n,j}(dx_j))$, where $\bar{x}_0 = \emptyset$.

Let $K_{n,j} \in \mathcal{A}_{j,n}$ for $1 \le j \le m_n - 1$. Let $D_{j,h}(\bar{x}_{j-1}, t) = d_j(\bar{x}_{j-1}, \alpha_{th})$ for $t \in [0,1]$ and $h \in \mathbb{R}^d$ such that $(\alpha_{th})_{t \in [0,1]} \subset \Theta$. Let

$$\zeta_{j,t}^{l,h} = \frac{(\frac{d}{dt})^l \tilde{p}_j(\alpha_{th})}{\tilde{p}_j(\alpha_{th})} 1_{\{\tilde{p}_j(\alpha_{th}) \neq 0\}}$$

for $l \in \mathbb{N}$ and $h \in \mathbb{R}^d$.

Assumption (B1). For any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$\sup_{\alpha \in \Theta} P_{\alpha,n}(\cup_{j=1}^{m_n-1} K_{n,j}^c) < \epsilon \tag{4.1}$$

and

$$\sup_{\alpha \in \Theta} \tilde{P}_{\alpha,n}(\cup_{j=1}^{m_n-1} K_{n,j}^c) < \epsilon \tag{4.2}$$

for $n \geq N$. Moreover,

$$m_n \max_{1 \leq j \leq m_n} \sup_{\alpha \in \Theta, \bar{x}_{j-1} \in \bar{X}_{j-1}(K_{n,j-1})} \int |p_j(\alpha) - \tilde{p}_j(\alpha)| \mu_{n,j}(dx_j) \to 0 \tag{4.3}$$

as $n \to \infty$.

Here and in the following, we ignore $\bar{x}_0 \in \bar{X}_0(K_{n,0})$ in the range of the supremum. (B1) implies that $\tilde{P}_{\alpha,n}$ approximates $P_{\alpha,n}$ well except on a rare event. A typical example of a rare event is that $x_1, \dots, x_n, \dots$ have large magnitude. On such an event, it is often difficult to evaluate a difference in mass measured by $\tilde{P}_{\alpha,n}$ and $P_{\alpha,n}$. As for an application to jump-diffusion processes, we set

$$K_{n,j} = \left\{ (x_l)_{l=0}^n \subset \mathbb{R}^{m(n+1)} \middle| \max_{0 \leq l \leq j} |x_l| \leq n^\delta \right\}$$

for small enough $\delta > 0$, and this makes it possible to obtain (4.3). For more details, see Section 5.1.

Because the approximation probability measure $\tilde{P}_{\alpha,n}$ contains normalizing constants $d_1, \dots, d_{m_n}$, to check (4.2) may seem cumbersome. Then the following lemma is helpful.

Lemma 4.1. *Assume (4.3), that* $\tilde{p}_j(\alpha, x_j, \bar{x}_{j-1}) \leq p_j(\alpha, x_j, \bar{x}_{j-1})$ $\mu_{n,1} \otimes \cdots \otimes \mu_{n,j}$*-almost everywhere in* $\bar{x}_j$ *for any* α, *and that for any* $\epsilon > 0$, *there exists* $N \in \mathbb{N}$ *such that (4.1) for* $n \geq N$. *Then, for any* $\epsilon > 0$, *there exists* $N' \in \mathbb{N}$ *such that (4.2) for* $n \geq N'$.

The following theorem ensures that the LAMN property of $(\tilde{P}_{\alpha,n})_{\alpha,n}$ implies the LAMN property of $(P_{\alpha,n})_{\alpha,n}$ under (B1). Let $\|P - Q\| = \sup_{|f| \leq 1} |\int f dP - \int f dQ|$ for probability measures P and Q.

Theorem 4.2. *Assume (B1). Then,* $\sup_{\alpha \in \Theta} \|P_{\alpha,n} - \tilde{P}_{\alpha,n}\| \to 0$ *as* $n \to \infty$. *If further, for any* $\epsilon > 0$ *and* $h \in \mathbb{R}^d$, *there exists* $\delta > 0$ *such that*

$$\limsup_{n \to \infty} \tilde{P}_{\alpha_0,n} \left(\frac{d\tilde{P}_{\alpha_h,n}}{d\tilde{P}_{\alpha_0,n}} < \delta \right) < \epsilon, \tag{4.4}$$

then

$$\log \frac{dP_{\alpha_h,n}}{dP_{\alpha_0,n}} - \log \frac{d\tilde{P}_{\alpha_h,n}}{d\tilde{P}_{\alpha_0,n}} \to 0$$

as $n \to \infty$ *in* $P_{\alpha_0,n}$*- and* $\tilde{P}_{\alpha_0,n}$*-probability for any* $h \in \mathbb{R}^d$.

For a vector $x = (x_1, \cdots, x_k)$, we denote $\partial_x^l = (\frac{\partial^l}{\partial x_{i_1} \cdots \partial x_{i_l}})_{i_1, \cdots, i_l = 1}^k$.

Assumption (B2). For any $h \in \mathbb{R}^d$, there exists $N \in \mathbb{N}$ such that $(\frac{d}{dt})^l p_j(\alpha_{th})$, $(\frac{d}{dt})^l \tilde{p}_j(\alpha_{th})$, and $\partial_t^l D_{j,h}$ exist and are continuous for $n \geq N$, $t \in [0,1]$, and $l \in \{0,1,2\}$ almost everywhere in $\bar{x}_{j-1} \in \bar{X}_{j-1}(K_{n,j-1})$. Moreover, there exists $\delta > 0$ such that

$$\sup_{|t|\leq\delta, \bar{x}_{j-1}\in\bar{X}_{j-1}(K_{n,j-1})} \int |\zeta_{j,t}^{l,h}| \tilde{p}_j(\alpha_{th}) \mu_{n,j}(dx_j) < \infty \tag{4.5}$$

and

$$m_n^{1/l} \max_{1\leq j\leq m_n} \sup_{t\in[0,1], \bar{x}_{j-1}\in\bar{X}_{j-1}(K_{n,j-1})} |\partial_t^l D_{j,h}(\bar{x}_{j-1},t)| \to 0 \quad \text{as} \quad n\to\infty \tag{4.6}$$

for $l \in \{1,2\}$ and $n \geq N$.

Lebesgue's dominated convergence theorem yields (4.6) if

$$m_n^{1/l} \max_{1\leq j\leq m_n} \sup_{t\in[0,1], \bar{x}_{j-1}\in\bar{X}_{j-1}(K_{n,j-1})} \int |\partial_\alpha^l p_j(\alpha_{th}) - \partial_\alpha^l \tilde{p}_j(\alpha_{th})| \mu_{n,j}(dx_j) \to 0 \tag{4.7}$$

for $h \in \mathbb{R}^d$, $l \in \{1,2,3\}$, and $n \geq N$ (see (C.8) and (C.9) in the supplementary material for the details [26]). For the setting in Section 2, it is not easy to check (4.7) because of the heavy-tailed behavior. So we directly check (4.6) in Section 5.

Let $(e_i)_{i=1}^d$ be the standard unit vectors in $\mathbb{R}^d$, and let

$$\tilde{\eta}_j(\bar{x}_{j-1}, x_j) = (\zeta_{j,0}^{1,e_1}, \cdots, \zeta_{j,0}^{1,e_d}) 1_{\bar{X}_{j-1}(K_{n,j-1})}(\bar{x}_{j-1}).$$

Let $\tilde{E}_\alpha$ denote the expectation with respect to $\tilde{P}_{\alpha,n}$. Let

$$\tilde{\mathcal{T}}_n = \sum_{j=1}^{m_n} \tilde{E}_{\alpha_0}[\tilde{\eta}_j \tilde{\eta}_j^\top | \mathcal{A}_{j-1,n}] \quad \text{and} \quad \tilde{V}_n = \sum_{j=1}^{m_n} \tilde{\eta}_j. \tag{4.8}$$

We further assume the following conditions.

