Mighty convergence of the Gaussian quasi-likelihood random fields for ergodic Lévy driven SDE observed at high frequency

Masuda, Hiroki
Institute of Mathematics for Industry, Kyushu University

出版情報：MI Preprint Series. 2012-2, 2012-02-10. 九州大学大学院数理学研究院
バージョン：
権利関係：

MI Preprint Series

Kyushu University
The Global COE Program
Math-for-Industry Education & Research Hub

Mighty convergence of the Gaussian quasi-likelihood random fields for ergodic Levy driven SDE observed at high frequency

Hiroki Masuda

MI 2012-2

(Received February 10, 2012)

Faculty of Mathematics
Kyushu University
Fukuoka, JAPAN

Mighty convergence of the Gaussian quasi-likelihood random fields for ergodic Lévy driven SDE observed at high frequency *

Hiroki Masuda

Institute of Mathematics for Industry, Kyushu University,
744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan.
Email: hiroki@imi.kyushu-u.ac.jp

Abstract

This paper investigates the Gaussian quasi-likelihood estimation of an exponentially ergodic multidimensional Markov process, which is expressed as a solution to a Lévy driven stochastic differential equations whose coefficients are supposed to be known except for the finite-dimensional parameters to be estimated. We suppose that the process is observed under the condition for the rapidly increasing experimental design. By means of the polynomial type large deviation inequality, the mighty convergence of the corresponding statistical random fields is derived, which especially leads to the asymptotic normality at rate $\sqrt{nh_n}$ for all the target parameters, and also to the convergence of their moments. In our results, the diffusion coefficient may be degenerate, or even null. Although the resulting estimator is not asymptotically efficient in the presence of jumps, we do not require any specific form of the driving Lévy measure, rendering that the proposed estimation procedure is practical and somewhat robust to underlying model specification.

Keywords. Exponential ergodicity, Gaussian quasi-likelihood estimation, high-frequency sampling, Lévy driven stochastic differential equation, polynomial type large deviation inequality.

2010 Mathematics Subject Classification. 62M05.

1 Introduction

Let $X = (X_t)_{t\in\mathbb{R}_+}$ be a solution to the Stochastic Differential Equation (SDE)

$$dX_t = a(X_t, \alpha)dt + b(X_t, \beta)dw_t + c(X_{t-}, \beta)dJ_t, \tag{1}$$

where the ingredients involved are as follows.

- The unknown finite-dimensional unknown parameter

  $$\theta = (\alpha, \beta) \in \Theta_\alpha \times \Theta_\beta =: \Theta,$$

  where, for simplicity, the parameter spaces $\Theta_\alpha \subset \mathbb{R}^{p_\alpha}$ and $\Theta_\beta \subset \mathbb{R}^{p_\beta}$ are supposed to be bounded convex domains; the parameter α (resp. β) affects local trend (resp. local dispersion).

- An r'-dimensional standard Wiener process w and an r''-dimensional centered pure-jump Lévy process J, whose Lévy measure is denoted by ν.

- The initial variable X_0 independent of (w, J), with $\eta := \mathcal{L}(X_0)$ possibly depending on θ.

- The measurable functions $a : \mathbb{R}^d \times \Theta_\alpha \to \mathbb{R}^d$, $b : \mathbb{R}^d \times \Theta_\beta \to \mathbb{R}^d \otimes \mathbb{R}^{r'}$, and $c : \mathbb{R}^d \times \Theta_\beta \to \mathbb{R}^d \otimes \mathbb{R}^{r''}$, all of which are known except for θ.

*This version: February 10, 2012

Incorporation of the jump part extends a continuous-path diffusion parametric model, which are nowadays widely used in many application fields. We denote by P_θ the image measure of a solution process X associated with $\theta \in \Theta \subset \mathbb{R}^p$, where $p := p_\alpha + p_\beta$. Suppose that the true parameter $\theta_0 = (\alpha_0, \beta_0) \in \Theta$ does exist, with P_0 denoting the shorthand for the true image measure P_{θ_0}, and that X is not completely (continuously) observed but only discretely at high frequency under the condition for the rapidly increasing experimental design: we are given a sample $(X_{t_0}, X_{t_1}, \dots, X_{t_n})$, where $t_j = t_j^n = jh_n$ for some $h_n > 0$ such that

$$T_n := nh_n \to \infty \quad \text{and} \quad nh_n^2 \to 0 \tag{2}$$

for $n \to \infty$. The main objective of this paper is to estimate θ_0 under the exponential ergodicity of X; the equidistant sampling assumption can be weaken to some extent as long as the long-term and high-frequency framework is concerned, however, it is just a technical extension just making the presentation notationally messy, hence we do not deal with it here to make the presentation clearer. It is common knowledge that the maximum likelihood estimation is generally infeasible since the transition probability is most often unavailable in a closed form. This implies that the conventional statistical analyses based on the genuine likelihood have no utility. For this reason, we have to resort to some other feasible estimation procedure, which could be a lot of things. Among several possibilities, we are concerned here with the *Gaussian Quasi Likelihood (GQL)* function defined as if the conditional distributions of X_{t_j} given $X_{t_{j-1}}$ are Gaussian with approximate but explicit mean vector and covariance matrix; see (10) below.

The terminology "Quasi Likelihood" has originated as the pioneering work Wedderburn [39], the concept of which formed a basis of generalized linear regression. The GQL based estimation has been known to have the advantage of computational simplicity and robustness for misspecification of the noise distribution, and well-established as a fundamental tool in estimating possibly non-Gaussian and dependent statistical models. Just to be a little more precise, consider a time-series $Y_1, \dots, Y_n$ in $\mathbb{R}$ with a fixed Y_0, and denote by $m_{j-1}(\theta) \in \mathbb{R}$ and $v_{j-1}(\theta) > 0$ the conditional mean and conditional variance of Y_j given $(Y_0, \dots, Y_{j-1})$, where θ is an unknown parameter of interest. Then, the *Gaussian Quasi Maximum Likelihood Estimator (GQMLE)* is defined to be a maximizer of the function

$$\theta \mapsto \sum_{j=1}^n \log \left\{ \frac{1}{\sqrt{2\pi v_{j-1}(\theta)}} \exp\left(-\frac{(Y_j - m_{j-1}(\theta))^2}{2v_{j-1}(\theta)} \right) \right\}.$$

Namely, we compute the likelihood of $(Y_1, Y_2, \dots, Y_n)$ as if the conditional law of Y_j given $(Y_1, \dots, Y_{j-1})$ is Gaussian with mean $m_{j-1}(\theta)$ and variance $v_{j-1}(\theta)$, so that only the structures of the conditional mean and variance do matter. Although it is not asymptotically efficient in general, it can serve as a widely applicable estimation procedure. One can consult Heyde [9] for an extensive and systematic account of statistical inference based on the GQL. The GQL has been a quite popular tool for (semi)parametric estimation, and especially there exist vast amounts of literatures concerning asymptotics of the GQL for time series models with possibly non-Gaussian error sequence; among others, we refer to Straumann and Mikosch [35] for a class of conditionally heteroscedastic time series models, and Bardet and Wintenburger [1] for multidimensional causal time series, as well as the references therein.