Assumption (B3). $\tilde{E}_{\alpha_0}[|\zeta_{j,t}^{1,h}|^2 | \mathcal{A}_{j-1,n}] < \infty$ and the zero points of $\tilde{p}_j$ do not depend on $\alpha \in \Theta$ $\tilde{P}_{\alpha_0,n}$-almost surely for $1 \leq j \leq m_n$, and

$$\sum_{j=1}^{m_n} \sup_{t\in[0,1]} \tilde{E}_{\alpha_{th}}[|\zeta_{j,t}^{2,h}|^2 + |\zeta_{j,t}^{1,h}|^4 | \mathcal{A}_{j-1,n}] \to 0$$

as $n \to \infty$ in $\tilde{P}_{\alpha_0,n}$-probability.

Assumption (B4). There exists a random $d \times d$ symmetric matrix $\mathcal{T}$ such that

$$P[\mathcal{T} \text{ is nonnegative definite}] = 1$$

and

$$\mathcal{L}((\tilde{V}_n, \tilde{\mathcal{T}}_n) | \tilde{P}_{\alpha_0,n}) \to \mathcal{L}(\mathcal{T}^{1/2} W, \mathcal{T}),$$

where $W \sim N(0, I_d)$ independent of $\mathcal{T}$.

Conditions (B3) and (B4) are conditions for the asymptotic behavior of functions of $\partial_t \tilde{p}_j$ (not $\partial_t p_j$). This fact is important when we discuss the LAN property of jump-diffusion processess in the following section. While the asymptotic behavior of the transition density functions of jump-diffusion processes is difficult to deal with, that of thresholding density functions is much easier to handle. The next theorem ensures that we need to consider only the latter when we show the LAMN property of the original model; the proof is left to the appendix.

Theorem 4.3. *Assume (B1)–(B4). Then, the family $\{P_{\alpha,n}\}_{\alpha,n}$ of probability measures satisfies Condition (L) with $\tilde{\mathcal{T}}_n$ and $\tilde{V}_n$ in (4.8). If further $\mathcal{T}$ in (B4) is positive definite almost surely and ϵ_n is symmetric and positive definite for any $n \in \mathbb{N}$, then $\{P_{\alpha,n}\}_{\alpha,n}$ satisfies the LAMN property at $\alpha = \alpha_0$.*

If $\mathcal{T}$ is nonrandom (which corresponds to the case of LAN), we can simplify Condition (B4).

Assumption (B4′). There exists a nonrandom $d \times d$ symmetric, nonnegative definite matrix $\mathcal{T}$ such that $\tilde{\mathcal{T}}_n \to \mathcal{T}$ in $\tilde{P}_{\alpha_0,n}$-probability.

Corollary 4.4. *Assume (B1)–(B3) and (B4′). Then $\{P_{\alpha,n}\}_{\alpha,n}$ satisfies Condition (L). If further $\mathcal{T}$ in (B4′) is positive definite and ϵ_n is symmetric and positive definite for any $n \in \mathbb{N}$, then $\{P_{\alpha,n}\}_{\alpha,n}$ satisfies the LAN property at $\alpha = \alpha_0$.*

Remark 4.5. Even when $\mathcal{T}$ is random, Sweeting [31] is useful for not having to check the convergence of $\tilde{V}_n$ in (B4).

Remark 4.6. We expect that such techniques of transition density approximation can be applied to models other than jump-diffusion models, and we expect the following examples of applications.

1. If we can find an approximation of the transition density function of a statistical model such that the asymptotic behavior of the approximation can be specified, then these techniques enable us to show the LAMN property of the statistical model. One such example is the statistical model of nonsynchronously observed diffusion processes in Ogihara [25]. The likelihood function is given by the integral of the likelihood function for synchronized observations with respect to unobserved variables. The LAMN property for this model is shown by introducing the likelihood approximation obtained by cutting off the domain of integration, and identifying the asymptotic behavior of the approximated likelihood function (see Lemma 4.3 and subsequent discussions in [25]). So the techniques in this section enable us to simplify the proof of LAMN for this model.
2. The integrals in the left-hand side of (3.2) are functionals of $\bar{x}_{j-1}$, and the L^2 regularity conditions in [18] or [10] are not applicable if convergence of the expectation of these integrals cannot be shown. We expect this to be the case when we consider estimation of Lévy driven stochastic differential equations or stochastic volatility diffusion models because the functionals of $\bar{x}_{j-1}$ seem to follow heavy-tailed distributions. In such models, the LAMN property is shown by our approach if the likelihood or its approximation satisfies (B1)–(B4).

5. Proof of the LAN property for jump-diffusion processes

In this section, we show the LAN property of jump-diffusion processes based on the scheme proposed in Section 4. We approximate the genuine likelihood by a thresholding likelihood that can roughly distinguish whether the increments contain at least one jump or not. We introduce some conventions used in the rest of this paper.

- For a matrix A and a vector v, we denote element (i,j) of A by $[A]_{ij}$ and element i of v by $[v]_i$. We often regard an r-dimensional vector v as an $r\times 1$ matrix.
- C and C_p denote generic positive constants whose values may vary depending on context.

From now on, we show the LAN property of jump-diffusion processes by applying Corollary 4.4 with the transition density function $p_j(\alpha)$ and the thresholding quasi-likelihood function

$$\tilde{p}_j(\alpha)=\tilde{p}_j(\alpha,x_j,x_{j-1})=p_j^0(x_{j-1},x_j,\alpha)1_{L_n}(\Delta x_j)+p_j^1(x_{j-1},x_j,\alpha)1_{L_n^c}(\Delta x_j).$$

In this setting, $d_j=\int\tilde{p}_j dy\le 1$ and Proposition A.4 in the supplementary material ensure $d_j>0$ under (H1) and (H2). We also remark that we set $p_0(x_0,\alpha)=1$ and $\mu_{n,0}=P(X_0^\alpha\in\cdot)$, and can ignore p_0 when we apply Corollary 4.4 because the distribution of X_0^α does not depend on α by the assumption. Let $\epsilon_n=\mathrm{diag}(n^{-1/2}I_{d_1},(nh_n)^{-1/2}I_{d_2}))$ and $T_n=nh_n$.

5.1. Verifying Conditions (B1) and (B2)

First we observe (B1). For a constant $\delta\in(0,1/4)$, we define

$$K_{n,j}=\left\{(x_l)_{l=0}^n\subset\mathbb{R}^{m(n+1)}\middle|\max_{0\le l\le j}|x_l|\le n^\delta\right\}\quad\text{and}\quad K''_{n,j}=\{x_j\in\mathbb{R}^m||x_j|\le n^\delta\}.$$

For this set, (4.1) follows from the following lemma.

Lemma 5.1. *Assume (H3). Then for any $\epsilon,\delta>0$, there exists a positive integer N such that*

$$\sup_{\alpha\in\Theta}P\left((X_{t_k}^\alpha)_{k=0}^n\in\cup_{j=1}^{n-1}K_{n,j}^c\right)=\sup_{\alpha\in\Theta}P\left[\max_{0\le k\le n}|X_{t_k}^\alpha|>n^\delta\right]<\epsilon$$

for all $n\ge N$.

Proof. Pick a positive constant q fulfilling $q\delta>1$. Then Chebyshev's inequality gives

$$\sup_{\alpha\in\Theta}P\left[\max_k|X_{t_k}^\alpha|>n^\delta\right]\le\frac{1}{n^{q\delta}}\sup_{\alpha\in\Theta}E\left[\max_k|X_{t_k}^\alpha|^q\right]\le n^{1-q\delta}\sup_{t,\alpha}E[|X_t^\alpha|^q]\to 0$$

as $n\to\infty$. □

Next we observe (4.3). Note that by the definition of the threshold quasi-likelihood $\tilde{p}_j(\alpha)$ and Markov property of X, we can replace the supremum over $\bar{X}_{j-1}(K_{n,j-1})$ by the supremum over $K''_{n,j-1}$ in (4.3). Under (H1)–(H4), Proposition A.4 in the supplementary material implies that

$$\begin{aligned}\int|p_j-\tilde{p}_j|dx_j&=\int p_j^0 1_{L_n^c}(\Delta x_j)dx_j+P(N_\theta((t_{j-1},t_j]\times E)\ge 2)+\int p_j^1 1_{L_n}(\Delta x_j)dx_j\\&\le Ch_n^2+1-e^{-\lambda h_n}(1+\lambda h_n)\\&\quad+P(N_\theta((t_{j-1},t_j]\times E)=1\text{ and }|X_{t_j}^\alpha-X_{t_{j-1}}^\alpha|\le h_n^\rho|X_{t_{j-1}}^\alpha=x_{j-1})\end{aligned}$$

for any $x_{j-1}\in K''_{n,j-1}$ and $\alpha\in\Theta$, where C does not depend on x_{j-1}. By applying the triangular inequality, we have

$$|X_{t_j}^\alpha-X_{t_{j-1}}^\alpha|\ge|X_{\tau_j}^\alpha-X_{\tau_j-}^\alpha|-|X_{\tau_j-}^\alpha-X_{t_{j-1}}^\alpha|-|X_{t_j}^\alpha-X_{\tau_j}^\alpha|,$$

where τ_j denotes the first jump time on $(t_{j-1}, t_j]$. Hence, by using (H4), we obtain

$$\begin{aligned} &P(N_\theta((t_{j-1},t_j]\times E)=1 \text{ and } |X^\alpha_{t_j}-X^\alpha_{t_{j-1}}|\le h_n^\rho | X^\alpha_{t_{j-1}}=x_{j-1}) \\ &\quad\le P(N_\theta((t_{j-1},t_j]\times E)=1 \text{ and } |X^\alpha_{\tau_j-}-X^\alpha_{t_{j-1}}|+|X^\alpha_{t_j}-X^\alpha_{\tau_j}|> h_n^\rho | X^\alpha_{t_{j-1}}=x_{j-1}) \\ &\quad+P(N_\theta((t_{j-1},t_j]\times E)=1 \text{ and } |X^\alpha_{\tau_j}-X^\alpha_{\tau_j-}|\le 2h_n^\rho | X^\alpha_{t_{j-1}}=x_{j-1}) \\ &\quad\le Ch_n^2+\lambda h_n e^{-\lambda h_n}\int_{|z|\le 2h_n^\rho} F_\theta(z)dz \\ &\quad\le Ch_n^2+Ch_n^{1+(m+\gamma)\rho}=o(n^{-1}) \end{aligned} \tag{5.1}$$

for $x_{j-1}\in K''_{n,j-1}$. Hence we obtain (4.3) and Lemma 4.1 yields (4.2). Thus (B1) holds.

We next check (B2). For each $l\in\{0,1,2,3\}$, we have

$$\sup_t \int |\partial_t^l \tilde{p}_j(\alpha_{th})|dx_j \le \sup_t \int \{|\partial_t^l p^0_{j,t}|1_{L_n}(\Delta x_j)+|\partial_t^l p^1_{j,t}|1_{L_n^c}(\Delta x_j)\}dx_j,$$

where $p^l_{j,t}=p^l_j(x_{j-1},x_j,\alpha_{th})$. Proposition A.2 and Remark A.3 in the supplementary material lead to

$$\sup_t \int |\partial_t^l p^0_{j,t}|dx_j<\infty.$$

It follows from Proposition A.2 in the supplementary material and (H4) that

$$\begin{aligned} &\int |\partial_t p^1_{j,t}|dx_j \\ &\quad\le \int\int_{t_{j-1}}^{t_j}\int\int \left|\frac{\partial_t p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}}{p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}}(x)+\frac{\partial_t f_{\theta_{th}}}{f_{\theta_{th}}}1_{\{f_{\theta_{th}}\ne 0\}}(y)+\frac{\partial_t p^{c,t_j-\tau}_{x+y,\alpha_{th}}}{p^{c,t_j-\tau}_{x+y,\alpha_{th}}}(x_j)\right| \\ &\qquad\qquad \times p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}} f_{\theta_{th}} p^{c,t_j-\tau}_{x+y,\alpha_{th}}dxdyd\tau dx_j+\frac{Ch_n^2}{\sqrt{T_n}} \\ &\quad\le Cn^{-1/2}h_n(1+|x_{j-1}|)^C+\frac{Ch_n}{\sqrt{T_n}}\int(1+|y|)^C f_{\theta_{th}}(y)dy \\ &\qquad+Cn^{-1/2}\int_{t_{j-1}}^{t_j}\int\int(1+|x+y|)^C p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}} f_{\theta_{th}}dxdyd\tau+\frac{Ch_n^2}{\sqrt{T_n}} \\ &\quad<\infty, \end{aligned} \tag{5.2}$$

where $(\sigma_{th},\theta_{th})=\alpha_{th}$. In a similar manner, we can obtain $\sup_t \int |\partial_t^l p^1_{j,t}|dx_j<\infty$ for $l\in\{0,1,2,3\}$ by Remark A.3 in the supplementary material and (H4), and thus (4.5) holds. $D_{j,h}$ can be decomposed as

$$D_{j,h}=1-e^{-\lambda_{th}h_n}[1+\lambda_{th}h_n]+\int(\tilde{p}_j(\alpha_{th})-p^0_{j,t}-p^1_{j,t})dx_j, \tag{5.3}$$

where $\lambda_{th} = \lambda(\theta_{th})$. Because

$$\left|\partial_t^l(e^{-\lambda_{th}h_n}[1+\lambda_{th}h_n])\right| = \left|\partial_t^{l-1}\left(\lambda_{th}h_n^2 e^{-\lambda_{th}h_n}\frac{\partial_\theta\lambda(\theta_{th})\cdot h}{\sqrt{T_n}}\right)\right| = O(h_n^2)$$

for $l \in \{1,2\}$, Hölder's inequality gives

$$\begin{aligned}
|\partial_t^l D_{j,h}| &\le \left|\partial_t^l(e^{-\lambda_{th}h_n}[1+\lambda_{th}h_n])\right| + \int |\partial_t^l(\tilde{p}_j(\alpha_{th}) - p_{j,t}^0 - p_{j,t}^1)|dx_j \\
&\le Ch_n^2 + \int\left|\frac{\partial_t^l p_{j,t}^0}{p_{j,t}^0}\right| p_{j,t}^0 1_{L_n^c}dx_j + \int\left|\frac{\partial_t^l p_{j,t}^1}{p_{j,t}^1}\right| p_{j,t}^1 1_{L_n}dx_j \\
&\le Ch_n^2 + \left(\int\left|\frac{\partial_t^l p_{j,t}^0}{p_{j,t}^0}\right|^p p_{j,t}^0 dx_j\right)^{1/p}\left(\int 1_{L_n^c}p_{j,t}^0 dx_j\right)^{1/q} \\
&\quad + \left(\int\left|\frac{\partial_t^l p_{j,t}^1}{p_{j,t}^1}\right|^p p_{j,t}^1 dx_j\right)^{1/p}\left(\int 1_{L_n}p_{j,t}^1 dx_j\right)^{1/q}
\end{aligned} \tag{5.4}$$

for $x_{j-1} \in K''_{n,j-1}$, where $p,q > 1$ with $1/p + 1/q = 1$.