Let us return to our framework. On the one hand, for the diffusion case (where $c \equiv 0$), the the GQL-estimation issue has been solved under some regularity conditions. Especially, the GQL leads to an asymptotically efficient estimator, where the crucial point is that the optimal rates of convergence for estimating α and β are different and given by $\sqrt{T_n}$ and $\sqrt{n}$, respectively; see Gobet [8] for the local asymptotic normality of the corresponding statistical experiments. For how to construct an explicit contrast function, we refer to Yoshida [40] and Kessler [15] as well as the references therein; specifically, they employed a discretized version of the continuous-observation likelihood process, or a higher-order local-Gauss approximation of the transition density, respectively. Sørensen [34] includes an extensive bibliography of many existing results including explicit martingale estimating functions for discretely observed diffusions (not necessarily at high frequency). On the other hand, the issue has not been addressed enough in the presence of jumps (possibly of infinite variation). The question we should then ask is what will occur when one adopts the GQL function. In this paper, we will provide sufficient conditions under which the GQL random field associated with our statistical experiments converges in a mighty mode. We will apply Yoshida [41] to derive the mighty convergence with the limit being shifted Gaussian. As results, we will obtain: an asymptotically normally distributed estimator at rate

$\sqrt{T_n}$ for both α and β; and also, very importantly, the convergence of their moments to the corresponding ones of the limit centered Gaussian distribution. Different from the diffusion case, the GQL does not lead to an asymptotically efficient estimator in the presence of jumps, and is not even rate-efficient for β. Nevertheless, as mentioned before, it has at least two practically important advantages: first, the computation of estimates is straightforward; second, the estimation procedure is robust to modelling Lévy measure, which we actually do not need to specify.

We should mention that the convergence of moments especially serves as a fundamental tool when analyzing asymptotic behavior of the expectations of statistics depending on the estimator; for example, asymptotic bias and mean squared prediction error, model-selection devices (information criteria), remainder estimation in higher-order inference. In the past, several authors have investigated such a strong mode of convergence of estimators: see Bhansali and Papangelou [2], Chan and Ing [3], Findley and Wei [5], Inagaki and Ogata [11], Jeganathan [13, 14], Ogata and Inagaki [30], Sieders and Dzhaparidze [33], and Uchida [36], as well as Ibragimov and Has'minski [10], Kutoyants [19, 20], and Yoshida [41]. See also the recent paper Uchida and Yoshida [37] for an adaptive parametric estimation of diffusions with moment convergence of estimators under the sampling design $nh_n^k \to 0$ for arbitrary integer $k \geq 2$.

The rest of this paper is organized as follows. Section 2 introduces our GQL random field and presents its asymptotic behavior, together with a small numerical example for observing finite-sample performance of the GQMLE. Section 3 provides a somewhat general result concerning the mighty convergence, based on which we prove our main result in Section 4. In Section 5, we prove a fairly simple criterion for the exponential ergodicity assumption in dimension one, only in terms of the coefficient (a, b, c) and the Lévy measure $\nu(dz)$.

Throughout this paper, asymptotics are taken for $n \to \infty$ unless otherwise mentioned, and the following notation is used.

- I_r denotes the $r \times r$-identity matrix.
- Given a multilinear form $M = \{M^{(i_1 i_2 \dots i_K)} : i_k = 1, \dots, d_k; k = 1, \dots, K\} \in \mathbb{R}^{d_1} \otimes \dots \otimes \mathbb{R}^{d_K}$ and variables $u_k = \{u_k^{(i)}\}_{i \leq d_k} \in \mathbb{R}^{d_k}$, we write

$$M[u_1, \dots, u_K] = \sum_{i_1=1}^{d_1} \dots \sum_{i_K=1}^{d_K} M^{(i_1 i_2 \dots i_K)} u_1^{(i_1)} \dots u_K^{(i_K)}.$$

  The correspondences of indices of M and u_k will be clear from each context. Some of u_k may be missing in "$M[u_1, \dots, u_K]$", so that the resulting form again defines a multilinear form; for example, $M[u_3, \dots, u_K] \in \mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}$. When $K \leq 2$, identifying M as a vector or matrix, we write: $M^{\otimes 2} = MM^\top$ with $\top$ denoting the transpose; furthermore, $|M|$ denotes either, depending on the context, $\det(M)$ when $d_1 = d_2$, or any matrix norm of M.
- ∂_a^m stands for the bundled mth partial differential operator with respect to $a = \{a^{(i)}\}$.
- C denotes generic positive constant possibly varying from line to line, and we write $x_n \lesssim y_n$ if $x_n \leq C y_n$ a.s. for every n large enough.

2 Gaussian quasi-likelihood estimation

We denote by $(\Omega, \mathcal{F}, \mathbf{F} = (\mathcal{F}_t)_{t \in \mathbb{R}_+}, P)$ be a complete filtered probability space on which the process X given by (1) is defined: the initial variable X_0 being $\mathcal{F}_0$-measurable, and (w, J) is $\mathbf{F}$-adapted.

2.1 Assumptions

Here we list up our assumptions. We will give remarks on some of them in Section 2.3.

Assumption 2.1 (Moments). *$E[J_1] = 0$, $E[J_1^{\otimes 2}] = I_{r''}$, and $E[|J_1|^q] < \infty$ for all $q > 0$.*

We introduce the function $V:\mathbb{R}^d\times\Theta_\beta\to\mathbb{R}^d\otimes\mathbb{R}^d$ by

$$V=b^{\otimes 2}+c^{\otimes 2}.$$

For each θ, the function $x\mapsto V(x,\beta)$ can be viewed as an approximate local covariance matrix of the law of $h_n^{-1/2}(X_{h_n}-x)$ under $P_\theta[\cdot|X_0=x]$.

Assumption 2.2 (Smoothness). *(a) The coefficient (a,b,c) admits partial derivatives jointly continuous in x and θ, such that*

$$\sup_{(x,\alpha)\in\mathbb{R}^d\times\Theta}\{|\partial_x a(x,\alpha)|+|\partial_x b(x,\beta)|+|\partial_x c(x,\beta)|\}<\infty,$$

and that, for each $k\in\{0,1,2\}$ and $l\in\{0,1,\dots,5\}$ there exists a constant $C(k,l)\geq 0$ for which

$$\sup_{(x,\theta)\in\mathbb{R}^d\times\Theta}(1+|x|)^{-C(k,l)}\left\{|\partial_x^k\partial_\alpha^l a(x,\alpha)|+|\partial_x^k\partial_\beta^l b(x,\beta)|+|\partial_x^k\partial_\beta^l c(x,\beta)|\right\}<\infty.$$

(b) $V(x,\beta)$ is invertible for each (x,β), and there exists a constant $C(V)\geq 0$ such that

$$\sup_{(x,\theta)\in\mathbb{R}^d\times\Theta}(1+|x|)^{-C(V)}|V^{-1}(x,\beta)|<\infty.$$

When we consider large-time asymptotics, the stability property of X much affects the statistical analyses in essential ways. A typical situation to be considered is that X is ergodic. We here impose the stronger stability condition. Let $P_t(x,dy)$ denote the transition probability $P_0[X_t\in dy|X_0=x]$, and $\|m\|_\rho:=\sup_{|f|\leq\rho}|m(f)|$ for a signed measure m on the d-dimensional Borel space.

Assumption 2.3 (Stability). *For any $q\geq 2$, the following conditions hold true for $g(x):=1+|x|^q$:*

(a) There exists a probability measure π_0 and a constant $a>0$ such that

$$\sup_{t\in\mathbb{R}_+}e^{at}\|P_t(x,\cdot)-\pi_0(\cdot)\|_g\lesssim g(x),\quad x\in\mathbb{R}^d. \tag{3}$$

(b)

$$\sup_{t\in\mathbb{R}_+}E_0\left[g(X_t)\right]<\infty. \tag{4}$$

The condition (3) with g_q being replaced by the constant 1 is the *exponential ergodicity*, which in particular entails the ergodic theorem: the limit π_0 is a unique invariant distribution such that, for every $f\in L^1(\pi_0)$

$$\frac{1}{T_n}\int_0^{T_n}f(X_t)dt\to^p\int f(x)\pi_0(dx), \tag{5}$$

where $\to^p$ stands for the convergence in probability. We also note that Assumption 2.3 entails the *exponential absolute regularity*, also referred to as the *exponential β-mixing property*. This means that $\beta_X(t)=O(e^{-at})$ as $t\to\infty$ for some $a>0$, where β_X denotes the β-mixing coefficient

$$\beta_X(t):=\sup_{s\in\mathbb{R}_+}\int\|P_t(x,\cdot)-\eta P_{s+t}(\cdot)\|\eta P_s(dx),$$

where $\eta P_t:=\mathcal{L}(X_t)$ and $\|m\|:=\|m\|_1$. Let us recall that the exponential absolute regularity implies the *exponential strong-mixing property*, which plays an essential role in Yoshida [41, Lemma 4], which we will apply in the proof of Theorem 2.7.