Proposition A.2 in the supplementary material, Jensen's inequality, and a similar argument to (5.2) yield

$$\left(\int\left|\frac{\partial_t^l p_{j,t}^k}{p_{j,t}^k}\right|^p p_{j,t}^k dx_j\right)^{1/p} \le C_p(1+|x_{j-1}|)^{C_p}\cdot\frac{1}{\sqrt{nh_n^k}}$$

for $k \in \{0,1\}$. Then as in (5.1), we have

$$\begin{aligned}
|\partial_t^l D_{j,h}| &\le o(n^{-1}) + Cn^{\delta C_p}(nh_n)^{-1/2}h_n^{(1+(m+\gamma)\rho)/q} \\
&\le o(n^{-1}) + Cn^{\epsilon}(nh_n)^{-1/2}h_n^{1+(m+\gamma)\rho} = o(n^{-1})
\end{aligned}$$

for $x_{j-1} \in K''_{n,j-1}$ and q satisfying $(1+(m+\gamma)\rho)/q > 1+(m+\gamma)\rho - \epsilon/2$, where ϵ is a positive constant satisfying $n^{\epsilon+1+(m+\gamma)\rho} = o(1)$ (δ in $K_{n,j}$ should be reset to satisfy $\delta C_p < \epsilon/2$ for C_p and ϵ). Therefore, we have (4.6), and hence (B2) holds true.

5.2. Verifying Conditions (B3) and (B4′)

In this subsection, we look at Conditions (B3) and (B4′). Let $\tilde{f}_t(y) = h_n e^{-\lambda_{th}h_n} f_{\theta_{th}}(y)$. Then we have

$$p_{j,t}^1 = h_n^{-1}\int_{t_{j-1}}^{t_j}\int\int p_{x_{j-1},\alpha_{th}}^{c,\tau-t_{j-1}}(x)\tilde{f}_t(y)p_{x+y,\alpha_{th}}^{c,t_j-\tau}(x_j)dxdyd\tau.$$

By Proposition A.4 in the supplementary material, we can see that $p_{j,t}^1 > 0$ and hence $\tilde{p}_j(\alpha_{th}) > 0$ for any $x_{j-1}, x_j \in \mathbb{R}^m$ and $t \in [0,1]$. Therefore, we have

$$\frac{\partial_\theta^l\tilde{p}_j}{\tilde{p}_j}(\alpha_{th}) = \frac{\partial_\theta^l p_{j,t}^0}{p_{j,t}^0}1_{L_n}(\Delta x_j) + \frac{\partial_\theta^l p_{j,t}^1}{p_{j,t}^1}1_{L_n^c}(\Delta x_j).$$

For notational simplicity, we write

$$\Phi_j^t(\varphi(x_{j-1},x,y,x_j,\tau)) \\ = h_n^{-1}\int_{t_{j-1}}^{t_j}\int\int \varphi(x_{j-1},x,y,x_j,\tau)p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)\tilde{f}_t(y)p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)dxdyd\tau$$

for an integrable function φ. Then we can write

$$\frac{\partial_\theta p^1_{j,t}}{p^1_{j,t}} = \frac{1}{p^1_{j,t}}\Phi_j^t\left(\frac{\partial_\theta \tilde{f}_t}{\tilde{f}_t}(y) + \frac{\partial_\theta p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}{p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)} + \frac{\partial_\theta p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}{p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}\right), \tag{5.5}$$

and

$$\begin{aligned}\frac{\partial^2_\theta p^1_{j,t}}{p^1_{j,t}} = \frac{1}{p^1_{j,t}}\Phi_j^t\Bigg(&\frac{\partial^2_\theta \tilde{f}_t}{\tilde{f}_t}(y) + \frac{\partial^2_\theta p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}{p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)} + \frac{\partial^2_\theta p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}{p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)} + 2\frac{\partial_\theta p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}{p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}\frac{\partial_\theta \tilde{f}_t}{\tilde{f}_t}(y) \\ &+ 2\frac{\partial_\theta p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}{p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}\frac{\partial_\theta \tilde{f}_t}{\tilde{f}_t}(y) + 2\frac{\partial_\theta p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}{p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}\frac{\partial_\theta p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}{p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}\Bigg).\end{aligned} \tag{5.6}$$

We consider the limit of each term in the right-hand side. The following lemmas and proposition are useful when we deduce the limit of each term. We interpret $\partial^l_\theta \tilde{f}_t/\tilde{f}_t(z) = 0$ if $\tilde{f}_t(z) = 0$ or $z = 0$ for $l \in \{1,2\}$.

Lemma 5.2. *Assume (H1)–(H4). Then there exists $n_0 \in \mathbb{N}$ such that*

$$\int\left|\frac{1}{p^1_{j,t}}\Phi_j^t\left(\frac{\partial^l_\theta p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}{p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}\right)1_{L^c_n}\right|^{4/l}\tilde{p}_j(\alpha_{th})dx_j \le Ch_n^{4-l}(1+|x_{j-1}|)^C,$$

$$\int\left|\frac{1}{p^1_{j,t}}\Phi_j^t\left(\frac{\partial^l_\theta p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}{p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}\right)1_{L^c_n}\right|^{4/l}\tilde{p}_j(\alpha_{th})dx_j \le Ch_n^{4-l}(1+|x_{j-1}|)^C$$

for any $x_{j-1} \in \mathbb{R}^m$, $n \ge n_0$, $t \in [0,1]$, $1 \le j \le n$, and $l \in \{1,2\}$.

Proof. Jensen's inequality and Proposition A.2 in the supplementary material yield that

$$\begin{aligned}&\int\left|\frac{1}{p^1_{j,t}}\Phi_j^t\left(\frac{\partial^l_\theta p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}{p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}\right)1_{L^c_n}\right|^{4/l}\tilde{p}_j(\alpha_{th})dx_j \\ &\le \int \Phi_j^t\left(\left|\frac{\partial^l_\theta p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}{p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}\right|^{4/l}\right)dx_j \\ &= \lambda_{th}e^{-\lambda_{th}h_n}\int_{t_{j-1}}^{t_j}\int\left|\frac{\partial^l_\theta p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}{p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}\right|^{4/l}p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)dxd\tau \\ &\le Ch_n^{4-l}(1+|x_{j-1}|)^C\end{aligned}$$

for any $x_{j-1} \in \mathbb{R}^m$.

Similarly, Proposition A.2 in the supplementary material and (H4) yield

$$\begin{aligned}
&\int \left| \frac{1}{p^1_{j,t}} \Phi^t_j \left(\frac{\partial^l_\theta p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}{p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)} \right) 1_{L^c_n} \right|^{4/l} \tilde{p}_j(\alpha_{th}) dx_j \\
&\leq \int \Phi^t_j \left(\left| \frac{\partial^l_\theta p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}{p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)} \right|^{4/l} \right) dx_j \\
&\leq C h_n^{2-l} \int_{t_{j-1}}^{t_j} \int \int (1 + |x+y|)^C p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x) \tilde{f}_t(y) dx dy d\tau \\
&\leq C h_n^{4-l} (1 + |x_{j-1}|)^C
\end{aligned}$$

for any $x_{j-1} \in \mathbb{R}^m$. We also used the fact

$$\int (1 + |x|)^C p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x) dx = E[(1 + |X^{\alpha_{th},c}_{\tau-t_{j-1},x_{j-1}}|)^C] \leq C(1 + |x_{j-1}|)^C$$

by a similar argument to Proposition 3.1 in Shimizu and Yoshida [30]. □

Let $\Delta_j N_\theta = N_\theta((t_{j-1}, t_j] \times E)$.

Lemma 5.3. *Assume (H1)–(H4). Then, there exist positive constants* ι *and* C *such that for all* $j \in \{1, \dots, n\}$ *and* $x_{j-1} \in K''_{n,j-1}$,

$$\sup_{\alpha \in \Theta} P \left(X^\alpha_{t_j} - x_{j-1} \in \{z : F_\theta(z) = 0\} \cup \{0\} \middle| \Delta_j N_\theta = 1, X^\alpha_{t_{j-1}} = x_{j-1} \right) \leq C h_n^\iota. \tag{5.7}$$

Proof. Let ρ be the one in (H4). For the jump time τ_j on $(t_{j-1}, t_j]$ and large enough n, (H4) and a similar argument to (5.1) yield

$$\begin{aligned}
&P \left(X^\alpha_{t_j} - x_{j-1} \in \{z | F_\theta(z) = 0\} \cup \{0\} | \Delta_j N_\theta = 1, X^\alpha_{t_{j-1}} = x_{j-1} \right) \\
&\leq P \Bigg(\left\{ X^\alpha_{t_j} - x_{j-1} \in \{z | F_\theta(z) = 0\} \cup \{0\} \right\} \cap \left\{ \left| X^\alpha_{t_j} - X^\alpha_{\tau_j} + X^\alpha_{\tau_j -} - x_{j-1} \right| \leq h_n^\rho \right\} \\
&\qquad\qquad \Bigg| \Delta_j N_\theta = 1, X^\alpha_{t_{j-1}} = x_{j-1} \Bigg) \\
&\quad + P \left(\left| X^\alpha_{t_j} - X^\alpha_{\tau_j} + X^\alpha_{\tau_j -} - x_{j-1} \right| > h_n^\rho \middle| \Delta_j N_\theta = 1, X^\alpha_{t_{j-1}} = x_{j-1} \right) \\
&\leq P \left(d(X^\alpha_{\tau_j} - X^\alpha_{\tau_j -}, \partial \operatorname{supp}(F_\theta)) \leq h_n^\rho \middle| \Delta_j N_\theta = 1, X^\alpha_{t_{j-1}} = x_{j-1} \right) \\
&\quad + P \left(|X^\alpha_{\tau_j} - X^\alpha_{\tau_j -}| \leq h_n^\rho \middle| \Delta_j N_\theta = 1, X^\alpha_{t_{j-1}} = x_{j-1} \right) + C h_n^2 \\
&\leq h_n^\epsilon + C h_n^{(m+\gamma)\rho} + C h_n^2.
\end{aligned}$$

□

Proposition 5.4. *Assume (H1)–(H4). Then there exist a positive constant C, $n_0 \in \mathbb{N}$, and $\epsilon \in (0,1)$ such that*

$$\int \left| \frac{\partial_\theta^l p_{j,t}^1}{p_{j,t}^1} - \frac{\partial_\theta^l \tilde{f}_t}{\tilde{f}_t}(\Delta x_j) \right|^{4/l} 1_{L_n^c} \tilde{p}_j(\alpha_{th}) dx_j 1_{K''_{n,j-1}}(x_{j-1}) \leq C h_n^{1+\epsilon}(1+|x_{j-1}|)^C, \tag{5.8}$$

$$\int \left| \frac{\partial_\sigma^l p_{j,t}^1}{p_{j,t}^1} \right|^{4/l} 1_{L_n^c} \tilde{p}_j(\alpha_{th}) dx_j 1_{K''_{n,j-1}}(x_{j-1}) \leq C h_n (1+|x_{j-1}|)^C, \tag{5.9}$$

$$\int \left| \frac{\partial_\sigma \partial_\theta p_{j,t}^1}{p_{j,t}^1} \right|^{2} 1_{L_n^c} \tilde{p}_j(\alpha_{th}) dx_j 1_{K''_{n,j-1}}(x_{j-1}) \leq C h_n (1+|x_{j-1}|)^C \tag{5.10}$$

for $t \in [0,1]$, $x_{j-1} \in \mathbb{R}^m$, $l \in \{1,2\}$, $1 \leq j \leq n$, and $n \geq n_0$.

Proof. Let $Q_j = Q_j(x_{j-1}) = \{x_j \in E | \tilde{f}_t(\Delta x_j) \neq 0\} \cup \{0\}$. From Assumption (H4), Q_j does not depend on t. First we set $l = 1$. We decompose the following:

$$\begin{aligned}
&\int \left| \frac{\partial_\theta^l p_{j,t}^1}{p_{j,t}^1} - \frac{\partial_\theta^l \tilde{f}_t}{\tilde{f}_t}(\Delta x_j) \right|^{4} 1_{L_n^c} \tilde{p}_j(\alpha_{th}) dx_j 1_{K''_{n,j-1}} \\
&= \int \left\{ \left| \frac{\partial_\theta p_{j,t}^1}{p_{j,t}^1} - \frac{\partial_\theta \tilde{f}_t}{\tilde{f}_t}(\Delta x_j) \right|^{4} 1_{Q_j}(x_j) + \left| \frac{\partial_\theta p_{j,t}^1}{p_{j,t}^1} \right|^4 1_{Q_j^c}(x_j) \right\} 1_{L_n^c} \tilde{p}_j(\alpha_{th}) dx_j 1_{K''_{n,j-1}}.
\end{aligned}$$

(5.5), Jensen's inequality, (2.8), and Lemma 5.2 yield

$$\begin{aligned}
&\int \left| \frac{\partial_\theta p_{j,t}^1}{p_{j,t}^1} - \frac{\partial_\theta \tilde{f}_t}{\tilde{f}_t}(\Delta x_j) \right|^{4} 1_{Q_j}(x_j) 1_{L_n^c} \tilde{p}_j(\alpha_{th}) dx_j 1_{K''_{n,j-1}} \\
&\leq C \int \left| \frac{1}{p_{j,t}^1} \Phi_j^t \left(\frac{\partial_\theta \tilde{f}_t}{\tilde{f}_t}(y) - \frac{\partial_\theta \tilde{f}_t}{\tilde{f}_t}(\Delta x_j) \right) \right|^4 1_{Q_j}(x_j) p_{j,t}^1 dx_j 1_{K''_{n,j-1}} + C h_n^3 (1+|x_{j-1}|)^C \\
&\leq C \int \Phi_j^t \left(\left| \frac{\partial_\theta \tilde{f}_t}{\tilde{f}_t}(y) - \frac{\partial_\theta \tilde{f}_t}{\tilde{f}_t}(\Delta x_j) \right|^4 \right) 1_{Q_j}(x_j) dx_j 1_{K''_{n,j-1}} + C h_n^3 (1+|x_{j-1}|)^C \\
&\leq C \int \Phi_j^t (|y - \Delta x_j|^4 (1+|y|+|\Delta x_j|)^C) dx_j 1_{K''_{n,j-1}} + C h_n^3 (1+|x_{j-1}|)^C.
\end{aligned} \tag{5.11}$$