Several sufficient conditions for Assumption 2.3 are known; for diffusion processes, see the references of Masuda [25, 26] for some details. In the presence of the jump component, verification of (3) can become much more involved. Especially if the coefficients are nonlinear and the Lévy process J is of infinite variation, the verification may be far from being a trivial matter. We refer to Kulik [16, 17], Maruyama and Tanaka [22], Menaldi and Robin [28], Meyn and Tweedie [29], and Wang [38] as well as Masuda [25, 26] for some general results concerning the exponential ergodicity. For the sake of convenience, focusing on the univariate case and setting ease of verification above generality, we will provide in Proposition 5.1 sufficient conditions for Assumption 2.3, in a form enabling us to deal with cases of nonlinear coefficients and infinite-variation J; see also Remark 5.4.

Define $\mathbb{G}_\infty(\theta) = (\mathbb{G}_\infty^\alpha(\theta), \mathbb{G}_\infty^\beta(\beta)) \in \mathbb{R}^p$ by

$$\mathbb{G}_\infty^\alpha(\theta) = \int \partial_\alpha a(x,\alpha)\left[V^{-1}(x,\beta)\left[a(x,\alpha_0) - a(x,\alpha)\right]\right]\pi_0(dx), \tag{6}$$

$$\mathbb{G}_\infty^\beta(\theta) = \int \left\{V^{-1}(\partial_\beta V)V^{-1}(x,\beta)\right\}\left[V(x,\beta_0) - V(x,\beta)\right]\pi_0(dx). \tag{7}$$

(In (7), we regarded "$V^{-1}(\partial_\beta V)V^{-1}(x,\beta)$" as a bilinear form with dimensions of indices being p_β and d^2.) Further, let $\mathbb{G}'_\infty(\theta_0) := \mathrm{diag}\{\mathbb{G}_\infty'^\alpha(\theta_0), \mathbb{G}_\infty'^\beta(\theta_0)\} \in \mathbb{R}^p \otimes \mathbb{R}^p$, where, for each $v'_1, v'_2 \in \mathbb{R}^{p_\alpha}$ and $v''_1, v''_2 \in \mathbb{R}^{p_\beta}$,

$$\mathbb{G}_\infty'^\alpha(\theta_0)[v'_1, v'_2] = -\int V^{-1}(x,\beta_0)\left[\partial_\alpha a(x,\alpha_0)[v'_1], \partial_\alpha a(x,\alpha_0)[v'_2]\right]\pi_0(dx), \tag{8}$$

$$\mathbb{G}_\infty'^\beta(\theta_0)[v''_1, v''_2] = -\int \mathrm{trace}\left\{V^{-1}(\partial_\beta V)V^{-1}(\partial_\beta V)(x,\beta_0)[v''_1, v''_2]\right\}\pi_0(dx). \tag{9}$$

Assumption 2.4 (Identifiability). *There exist positive constants* $\chi_\alpha = \chi_\alpha(\theta_0)$ *and* $\chi_\beta = \chi_\beta(\theta_0)$ *such that* $|\mathbb{G}_\infty^\alpha(\theta)|^2 \ge \chi_\alpha|\alpha - \alpha_0|^2$ *and* $|\mathbb{G}_\infty^\beta(\beta)|^2 \ge \chi_\beta|\beta - \beta_0|^2$ *for every* $\theta \in \Theta$.

Assumption 2.5 (Nondegeneracy). *Both* $\mathbb{G}_\infty'^\alpha(\theta_0)$ *and* $\mathbb{G}_\infty'^\beta(\theta_0)$ *are invertible.*

Assumptions 2.4 and 2.5 are quite typical in statistical estimation. As seen in Lemma 2.6 below, the both assumptions are implied by a kind of uniform nonsingularity. Define two bilinear forms $\bar{A}(\alpha', \alpha'', \beta')$ and $\bar{B}(\beta', \beta'')$ by, just like (8) and (9),

$$\bar{A}(\alpha', \alpha'', \beta')[v'_1, v'_2] = \int V^{-1}(x,\beta')\left[\partial_\alpha a(x,\alpha')[v'_1], \partial_\alpha a(x,\alpha'')[v'_2]\right]\pi_0(dx),$$

$$\bar{B}(\beta', \beta'')[v''_1, v''_2] = \int \mathrm{trace}\left\{(V^{-1}(\partial_\beta V)V^{-1})(x,\beta')\partial_\beta V(x,\beta'')[v''_1, v''_2]\right\}\pi_0(dx).$$

Lemma 2.6. *Suppose that* $\bar{A}(\alpha', \alpha'', \beta')$ *and* $\bar{B}(\beta', \beta'')$ *are nonsingular uniformly in* $\alpha', \alpha'' \in \Theta_\alpha$ *and* $\beta', \beta'' \in \Theta_\beta$. *Then, both Assumptions 2.4 and 2.5 hold true.*

Proof. It is obvious that Assumption 2.5 follows. The mean-value theorem applied to (6) and (7) leads to $\mathbb{G}_\infty^\alpha(\theta) = \bar{A}(\alpha, \tilde{\alpha}, \beta)[\alpha_0 - \alpha]$ for some $\tilde{\alpha}$ lying the segment connecting α and α_0, with a similar form for $\mathbb{G}_\infty^\beta(\beta)$; recall that Θ_α and Θ_β are presupposed to be convex. Since $\inf_{\alpha',\alpha'',\beta'}\left||\bar{A}(\alpha', \alpha'', \beta')|\right| > 0$ and $\inf_{\beta',\beta''}\left||\bar{B}(\beta', \beta'')|\right| > 0$ under the assumption, the matrices $\bar{A}^{\otimes 2}$ and $\bar{B}^{\otimes 2}$ are uniformly positive definite, whence Assumption 2.4 follows. $\square$

2.2 Main result

In what follows, we write

$$\Delta_j Y = Y_{t_j} - Y_{t_{j-1}}$$

for any process Y, and

$$f_{j-1}(a) = f(X_{t_{j-1}}, a)$$

for a variable a in some set A and a measurable function f on $\mathbb{R}^d \times A$. The Euler approximation for SDE (1) is formally

$$X_{t_j} \approx X_{t_{j-1}} + a_{j-1}(\alpha)h_n + b_{j-1}(\beta)\Delta_j w + c_{j-1}(\beta)\Delta_j J$$

under P_θ, which leads us to consider the local-Gauss distribution approximation

$$\mathcal{L}(X_{t_j}|X_{t_{j-1}}) \approx \mathcal{N}_d\left(X_{t_{j-1}} + a_{j-1}(\alpha)h_n, h_n V_{j-1}(\beta)\right). \tag{10}$$

Put

$$\chi_j(\alpha) = \Delta_j X - h_n a_{j-1}(\alpha).$$

Based on (10), we define our GQL by

$$\mathbb{Q}_n(\theta) = -\sum_{j=1}^{n} \left\{ \log |V_{j-1}(\beta)| + \frac{1}{h_n} V_{j-1}^{-1}(\beta) \left[\chi_j(\alpha)^{\otimes 2} \right] \right\}, \tag{11}$$

and the corresponding GQMLE by any element

$$\hat{\theta}_n = (\hat{\alpha}_n, \hat{\beta}_n) \in \underset{\theta \in \Theta^-}{\operatorname{argmax}}\, \mathbb{Q}_n(\theta),$$

where Θ^- denotes the closure of Θ.