Obviously, $|y - \Delta x_j|^4 \leq C|x_j - x - y|^4 + C|x - x_{j-1}|^4$, and we can easily see that

$$\int |x_j - x - y|^p p_{x+y,\alpha_{th}}^{c,t_j-\tau}(x_j) dx_j = E_{x+y}[|X_{t_j-\tau}^{\alpha_{th},c} - x - y|^p] \leq C(t_j - \tau)^{p/2}(1+|x+y|)^p.$$

Hence it follows that

$$\begin{aligned}&\int \Phi_j^t(|y-\Delta x_j|^4(1+|y|+|\Delta x_j|)^C)dx_j 1_{K''_{n,j-1}}\\&\le Ch_n^{-1}\int_{t_{j-1}}^{t_j}\int\int (t_j-\tau)^2(1+|x+y|)^C p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)\tilde{f}_t\\&\qquad\qquad\times(1+|y|)^C(1+|x+y-x_{j-1}|)^C dxdyd\tau\\&\quad+Ch_n^{-1}\int_{t_{j-1}}^{t_j}\int\int p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)\tilde{f}_t|x-x_{j-1}|^4(1+|y|)^C(1+|x+y-x_{j-1}|)^C dxdyd\tau.\end{aligned}$$

Moreover, we similarly obtain the inequality

$$\int |x-x_{j-1}|^p p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)dx = E_{x_{j-1}}[|X^{\alpha_{th},c}_{\tau-t_{j-1}}-x_{j-1}|^p]\le C(\tau-t_{j-1})^{p/2}(1+|x_{j-1}|)^p,$$

and hence we have

$$\begin{aligned}&\int \Phi_j^t(|y-\Delta x_j|^4(1+|y|+|\Delta x_j|)^C)dx_j 1_{K''_{n,j-1}}\\&\le C\int_{t_{j-1}}^{t_j}(t_j-\tau)^2\int p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)(1+|x|)^C(1+|x-x_{j-1}|)^C dxd\tau\\&\quad+C\int_{t_{j-1}}^{t_j}\int p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)|x-x_{j-1}|^4(1+|x-x_{j-1}|)^C dxd\tau\\&\le Ch_n^3(1+|x_{j-1}|)^C+C\int_{t_{j-1}}^{t_j}(\tau-t_{j-1})^2d\tau(1+|x_{j-1}|)^C\\&\le Ch_n^3(1+|x_{j-1}|)^C.\end{aligned}\tag{5.12}$$

From the Cauchy–Schwarz inequality, (H4), Lemma 5.3, and a similar estimate to (5.11), we obtain

$$\begin{aligned}&\int\left|\frac{\partial_\theta p^1_{j,t}}{p^1_{j,t}}\right|^4 1_{Q_j^c}(x_j)1_{L_n^c}\tilde{p}_j dx_j 1_{K''_{n,j-1}}\\&\le\sqrt{\int\left|\frac{\partial_\theta p^1_{j,t}}{p^1_{j,t}}\right|^8 p^1_{j,t}dx_j}\\&\quad\times\sqrt{P\left(\{X_{t_j}-x_{j-1}\in\{z\in E|\tilde{f}_t(z)=0\}\cup\{0\}\}\cap\{\Delta_jN_\theta=1\}\,|X_{t_{j-1}}=x_{j-1}\right)}1_{K''_{n,j-1}}\\&\le Ch_n^{1+\iota/2}(1+|x_{j-1}|)^C,\end{aligned}\tag{5.13}$$

so that

$$\int\left|\frac{\partial_\theta p^1_{j,t}}{p^1_{j,t}}-\frac{\partial_\theta\tilde{f}_t}{\tilde{f}_t}(\Delta x_j)\right|^4 1_{L_n^c}\tilde{p}_j(\alpha_{th})dx_j1_{K''_{n,j-1}}(x_{j-1})\le Ch_n^{(1+\iota/2)\wedge 3}(1+|x_{j-1}|)^C.$$

Next, we show (5.8) for $l = 2$. A similar argument to Lemma 5.2 yields

$$\begin{aligned}&\int \left|\frac{1}{p^1_{j,t}}\Phi^t_j\left(\frac{\partial_\theta \mathcal{P}_{j,t}}{\mathcal{P}_{j,t}}\frac{\partial_\theta \tilde{f}_t}{\tilde{f}_t}(y)\right)\right|^2 1_{L^c_n}\tilde{p}_j(\alpha_{th})dx_j 1_{K''_{n,j-1}}\\&\quad\le \sqrt{\int \Phi^t_j\left(\left|\frac{\partial_\theta \mathcal{P}_{j,t}}{\mathcal{P}_{j,t}}\right|^4\right)dx_j}\sqrt{\int \Phi^t_j\left(\left|\frac{\partial_\theta \tilde{f}_t}{\tilde{f}_t}(y)\right|^4\right)dx_j}1_{K''_{n,j-1}} \le Ch^2_n(1+|x_{j-1}|)^C,\end{aligned}$$

where $\mathcal{P}_{j,t} = p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)$ or $p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)$.

Together with Jensen's inequality, (5.6), Lemma 5.2, (H4), and a similar argument to (5.12), we have

$$\begin{aligned}&\int \left|\frac{\partial^2_\theta p^1_{j,t}}{p^1_{j,t}} - \frac{\partial^2_\theta \tilde{f}_t}{\tilde{f}_t}(\Delta x_j)\right|^2 1_{Q_j}(x_j)1_{L^c_n}\tilde{p}_j(\alpha_{th})dx_j 1_{K''_{n,j-1}}\\&\quad\le C\int \left|\frac{1}{p^1_{j,t}}\Phi^t_j\left(\frac{\partial^2_\theta \tilde{f}_t}{\tilde{f}_t}(y) - \frac{\partial^2_\theta \tilde{f}_t}{\tilde{f}_t}(\Delta x_j)\right)\right|^2 1_{Q_j}(x_j)p^1_{j,t}dx_j 1_{K''_{n,j-1}} + Ch^2_n(1+|x_{j-1}|)^C\\&\quad\le C\int \Phi^t_j(|y-\Delta x_j|^2(1+|y|+|\Delta x_j|)^C)dx_j + Ch^2_n(1+|x_{j-1}|)^C\\&\quad\le Ch^2_n(1+|x_{j-1}|)^C.\end{aligned}$$

Together with a similar argument to (5.13), we obtain (5.8) for $l = 2$.

For the estimate for $\partial^l_\sigma p^1_{j,t}$, we first have

$$\begin{aligned}\frac{\partial_\sigma p^1_{j,t}}{p^1_{j,t}} &= \frac{1}{p^1_{j,t}}\Phi^t_j\left(\frac{\partial_\sigma p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}{p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)} + \frac{\partial_\sigma p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}{p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}\right),\\ \frac{\partial^2_\sigma p^1_{j,t}}{p^1_{j,t}} &= \frac{1}{p^1_{j,t}}\Phi^t_j\left(\frac{\partial^2_\sigma p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}{p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)} + \frac{\partial^2_\sigma p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}{p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)} + 2\frac{\partial_\sigma p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}{p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}\frac{\partial_\sigma p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}{p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}\right).\end{aligned}$$

Thanks to Jensen's inequality and Proposition A.2 in the supplementary material, we obtain

$$\begin{aligned}&\int \left|\frac{\partial_\sigma p^1_{j,t}}{p^1_{j,t}}\right|^4 1_{L^c_n}\tilde{p}_j(\alpha_{th})dx_j 1_{K''_{n,j-1}}\\&\quad\le C\int \Phi^t_j\left(\left|\frac{\partial_\sigma p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)}{p^{c,t_j-\tau}_{x+y,\alpha_{th}}(x_j)} + \frac{\partial_\sigma p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}{p^{c,\tau-t_{j-1}}_{x_{j-1},\alpha_{th}}(x)}\right|^4\right)dx_j 1_{K''_{n,j-1}}\\&\quad\le Ch_n(1+|x_{j-1}|)^C\end{aligned}$$

and

$$\int \left|\frac{\partial_\sigma^2 p_{j,t}^1}{p_{j,t}^1}\right|^2 1_{L_n^c} \tilde{p}_j(\alpha_{th}) dx_j 1_{K''_{n,j-1}}$$

$$\leq C \int \Phi_j^t \left(\left| \frac{\partial_\sigma^2 p_{x+y,\alpha_{th}}^{c,t_j-\tau}(x_j)}{p_{x+y,\alpha_{th}}^{c,t_j-\tau}(x_j)} + \frac{\partial_\sigma^2 p_{x_{j-1},\alpha_{th}}^{c,\tau-t_{j-1}}(x)}{p_{x_{j-1},\alpha_{th}}^{c,\tau-t_{j-1}}(x)} + 2 \frac{\partial_\sigma p_{x_{j-1},\alpha_{th}}^{c,\tau-t_{j-1}}(x)}{p_{x_{j-1},\alpha_{th}}^{c,\tau-t_{j-1}}(x)} \frac{\partial_\sigma p_{x+y,\alpha_{th}}^{c,t_j-\tau}(x_j)}{p_{x+y,\alpha_{th}}^{c,t_j-\tau}(x_j)} \right|^2 \right) dx_j 1_{K''_{n,j-1}}$$

$$\leq C h_n (1 + |x_{j-1}|)^C.$$

Similarly, we have

$$\int \left|\frac{\partial_\sigma \partial_\theta p_{j,t}^1}{p_{j,t}^1}\right|^2 1_{L_n^c} \tilde{p}_j(\alpha_{th}) dx_j 1_{K''_{n,j-1}} \leq C h_n (1 + |x_{j-1}|)^C. \qquad \square$$

Proposition 5.5. *Assume (H1)–(H4). Then (B3) holds true.*

Proof. (C.2) and Proposition A.2 in the supplementary material and Proposition 5.4 yield

$$\sum_{j=1}^n \sup_{t\in[0,1]} D_{j,h}^{-1} \int \left(|\zeta_{j,t}^{2,h}|^2 + |\zeta_{j,t}^{1,h}|^4\right) \tilde{p}_j(\alpha_{th}) dx_j 1_{K''_{n,j-1}}$$

$$\leq C \sum_{j=1}^n \sup_{t\in[0,1]} D_{j,h}^{-1} \int \left(\left|\frac{1}{n} \frac{\partial_\sigma^2 \tilde{p}_j}{\tilde{p}_j}\right|^2 + \left|\frac{1}{T_n} \frac{\partial_\theta^2 \tilde{p}_j}{\tilde{p}_j}\right|^2 + \left|\frac{1}{\sqrt{nT_n}} \frac{\partial_\theta \partial_\sigma \tilde{p}_j}{\tilde{p}_j}\right|^2 + \frac{1}{n^2}\left|\frac{\partial_\sigma \tilde{p}_j}{\tilde{p}_j}\right|^4 + \frac{1}{T_n^2}\left|\frac{\partial_\theta \tilde{p}_j}{\tilde{p}_j}\right|^4 \right) \tilde{p}_j(\alpha_{th}) dx_j 1_{K''_{n,j-1}}$$

$$\leq \frac{C}{T_n^2} \sum_{j=1}^n \sup_{t\in[0,1]} D_{j,h}^{-1} \int \left(\left|\frac{\partial_\theta^2 \tilde{f}_t}{\tilde{f}_t}\right|^2 + \left|\frac{\partial_\theta \tilde{f}_t}{\tilde{f}_t}\right|^4 \right) (\Delta x_j) 1_{L_n^c} \tilde{p}_j(\alpha_{th}) dx_j 1_{K''_{n,j-1}} + o_p(1).$$

Then (C.2) in the supplementary material, (B1), (H4), and a similar argument to (5.12) yield the conclusion because $\partial_\theta^2 \tilde{f}_t / \tilde{f}_t = \partial_\theta^2 \log \tilde{f}_t + (\partial_\theta \log \tilde{f}_t)^2$. $\square$

We turn to observe (B4′). Let $\tilde{p}_{j,0} = \tilde{p}_j(\alpha_0)$. (B1), (C.2) in the supplementary material, Proposition 5.4, and the Cauchy–Schwarz inequality yield

$$\sum_{j=1}^n D_{j,0}^{-1} \int \tilde{\eta}_j \tilde{\eta}_j^\top \tilde{p}_{j,0} dx_j$$

$$= \sum_{j=1}^n D_{j,0}^{-1} \int \begin{pmatrix} \frac{\partial_\sigma p_{j,0}^0 (\partial_\sigma p_{j,0}^0)^\top}{n (p_{j,0}^0)^2} 1_{L_n} & \frac{\partial_\sigma p_{j,0}^0 (\partial_\theta p_{j,0}^0)^\top}{n\sqrt{h_n} (p_{j,0}^0)^2} 1_{L_n} \\ \frac{\partial_\theta p_{j,0}^0 (\partial_\sigma p_{j,0}^0)^\top}{n\sqrt{h_n} (p_{j,0}^0)^2} 1_{L_n} & \frac{\partial_\theta \tilde{f}_0 \partial_\theta \tilde{f}_0^\top}{n h_n \tilde{f}_0^2} (\Delta x_j) 1_{L_n^c} + \frac{\partial_\theta p_{j,0}^0 (\partial_\theta p_{j,0}^0)^\top}{n h_n (p_{j,0}^0)^2} 1_{L_n} \end{pmatrix} \tilde{p}_{j,0} dx_j + o_p(1).$$

Hence, (B4′) follows if we show

$$\frac{1}{n}\sum_{j=1}^{n} D_{j,0}^{-1}\int \frac{\partial_\sigma p_{j,0}^0(\partial_\sigma p_{j,0}^0)^\top}{(p_{j,0}^0)^2} 1_{L_n}\tilde{p}_{j,0}dx_j \to^{\tilde{P}_{\alpha_0,n}} \Gamma_1, \tag{5.14}$$

$$\frac{1}{nh_n}\sum_{j=1}^{n} D_{j,0}^{-1}\int \left\{\frac{\partial_\theta \tilde{f}_0 \partial_\theta \tilde{f}_0^\top}{\tilde{f}_0^2}(\Delta x_j)1_{L_n^c} + \frac{\partial_\theta p_{j,0}^0(\partial_\theta p_{j,0}^0)^\top}{(p_{j,0}^0)^2} 1_{L_n}\right\}\tilde{p}_{j,0}dx_j \to^{\tilde{P}_{\alpha_0,n}} \Gamma_2, \tag{5.15}$$

$$\frac{1}{n\sqrt{h_n}}\sum_{j=1}^{n} D_{j,0}^{-1}\int \frac{\partial_\sigma p_{j,0}^0(\partial_\theta p_{j,0}^0)^\top}{(p_{j,0}^0)^2} 1_{L_n}\tilde{p}_{j,0}dx_j \to^{\tilde{P}_{\alpha_0,n}} 0. \tag{5.16}$$

Proofs of (5.14)–(5.16) are given in the supplementary material, which complete the proof of Theorem 2.3.