Under Assumption 2.1 we have $\int z^{(k)} z^{(l)} \nu(dz) = \delta_{kl}$ for $k, l \in \{1, \dots, r''\}$. We need some further notation in this direction. For $i_1, \dots, i_m \in \{1, \dots, r''\}$ with $m \geq 3$, we write $\nu(m)$ for the mth mixed moments of ν:

$$\nu(m) = \{\nu_{i_1 \dots i_m}(m)\}_{i_1, \dots, i_m} := \left\{ \int z^{(i_1)} \dots z^{(i_m)} \nu(dz) \right\}_{i_1, \dots, i_m}.$$

Let $c^{(\cdot k)}(x, \beta) \in \mathbb{R}^d$ denote the kth column of $c(x, \beta)$. We introduce the matrix

$$\mathbb{V}(\theta_0) := \begin{pmatrix} \mathbb{G}'^{\alpha}_{\infty}(\theta_0) & \mathbb{V}_{\alpha\beta} \\ \mathbb{V}_{\alpha\beta}^{\top} & \mathbb{V}_{\beta\beta} \end{pmatrix}, \tag{12}$$

where, for each $v' \in \mathbb{R}^{p_\alpha}$ and $v''_1, v''_2 \in \mathbb{R}^{p_\beta}$,

$$\begin{aligned} \mathbb{V}_{\alpha\beta}[v', v''_1] &:= -\int \sum_{k', l', s'} \nu_{k' l' s'}(3) V^{-1}(x, \beta_0) \left[\partial_\alpha a(x, \alpha_0)[v'], c^{(\cdot s')}(x, \beta_0) \right] \\ &\quad \cdot \left\{ \partial_\beta V^{-1}(x, \beta) \right\} [v''_1, c^{(\cdot k')}(x, \beta_0), c^{(\cdot l')}(x, \beta_0)] \pi_0(dx), \\ \mathbb{V}_{\beta\beta}[v''_1, v''_2] &:= \int \sum_{s, t, s', t'} \nu_{s t s' t'}(4) \left\{ \partial_\beta V^{-1}(x, \beta_0)[v''_1, c^{(\cdot s)}(x, \beta_0), c^{(\cdot t)}(x, \beta_0)] \right\} \\ &\quad \cdot \left\{ \partial_\beta V^{-1}(x, \beta_0)[v''_2, c^{(\cdot s')}(x, \beta_0), c^{(\cdot t')}(x, \beta_0)] \right\} \pi_0(dx). \end{aligned}$$

Finally, put

$$\Sigma_0 = \begin{pmatrix} (\mathbb{G}'^{\alpha}_{\infty})^{-1}(\theta_0) & (\mathbb{G}'^{\alpha}_{\infty})^{-1} \mathbb{V}_{\alpha\beta} (\mathbb{G}'^{\beta}_{\infty})^{-1}(\theta_0) \\ \text{Sym.} & (\mathbb{G}'^{\beta}_{\infty})^{-1} \mathbb{V}_{\beta\beta} (\mathbb{G}'^{\beta}_{\infty})^{-1}(\theta_0) \end{pmatrix}.$$

Now we can state our main result, the proof of which is deferred to Section 4.1.

Theorem 2.7. *Suppose the conditions 2.1, 2.2, 2.3, 2.4, and 2.5. Then we have*

$$E_0 \left[f \left(\sqrt{T_n} (\hat{\theta}_n - \theta_0) \right) \right] \to \int f(u) \phi(u; 0, \Sigma_0)\, du, \quad n \to \infty,$$

for every continuous function $f : \mathbb{R}^p \to \mathbb{R}$ of at most polynomial growth.

We immediately see that $\hat{\alpha}_n$ and $\hat{\beta}_n$ are asymptotically independent if $\nu(3) = 0$, implying that $\hat{\alpha}_n$ and $\hat{\beta}_n$ may not be asymptotically independent if ν is skewed. If $c \equiv 0$ so that X is a diffusion, then $\nu(4) = 0$, so that $\mathbb{V}_{\beta\beta} = 0$ and $\sqrt{T_n}(\hat{\beta}_n - \beta_0)$ is asymptotically degenerate at 0. This is in accordance with the case of diffusion, where the GQMLE of β is $\sqrt{n}$-consistent.

In order to construct confidence regions for θ_0 as well as to perform statistical tests, we need a consistent estimator of the asymptotic covariance matrix Σ_0. Although Σ_0 contains unknown third and fourth mixed moments of ν, It turns out to be possible to provide a consistent estimator of Σ_0 without any specific knowledge of ν other than Assumption 2.1. Let

$$\hat{\Sigma}_n = \begin{pmatrix} (\hat{\mathbb{G}}'^{\alpha}_n)^{-1} & (\hat{\mathbb{G}}'^{\alpha}_n)^{-1} \hat{\mathbb{V}}_{\alpha\beta,n} (\hat{\mathbb{G}}'^{\beta}_n)^{-1} \\ \text{Sym.} & (\hat{\mathbb{G}}'^{\beta}_n)^{-1} \hat{\mathbb{V}}_{\beta\beta,n} (\hat{\mathbb{G}}'^{\beta}_n)^{-1} \end{pmatrix}.$$

where, for each $v'_1, v'_2 \in \mathbb{R}^{p_\alpha}$ and $v''_1, v''_2 \in \mathbb{R}^{p_\beta}$,

$$\hat{\mathbb{G}}'^{\alpha}_n[v'_1, v'_2] := -\frac{1}{n}\sum_{j=1}^{n} V_{j-1}^{-1}(\hat{\beta}_n)\left[\partial_\alpha a_{j-1}(\hat{\alpha}_n)[v'_1], \partial_\alpha a_{j-1}(\hat{\alpha}_n)[v'_2]\right],$$

$$\hat{\mathbb{G}}'^{\beta}_n[v''_1, v''_2] := -\frac{1}{n}\sum_{j=1}^{n} \mathrm{trace}\left\{\left(V_{j-1}^{-1}(\partial_\beta V_{j-1})V_{j-1}^{-1}(\partial_\beta V_{j-1})\right)(\hat{\beta}_n)[v''_1, v''_2]\right\},$$

$$\hat{\mathbb{V}}_{\alpha\beta,n}[v'_1, v''_1] := -\sum_{j=1}^{n}\frac{1}{T_n}\left(V_{j-1}^{-1}\partial_\beta V_{j-1}^{-1}\right)(\hat{\beta}_n)[\partial_\alpha a_{j-1}(\hat{\alpha}_n)[v'_1], \chi_j(\hat{\alpha}_n), v''_2, \chi_j(\hat{\alpha}_n)^{\otimes 2}],$$

$$\hat{\mathbb{V}}_{\beta\beta,n}[v''_1, v''_2] := \sum_{j=1}^{n}\frac{1}{T_n}\left(\partial_\beta V_{j-1}^{-1}\partial_\beta V_{j-1}^{-1}\right)(\hat{\beta}_n)[v''_1, \chi_j(\hat{\alpha}_n)^{\otimes 2}, v''_2, \chi_j(\hat{\alpha}_n)^{\otimes 2}].$$

Corollary 2.8. *Under the conditions of Theorem 2.7, we have $\hat{\Sigma}_n \to^p \Sigma_0$, so that the weak convergence*

$$\hat{\Sigma}_n^{-1/2}\sqrt{T_n}(\hat{\theta}_n - \theta_0) \to^{\mathcal{L}} \mathcal{N}_p(0, I_p)$$

holds true.

The proof of Corollary 2.8 is given in Section 4.2.

2.3 Remarks

1. The revealed convergence rate $\sqrt{T_n}$ of the GQMLE $\hat{\beta}_n$ alerts someone to take precautions against the presence of jumps. For instance, suppose that one has adopted the parametric diffusion model (i.e. (1) with $c \equiv 0$) although there actually does exist a nonnull jump part. Then, he/she takes $\sqrt{n}$ for the convergence rate of $\hat{\beta}_n$, although the true one is $\sqrt{T_n}$, which may lead to a seriously inappropriate confidence zone. This point can be sufficient grounds for importance of testing the presence of jumps. In case of one-dimensional X, Masuda [27, Section 4] constructed an analogue to Jarque-Bera normality test and studied its asymptotic behavior. We will report a multivariate extension of the result in a subsequent paper.

2. Here are some remarks concerning the stability condition. The exponential mixing property can be drastically relaxed if we are only interested in the asymptotic normality of $\hat{\theta}_n$. Then, the mere ergodicity (i.e. $\|P_t(x,\cdot) - \pi_0(\cdot)\| \to 0$ as $t \to \infty$ for every x, without knowledge of its decreasing rate) suffices, and we could exactly proceed as in Masuda [27, Theorem 3.4]. Although the paper just cited dealt with one-dimensional X, a multivariate extension in deducing the asymptotic normality would be straightforward. See Masuda [27, p.115] for some discussion concerning a relation between the uniform (in time) moment estimates and the boundedness of the coefficients. See also Remark 5.4.

3. Our identifiability condition Assumption 2.4 can be stringent; for example, it excludes the case where $b(x,\beta) = \beta'$ and $c(x,\beta) = \beta''$, so that there exist infinitely many $\beta = (\beta', \beta'')$ such that $V(x,\beta) - V(x,\beta_0) = 0$ for every x. However, this is unavoidable as our contrast function $\mathbb{M}_n$ is constructed solely based on fitting local conditional mean and covariance matrix. Roughly speaking, in order to estimate both parameters separately contained in b and c, we need "distinct" nonlinearities in x for $b(x,\beta)$ and $c(x,\beta)$ to fulfil Assumption 2.4.

4. Although we are considering "ergodic" X, it is obvious that we can target Lévy processes as well, according to the built-in independence of the increments $(\Delta_j X)_{j \le n}$.

5. A general form of the martingale estimating functions is

$$\theta \mapsto \sum_{j=1}^{n} W_{j-1}(\theta)\left\{g(X_{t_{j-1}}, X_{t_j}; \theta) - E_\theta^{j-1}[g(X_{t_{j-1}}, X_{t_j}; \theta)]\right\}$$

for some $W \in \mathbb{R}^p \otimes \mathbb{R}^m$ and $\mathbb{R}^m$-valued function g on $\mathbb{R}^d \times \mathbb{R}^d \times \Theta$. We would have a wide choice of W and g. When the conditional expectations involved do not admit closed forms, then the leading-term approximation of them via the Itô-Taylor expansion can be used. In view of this, as in Kessler [15], it would be formally possible to relax the condition $nh_n^2 \to 0$ in (2) by gaining the order of the Ito-Taylor expansions of the conditional mean and conditional covariance:

$$E_\theta[X_{t_j}|\mathcal{F}_{t_{j-1}}] = X_{t_{j-1}} + a_{j-1}(\alpha)h_n + \cdots,$$
$$V_\theta[X_{t_j}|\mathcal{F}_{t_{j-1}}] = V_{j-1}(\beta)h_n + \cdots,$$

which we have implicitly used up to the h_n-order terms to build $\mathbb{Q}_n$ of (11). However, we then need specific moment structures of ν, which appear in the higher orders of the above Itô-Taylor expansion. Moreover, we should note that the convergence rate $\sqrt{T_n}$ can be never improved for both α and β even if $E_\theta[X_{t_j}|\mathcal{F}_{t_{j-1}}]$ and $V_\theta[X_{t_j}|\mathcal{F}_{t_{j-1}}]$ have closed forms, such as the case of linear drifts, so that the rate of $h_n \to 0$ may not matter as long as $T_n \to \infty$.

2.4 A numerical example

For simulation purpose, we consider the following concrete model:

$$dX_t = \frac{-\alpha X_t}{\sqrt{1+X_t^2}}dt + \sqrt{\beta}dJ_t, \quad X_0 = 0, \tag{13}$$

where the true value is $(\alpha_0, \beta_0) = (1,1)$, the driving process is the normal inverse Gaussian Lévy process such that $\mathcal{L}(J_t) = NIG(\delta, 0, \delta t, 0)$, $\delta > 0$. It holds that $E[J_t] = 0$, $E[J_1^2] = t$, and $\mathcal{L}(J_t) \to \mathcal{N}(0,1)$ in total variation as $\delta \to \infty$, and that $\nu(3) = 0$ and $\nu(4) = 3/\delta^2$. The model (13) is a normal-inverse Gaussian counterpart to the hyperbolic diffusion, for which J is replaced by a standard Wiener process. For this X, we can verify all the assumptions.

We simulated 1000 independent paths by Euler scheme with sufficiently fine stepsize to obtain 1000 independent estimates $\hat{\theta}_n = (\hat{\alpha}_n, \hat{\alpha}_n)$, and then computed their empirical mean and standard deviations. Table 1 reports the results; just for comparison, we included the case of diffusion, where Z is a standard Wiener process. From the table, we can observe the following:

- The performance of $\hat{\alpha}_n$ are rather similar for all the three cases;
- The performance of $\hat{\beta}_n$ gets better for larger δ, which can be expected from the fact that the asymptotic variance of $\hat{\beta}_n$ is a constant multiple of $\nu(4) = 3\delta^{-2}$.

T_n	h_n	Diffusion		$\delta = 1$		$\delta = 10$	
		α	β	α	β	α	β
10	0.05	1.158	0.964	1.149	0.982	1.179	0.965
		(0.629)	(0.100)	(0.621)	(0.576)	(0.645)	(0.113)
	0.01	1.193	0.993	1.172	0.968	1.210	0.993
		(0.673)	(0.044)	(0.635)	(0.476)	(0.658)	(0.069)
100	0.05	0.999	0.970	0.997	0.976	0.996	0.971
		(0.184)	(0.031)	(0.188)	(0.174)	(0.177)	(0.035)
	0.01	1.018	0.994	1.017	0.996	1.017	0.994
		(0.184)	(0.014)	(0.189)	(0.174)	(0.181)	(0.022)

Table 1: Finite sample performance of $\hat{\theta}_n$ concerning the model (13); just for comparison, the case of diffusion is also included.

3 Mighty convergence of a class of continuous random fields

In this section, we will prove a fundamental result concerning the "single-norming" mighty convergence of a continuous statistical random fields associated with general vector-valued estimating functions; here, the "single-norming" means that the rates of convergence are the same for all the arguments of the corresponding estimator. Theorem 3.5 below will serve as a fundamental tool in the proof of Theorem 2.7. The content of this section can be read independently of the main body of this paper.

To proceed, we need some notation. Denote by $\{\mathcal{X}_n, \mathcal{A}_n, (P_\theta)_{\theta\in\Theta}\}_{n\in\mathbb{N}}$ underlying statistical experiments, where $\Theta \subset \mathbb{R}^p$ is a bounded convex domain. Let $\theta_0 \in \Theta$, and write $P_0 = P_{\theta_0}$. Let $\mathbb{G}_n = (\mathbb{G}_{j,n})_{j=1}^p : \mathcal{X}_n \times \Theta \to \mathbb{R}^p$ be vector-valued random functions; as usual, we will simply write $\mathbb{G}_n(\theta)$, dropping the argument of $\mathcal{X}_n$. Our target "contrast" function is

$$\mathbb{M}_n(\theta) := -\frac{1}{T_n}|\mathbb{G}_n(\theta)|^2, \tag{14}$$

where (T_n) is a nonrandom positive real sequence such that $T_n \to \infty$. The corresponding "M-estimator" is defined to be any measurable mapping $\hat{\theta}_n : \mathcal{X}_n \to \Theta^-$ such that

$$\hat{\theta}_n \in \operatorname*{argmax}_{\theta\in\Theta^-} \mathbb{M}_n(\theta).$$

Due to the compactness of Θ^- and the continuity of $\mathbb{M}_n$ imposed later on, we can always find such a $\hat{\theta}_n$. The estimate $\hat{\theta}_n$ can be any root of $\mathbb{G}_n(\theta) = 0$ as soon as it exists.