Acknowledgements

The authors would like to thank the anonymous referees, an Associate Editor and the Editor for their constructive comments that improved the quality of this paper. Teppei Ogihara and Yuma Uehara were supported by Japan Society for the Promotion of Science KAKENHI Grant Number JP21H00997.

Supplementary Material

Supplement to "Local asymptotic normality for ergodic jump-diffusion processes via transition density approximation" [26] (DOI: 10.3150/22-BEJ1544SUPP; .pdf). The supplementary material gives auxiliary results for diffusion processes, and the proofs of Theorem 3.1 and the results in Sections 4 and 5.2.

References

[1] Aït-Sahalia, Y. and Jacob, J. (2007). Volatility estimators for discretely sampled Lévy processes. *Ann. Statist.* **35** 355–392. MR2332279 https://doi.org/10.1214/009053606000001190

[2] Applebaum, D. (2009). *Lévy Processes and Stochastic Calculus*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **116**. Cambridge: Cambridge Univ. Press. MR2512800 https://doi.org/10.1017/CBO9780511809781

[3] Basawa, I.V. and Scott, D.J. (1983). *Asymptotic Optimal Inference for Nonergodic Models*. *Lecture Notes in Statistics* **17**. New York-Berlin: Springer. MR0688650

[4] Bibinger, M. and Winkelmann, L. (2015). Econometrics of co-jumps in high-frequency data with noise. *J. Econometrics* **184** 361–378. MR3291008 https://doi.org/10.1016/j.jeconom.2014.10.004

[5] Choi, S., Hall, W.J. and Schick, A. (1996). Asymptotically uniformly most powerful tests in parametric and semiparametric models. *Ann. Statist.* **24** 841–861. MR1394992 https://doi.org/10.1214/aos/1032894469

[6] Clément, E., Delattre, S. and Gloter, A. (2014). Asymptotic lower bounds in estimating jumps. *Bernoulli* **20** 1059–1096. MR3217438 https://doi.org/10.3150/13-BEJ515

[7] Clément, E. and Gloter, A. (2015). Local asymptotic mixed normality property for discretely observed stochastic differential equations driven by stable Lévy processes. *Stochastic Process. Appl.* **125** 2316–2352. MR3322866 https://doi.org/10.1016/j.spa.2015.01.002

[8] Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. *Chapman & Hall/CRC Financial Mathematics Series*. Boca Raton, FL: CRC Press/CRC. MR2042661

[9] Eguchi, S. and Masuda, H. (2018). Schwarz type model comparison for LAQ models. *Bernoulli* **24** 2278–2327. MR3757530 https://doi.org/10.3150/17-BEJ928

[10] Fukasawa, M. and Ogihara, T. Malliavin calculus techniques for local asymptotic mixed normality and their application to degenerate diffusions. Available at arXiv:2005.14599.

[11] Gloter, A., Loukianova, D. and Mai, H. (2018). Jump filtering and efficient drift estimation for Lévy-driven SDEs. *Ann. Statist.* **46** 1445–1480. MR3819106 https://doi.org/10.1214/17-AOS1591

[12] Gobet, E. (2001). Local asymptotic mixed normality property for elliptic diffusion: A Malliavin calculus approach. *Bernoulli* **7** 899–912. MR1873834 https://doi.org/10.2307/3318625

[13] Gobet, E. (2002). LAN property for ergodic diffusions with discrete observations. *Ann. Inst. Henri Poincaré Probab. Stat.* **38** 711–737. MR1931584 https://doi.org/10.1016/S0246-0203(02)01107-X

[14] Hájek, J. (1969/70). A characterization of limiting distributions of regular estimates. *Z. Wahrsch. Verw. Gebiete* **14** 323–330. MR0283911 https://doi.org/10.1007/BF00533669

[15] Hájek, J. (1972). Local asymptotic minimax and admissibility in estimation. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. I: Theory of Statistics* 175–194. Berkeley, Calif.: Univ. California Press. MR0400513

[16] Ibragimov, I.A. and Has'minskiĭ, R.Z. (1981). *Statistical Estimation*. *Applications of Mathematics* **16**. New York-Berlin: Springer. Asymptotic theory, Translated from the Russian by Samuel Kotz. MR0620321

[17] Inatsugu, H. and Yoshida, N. (2021). Global jump filters and quasi-likelihood analysis for volatility. *Ann. Inst. Statist. Math.* **73** 555–598. MR4247071 https://doi.org/10.1007/s10463-020-00768-x

[18] Jeganathan, P. (1982). On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal. *Sankhyā Ser. A* **44** 173–212. MR0688800

[19] Kawai, R. and Masuda, H. (2013). Local asymptotic normality for normal inverse Gaussian Lévy processes with high-frequency sampling. *ESAIM Probab. Stat.* **17** 13–32. MR3002994 https://doi.org/10.1051/ps/2011101

[20] Kohatsu-Higa, A., Nualart, E. and Tran, N.K. (2017). LAN property for an ergodic diffusion with jumps. *Statistics* **51** 419–454. MR3609328 https://doi.org/10.1080/02331888.2016.1239727

[21] Mai, H. (2014). Efficient maximum likelihood estimation for Lévy-driven Ornstein-Uhlenbeck processes. *Bernoulli* **20** 919–957. MR3178522 https://doi.org/10.3150/13-BEJ510

[22] Mancini, C. (2011). The speed of convergence of the threshold estimator of integrated variance. *Stochastic Process. Appl.* **121** 845–855. MR2770909 https://doi.org/10.1016/j.spa.2010.12.001

[23] Masuda, H. (2007). Ergodicity and exponential β-mixing bounds for multidimensional diffusions with jumps. *Stochastic Process. Appl.* **117** 35–56. MR2287102 https://doi.org/10.1016/j.spa.2006.04.010

[24] Masuda, H. (2008). On stability of diffusions with compound-Poisson jumps. *Bull. Inform. Cybernet.* **40** 61–74. MR2484446

[25] Ogihara, T. (2015). Local asymptotic mixed normality property for nonsynchronously observed diffusion processes. *Bernoulli* **21** 2024–2072. MR3378458 https://doi.org/10.3150/14-BEJ634

[26] Ogihara, T. and Uehara, Y. (2023). Supplement to "Local asymptotic normality for ergodic jump-diffusion processes via transition density approximation." https://doi.org/10.3150/22-BEJ1544SUPP

[27] Ogihara, T. and Yoshida, N. (2011). Quasi-likelihood analysis for the stochastic differential equation with jumps. *Stat. Inference Stoch. Process.* **14** 189–229. MR2832818 https://doi.org/10.1007/s11203-011-9057-z

[28] Prakasa Rao, B.L.S. (1999). *Semimartingales and Their Statistical Inference*. *Monographs on Statistics and Applied Probability* **83**. Boca Raton, FL: CRC Press/CRC. MR1689166

[29] Shimizu, Y. (2008). A practical inference for discretely observed jump-diffusions from finite samples. *J. Japan Statist. Soc.* **38** 391–413. MR2510946 https://doi.org/10.14490/jjss.38.391

[30] Shimizu, Y. and Yoshida, N. (2006). Estimation of parameters for diffusion processes with jumps from discrete observations. *Stat. Inference Stoch. Process.* **9** 227–277. MR2252242 https://doi.org/10.1007/s11203-005-8114-x

[31] Sweeting, T.J. (1980). Uniform asymptotic normality of the maximum likelihood estimator. *Ann. Statist.* **8** 1375–1381. MR0594652

Received February 2022 and revised August 2022