Set $U_n(\theta_0) := \{u \in \mathbb{R}^p : \theta_0 + T_n^{-1/2}u \in \Theta\}$ and define random fields $\mathbb{Z}_n : U_n(\theta_0) \to (0, \infty)$ by

$$\mathbb{Z}_n(u) = \mathbb{Z}_n(u;\theta_0) := \exp\left\{\mathbb{M}_n(\theta_0 + T_n^{-1/2}u) - \mathbb{M}_n(\theta_0)\right\}. \tag{15}$$

Obviously, it holds that

$$\hat{u}_n := \sqrt{T_n}(\hat{\theta}_n - \theta_0) \in \operatorname*{argmax}_{\theta\in\Theta^-} \mathbb{Z}_n(\theta).$$

We consider the following two conditions for the random fields $\mathbb{Z}_n$.

- *(Polynomial type Large Deviation Inequality (PLDI))* For every $M > 0$, we have

$$\sup_{r>0}\left\{r^M \sup_{n\in\mathbb{N}} P_0\left[\sup_{|u|>r} \mathbb{Z}_n(u) \ge e^{-r}\right]\right\} < \infty. \tag{16}$$

- *(Weak convergence on compact sets)* There exists a random field $\mathbb{Z}_0(\cdot) = \mathbb{Z}_0(\cdot;\theta_0)$ such that $\mathbb{Z}_n \to^{\mathcal{L}} \mathbb{Z}_0$ in $\mathcal{C}(B^-(R))$ for each $R > 0$, where $B^-(R) := \{u \in \mathbb{R}^p; |u| \ge R\}$.

Under these conditions, the mode of convergence of $\mathbb{Z}_n(\cdot)$ is mighty enough to deduce that the maximum-point sequence $(\hat{u}_n)_n$ is $L^q(P_0)$-bounded for every $q > 0$, which especially implies that $(\hat{u}_n)_n$ is tight: indeed, if (16) is in force,

$$\sup_{n\in\mathbb{N}} P_0[|\hat{u}_n| > r] \le \sup_{n\in\mathbb{N}} P_0\left[\sup_{|u|>r} \mathbb{Z}_n(u) \ge \mathbb{Z}_n(0)\right] = \sup_{n\in\mathbb{N}} P_0\left[\sup_{|u|>r} \mathbb{Z}_n(u) \ge 1\right] \lesssim \frac{1}{r^M}$$

for every $r > 0$, so that

$$\sup_{n\in\mathbb{N}} E_0[|\hat{u}_n|^q] = \int_0^\infty \sup_{n\in\mathbb{N}} P_0[|\hat{u}_n| > s^{1/q}]ds \lesssim 1 + \int_1^\infty s^{-M/q}ds < \infty.$$

If $u \mapsto \mathbb{Z}_0(u)$ is a.s. maximized at a unique point $\hat{u}_\infty$, then it follows from the tightness of $(\hat{u}_n)_{n\in\mathbb{N}}$ that $\hat{u}_n \to^{\mathcal{L}} \hat{u}_\infty$; let us remind that the weak convergence on any compact set alone is not enough to deduce the weak convergence of $\hat{u}_n$, since $U_n(\theta_0) \uparrow \mathbb{R}^p$ and we have no guarantee that $(\hat{u}_n)$ is tight. Moreover, owing to the PLDI, the moment of $f(\hat{u}_n)$ converges to that of $f(\hat{u}_\infty)$ for every continuous function f on $\mathbb{R}^p$ of at most polynomial growth. In our framework, $\log \mathbb{Z}_0$ admits a quadratic structure with a normally distributed linear term and a nonrandom positive definite quadratic term, so that $\hat{u}_\infty$ is asymptotically normally distributed.

We now introduce regularity conditions.

Assumption 3.1 (Smoothness)**.** *The functions $\theta \mapsto \mathbb{G}_n(\theta)$ are continuously extended to the boundary of Θ, and belong to $\mathcal{C}^3(\Theta)$, P_0-a.s.*

Assumption 3.2 (Bounded moments)**.** *For every $K > 0$,*

$$\sup_{n\in\mathbb{N}} E_0\left[\left|\frac{1}{\sqrt{T_n}}\mathbb{G}_n(\theta_0)\right|^K\right] + \sum_{k=0}^{3}\sup_{n\in\mathbb{N}} E_0\left[\sup_{\theta\in\Theta}\left|\frac{1}{T_n}\partial_\theta^k\mathbb{G}_n(\theta)\right|^K\right] < \infty.$$

Let $M > 0$ be a given constant.

Assumption 3.3 (Limits)**.** *(a) There exist a nonrandom function $\mathbb{G}_\infty : \Theta \to \mathbb{R}^p$ and positive constants $\chi = \chi(\theta_0)$ and ϵ such that: $\mathbb{G}_\infty(\theta_0) = 0$; $\sup_\theta|\mathbb{G}_\infty(\theta)| < \infty$; $|\mathbb{G}_\infty(\theta)|^2 \geq \chi|\theta-\theta_0|^2$ for every $\theta \in \Theta$; and*

$$\sup_{n\in\mathbb{N}} E_0\left[\sup_{\theta\in\Theta}\left|\sqrt{T_n}\left(\frac{1}{T_n}\mathbb{G}_n(\theta) - \mathbb{G}_\infty(\theta)\right)\right|^{M+\epsilon}\right] < \infty.$$

(b) There exists a nonrandom $\mathbb{G}'_\infty(\theta_0) \in \mathbb{R}^p \otimes \mathbb{R}^p$ of rank p such that

$$\sup_{n\in\mathbb{N}} E_0\left[\left|\sqrt{T_n}\left(\frac{1}{T_n}\partial_\theta\mathbb{G}_n(\theta_0) - \mathbb{G}'_\infty(\theta_0)\right)\right|^M\right] < \infty.$$

Assumption 3.4 (Weak convergence)**.** *$T_n^{-1/2}\mathbb{G}_n(\theta_0) \to^{\mathcal{L}} \mathcal{N}_p(0, \mathbb{V}(\theta_0))$ for some positive definite $\mathbb{V}(\theta_0) \in \mathbb{R}^p \otimes \mathbb{R}^p$.*

Let $\Sigma(\theta_0) := (\mathbb{G}'_\infty)^{-1}\mathbb{V}(\mathbb{G}'_\infty)^{-1\top}(\theta_0)$, and denote by $\phi(u; 0, \Sigma(\theta_0))$ the centered Gaussian density with covariance matrix $\Sigma(\theta_0)$. The main claim of this section is the following.

Theorem 3.5. *Let $M > 0$.*

(a) Suppose that Assumptions 3.1, 3.2, and 3.3. Then the PLDI (16) holds true.

(b) If Assumption 3.4 is additionally met, then

$$E_0[f(\hat{u}_n)] \to \int f(u)\phi\left(u; 0, \Sigma(\theta_0)\right)du$$

for every continuous function $f : \mathbb{R}^p \to \mathbb{R}$ satisfying that $\limsup_{|u|\to\infty}|u|^{-q}|f(u)| < \infty$ for some $q \in (0, M)$.

Proof. Applying Taylor expansion to (15), we get

$$\log\mathbb{Z}_n(u) = \Delta_n(\theta_0)[u] - \frac{1}{2}\Gamma(\theta_0)[u,u] + \xi_n(u), \tag{17}$$

where $\Delta_n(\theta_0) := T_n^{-1/2}\partial_\theta\mathbb{M}_n(\theta_0)$, $\Gamma_n(\theta_0) := -T_n^{-1}\partial_\theta^2\mathbb{M}_n(\theta_0)$, $\Gamma(\theta_0) := 2\mathbb{G}'_\infty(\theta_0)^\top\mathbb{G}'_\infty(\theta_0)$, and

$$\xi_n(u) := \frac{1}{2}\{\Gamma(\theta_0) - \Gamma_n(\theta_0)\}[u,u] - \int_0^1(1-s)\int\partial_\theta\Gamma_n(\theta_0 + stT_n^{-1/2}u)[sT_n^{-1/2}u, u^{\otimes 2}]dtds. \tag{18}$$

We will prove (a) by making use of Yoshida [41, Theorem 3(c)]. We will verify the conditions [A1″], [A4′], [A6], [B1] and [B2] of that paper, omitting the full description of the conditions. Assumption 3.3 assures [B1] (the positive definiteness of $\Gamma(\theta_0)$) and the convergence $T_n^{-1}\mathbb{M}_n(\theta) \to^p -|\mathbb{G}_\infty(\theta)|^2$ for each $\theta \in \Theta$. Let

$$\mathbb{Y}_n(\theta) := \frac{1}{T_n}\{\mathbb{M}_n(\theta) - \mathbb{M}_n(\theta_0)\} = \frac{1}{T_n}\log\mathbb{Z}_n\left(\sqrt{T_n}(\theta-\theta_0)\right),$$

then $\mathbb{Y}_n(\theta) \to^p -|\mathbb{G}_\infty(\theta)|^2 =: \mathbb{Y}(\theta)$. Obviously $\mathbb{Y}(\theta) \leq -\chi^2|\theta-\theta_0|^2$ for each $\theta \in \Theta$, verifying [B2] (the identifiability). Next, we will verify the conditions [A1″] and [A6] in the following form:

[A1″] (i) $\sup_n E_0\left[\sup_\theta|T_n^{-1}\partial_\theta^3\mathbb{M}_n(\theta)|^K\right] < \infty$ for every $K > 0$.

(ii) $\sup_n E_0\left[|\sqrt{T_n}(\Gamma_n(\theta_0)-\Gamma(\theta_0)|^{M-\epsilon_1}\right] < \infty$ for every $\epsilon_1 > 0$ small enough.

[A6] (i) $\sup_n E_0\left[|\Delta_n(\theta_0)|^K\right] < \infty$ for every $K > 0$.

(ii) $\sup_n E_0\left[\sup_\theta |\sqrt{T_n}(\mathbb{Y}_n(\theta)-\mathbb{Y}(\theta))|^{M+\epsilon/2}\right] < \infty$, for ϵ given in Assumption 3.3.

At this stage, we should remark that the remaining [A4′] is in force under [A1″] and [A6]: here, [A4′] is the "tuning-parameter controlling" condition concerned with the moment-order indices $M_1,\dots,M_4$ in the notation of Yoshida [41]. We only give a sketch of verification of [A4′]. The indices corresponding to M_1 and M_3 can be taken arbitrarily large under Assumption 3.2, so it suffices to look at M_2 and M_4. It turns out that, by taking the tuning parameters of Yoshida [41] as $\rho_1,\rho_2,\alpha \approx 0$ and $\beta_2 = 0$ and then $\beta_1 \approx 1/2$ with $\rho = 2$, we can pick a constant $\delta \in (0,\epsilon/2)$ small enough so that [A4′] follows with $M_2 = M+\delta$ and $M_4 = M-\delta$. Building on these observations, we are left to proving [A1″] and [A6] above.

We begin with [A1″]. Since $|T_n^{-1}\partial_\theta^3\mathbb{M}_n(\theta)| \lesssim |T_n^{-1}\mathbb{G}_n(\theta)||T_n^{-1}\partial_\theta^3\mathbb{G}_n(\theta)|+|T_n^{-1}\partial_\theta\mathbb{G}_n(\theta)||T_n^{-1}\partial_\theta^2\mathbb{G}_n(\theta)|$, we have for every $K>0$

$$\sup_{n\in\mathbb{N}} E_0\left[\sup_{\theta\in\Theta}\left|\frac{1}{T_n}\partial_\theta^3\mathbb{M}_n(\theta)\right|^K\right] < \infty.$$

Noting that $\partial_{\theta_i}\partial_{\theta_j}\mathbb{M}_n = -2T_n^{-1}\{\partial_{\theta_i}\partial_{\theta_j}\mathbb{G}_n[\mathbb{G}_n] + \partial_{\theta_i}\mathbb{G}_n[\partial_{\theta_j}\mathbb{G}_n]\}$, we also have

$$\begin{aligned}\sqrt{T_n}|\Gamma_n(\theta_0)-\Gamma(\theta_0)| \lesssim{}& \left|\frac{1}{\sqrt{T_n}}\mathbb{G}_n(\theta_0)\right|\left|\frac{1}{T_n}\partial_\theta^2\mathbb{G}_n(\theta_0)\right| \\ &+ \left(|\Gamma(\theta_0)| + \left|\frac{1}{T_n}\partial_\theta\mathbb{G}_n(\theta_0)\right|\right)\left|\sqrt{T_n}\left(\frac{1}{T_n}\partial_\theta\mathbb{G}_n(\theta_0)-\mathbb{G}'_\infty(\theta_0)\right)\right|.\end{aligned}$$

Therefore, Assumptions 3.2 and 3.3 combined with Hölder's inequality yield that for $\epsilon_1 \in (0,M)$

$$\begin{aligned}&\sup_{n\in\mathbb{N}} E_0\left[\left|\sqrt{T_n}(\Gamma_n(\theta_0)-\Gamma(\theta_0)\right|^{M-\epsilon_1}\right] \\ &\lesssim 1 + \left\{\sup_{n\in\mathbb{N}} E_0\left[\left|\sqrt{T_n}\left(\frac{1}{T_n}\partial_\theta\mathbb{G}_n(\theta_0)-\mathbb{G}'_\infty(\theta_0)\right)\right|^M\right]\right\}^{(M-\epsilon_1)/M} < \infty.\end{aligned}$$

Thus [A1″] follows.

Next we prove [A6]. The statement (i) is obvious from Assumption 3.2:

$$\sup_{n\in\mathbb{N}} E_0\left[|\Delta_n(\theta_0)|^K\right] \lesssim \sup_{n\in\mathbb{N}} E_0\left[\left|\frac{1}{T_n}\partial_\theta\mathbb{G}_n(\theta_0)\right|^K\left|\frac{1}{\sqrt{T_n}}\mathbb{G}_n(\theta_0)\right|^K\right] < \infty. \tag{19}$$

Using the estimate

$$\left|\sqrt{T_n}(\mathbb{Y}_n(\theta)-\mathbb{Y}(\theta))\right| \le \frac{1}{\sqrt{T_n}}\left|\frac{1}{\sqrt{T_n}}\mathbb{G}_n(\theta_0)\right|^2 + \left(|\mathbb{G}_\infty(\theta)| + \left|\frac{1}{T_n}\mathbb{G}_n(\theta)\right|\right)\left|\sqrt{T_n}\left(\frac{1}{T_n}\mathbb{G}_n(\theta)-\mathbb{G}_\infty(\theta)\right)\right|,$$

it follows under Assumptions 3.2 and 3.3 that

$$\begin{aligned}&\sup_{n\in\mathbb{N}} E_0\left[\sup_{\theta\in\Theta}\left|\sqrt{T_n}(\mathbb{Y}_n(\theta)-\mathbb{Y}(\theta))\right|^{M+\epsilon/2}\right] \\ &\lesssim 1 + \sup_{n\in\mathbb{N}} E_0\left[\sup_{\theta\in\Theta}\left|\sqrt{T_n}\left(\frac{1}{T_n}\mathbb{G}_n(\theta)-\mathbb{G}_\infty(\theta)\right)\right|^{M+\epsilon}\right]^{(M+\epsilon/2)/(M+\epsilon)} < \infty.\end{aligned}$$

Thus [A6] is ensured, and the proof of (a) is complete.

We now turn to the proof of (b). Fix any $R>0$. Since we know that the sequence $(\hat{u}_n)$ is $L^q(P_0)$-bounded for each $q \in (0,M)$ and that the set $\mathrm{argmax}_u \log\mathbb{Z}_\infty(u)$ a.s. consists of the only point

$$\hat{u}_\infty := \Gamma(\theta_0)^{-1}\Delta_\infty(\theta_0) \sim \mathcal{N}_p(0,\Sigma(\theta_0)),$$

it suffices to show that $\log \mathbb{Z}_n \to^{\mathcal{L}} \log \mathbb{Z}_\infty$ in $\mathcal{C}(B^-(R))$, where

$$\log \mathbb{Z}_\infty(u) := \Delta_\infty(\theta_0)[u] - \frac{1}{2}\Gamma(\theta_0)[u,u], \quad \Delta_\infty(\theta_0) \sim \mathcal{N}_p\left(0, 4\mathbb{G}'_\infty(\theta_0)^\top \mathbb{V}(\theta_0)\mathbb{G}'_\infty(\theta_0)\right)$$

(e.g. Yoshida [41, Theorem 5]). We have $T_n^{-1}\partial_\theta \mathbb{G}_n(\theta_0) \to^p \mathbb{G}'_\infty(\theta_0)$ from Assumption 3.3, hence Slutsky's lemma and Assumption 3.4 imply that

$$\Delta_n(\theta_0) = -\frac{2}{T_n}\partial_\theta \mathbb{G}_n(\theta_0)\left[\frac{1}{\sqrt{T_n}}\mathbb{G}_n(\theta_0)\right] \to^{\mathcal{L}} \Delta_\infty(\theta_0).$$

Also, we have

$$|\xi_n(u)| \lesssim |u|^2|\Gamma_n(\theta_0) - \Gamma(\theta_0)| + \frac{|u|^3}{\sqrt{T_n}}\sup_{\theta\in\Theta}\left|\frac{1}{T_n}\partial_\theta^3 \mathbb{M}_n(\theta)\right| = o_p(1) \tag{20}$$

for every $u \in B^-(R)$. Thus, recalling the expression (17), we get $\log \mathbb{Z}_n(u) \to^{\mathcal{L}} \log \mathbb{Z}_0(u)$ for every $u \in B^-(R)$, and moreover, due to the linearity in u of the weak convergence term $\Delta_n(\theta_0)[u]$, the Cramér-Wold device ensures the finite-dimensional convergence. Therefore, it remains to check the tightness of $\{\log \mathbb{Z}_n(u)\}_{u\in B^-(R)}$. In view of the classical Kolmogorov's tightness criterion for continuous random fields (e.g. Kunita [18, Theorem 1.4.7]), it suffices to show that there exists a constant $\gamma > p(= \dim \Theta)$ such that

$$\sup_{u\in B^-(R)}\sup_{n\in\mathbb{N}} E_0\left[|\log \mathbb{Z}_n(u)|^\gamma\right] + \sup_{n\in\mathbb{N}} E_0\left[\sup_{|u|\le R}|\partial_u \log \mathbb{Z}_n(u)|^\gamma\right] < \infty. \tag{21}$$

In view of the estimates in (19) and (20) as well as the expressions (17) and (18),

$$\begin{aligned}\sup_{u\in B^-(R)}\sup_{n\in\mathbb{N}} E_0\left[|\log \mathbb{Z}_n(u)|^\gamma\right] &\lesssim \sup_{n\in\mathbb{N}} E_0\left[|\Delta_n(\theta_0)|^\gamma\right] + 1 + \sup_{u\in B^-(R)}\sup_{n\in\mathbb{N}} E_0\left[|\xi_n(u)|^\gamma\right] \\ &\lesssim 1 + E_0\left[|\Gamma_n(\theta_0) - \Gamma(\theta_0)|^\gamma\right] + \sup_{n\in\mathbb{N}} E_0\left[\sup_{\theta\in\Theta}\left|\frac{1}{T_n}\partial_\theta^3 \mathbb{M}_n(\theta)\right|^\gamma\right] < \infty.\end{aligned}$$

Furthermore, since

$$\begin{aligned}\partial_u \log \mathbb{Z}_n(u) &= \partial_u\left\{\mathbb{M}_n\left(\theta_0 + \frac{1}{\sqrt{T_n}}u\right) - \mathbb{M}_n(\theta_0)\right\} \\ &= \frac{1}{\sqrt{T_n}}\partial_\theta \mathbb{M}_n\left(\theta_0 + \frac{1}{\sqrt{T_n}}u\right) \\ &= \frac{1}{\sqrt{T_n}}\left\{\partial_\theta \mathbb{M}_n(\theta_0) + \frac{1}{\sqrt{T_n}}\int_0^1 \partial_\theta^2 \mathbb{M}_n\left(\theta_0 + \frac{s}{\sqrt{T_n}}u\right)[u]ds\right\},\end{aligned}$$

the finiteness of $\sup_n E_0\left[\sup_{|u|\le R}|\partial_u \log \mathbb{Z}_n(u)|^\gamma\right]$ follows on applying Assumption 3.2 to the estimate

$$\begin{aligned}\sup_{|u|\le R}|\partial_u \log \mathbb{Z}_n(u)| &\lesssim \left|\frac{1}{\sqrt{T_n}}\mathbb{G}_n(\theta_0)\right|\left|\frac{1}{T_n}\partial_\theta \mathbb{G}_n(\theta_0)\right| + \sup_{\theta\in\Theta}\left|\frac{1}{T_n}\partial_\theta^2 \mathbb{M}_n(\theta)\right| \\ &\lesssim \left|\frac{1}{\sqrt{T_n}}\mathbb{G}_n(\theta_0)\right|\left|\frac{1}{T_n}\partial_\theta \mathbb{G}_n(\theta_0)\right| + \sup_{\theta\in\Theta}\left\{\left|\frac{1}{T_n}\mathbb{G}_n(\theta)\right|\left|\frac{1}{T_n}\partial_\theta^2 \mathbb{G}_n(\theta)\right| + \left|\frac{1}{T_n}\partial_\theta \mathbb{G}_n(\theta)\right|^2\right\}.\end{aligned}$$

Thus we have obtained (21), thereby achieving the proof of (b). □

Remark 3.6. *We have confined ourselves to the "single-norming (i.e. scalar-T_n)" case for the squared quasi-score function. Nevertheless, as in the original formulation of Yoshida [41, Theorem 1], it would be also possible to deal with "multi-norming" cases where elements of $\hat{\theta}_n$ possibly converge at different rates, i.e., cases of a matrix norming instead of the scalar norming $\sqrt{T_n}$. This would require somewhat more complicated arguments, but we do not need such an extension in this paper.*

4 Proofs of Theorem 2.7 and Corollary 2.8

4.1 Proof of Theorem 2.7

The proof of Theorem 2.7 is achieved by applying Theorem 3.5. When we have a reasonable estimating function $\theta \mapsto \mathbb{G}_n(\theta)$ with which an estimator of θ is defined by a random root of the estimating equation $\mathbb{G}_n(\theta) = 0$, it may be unclear what is the "single" associated contrast function to be maximized or minimized; for example, it would be often the case when $\mathbb{G}_n$ is constructed via a kind of (conditional-) moment fittings. The setup (24) provides a way of converting the situation from Z-estimation to M-estimation.

4.1.1 Introductory remarks

At first glance, it seems that, in order to investigate the asymptotic behavior of $\hat{\theta}_n$, we may proceed as in the case of diffusions, expanding the GQL $\mathbb{Q}_n$ of (11) and then investigating asymptotic behaviors of the derivatives $\partial_\theta^k \mathbb{Q}_n$; see Yoshida [41, Section 6] for details. Following this route however leads to a lesser evil, essentially due to the fact that $(h_n^{-1/2}\Delta_j X)_{j\le n}$ is not $L^q(P_0)$-bounded for $q > 2$. To see this more precisely, let us take a brief look at the simple one-dimensional Lévy process $X_t = \alpha t + \sqrt{\beta} J_t$, with $\theta = (\alpha, \beta) \in \mathbb{R} \times (0, \infty)$ and $\mathcal{L}(J_1)$ admitting finite moments. In this case, $\mathbb{Q}_n(\theta) = -\sum_j \{(\log \beta) + (\beta h_n)^{-1}(\Delta_j X - \alpha h_n)^2\}$:

$$
\begin{aligned}
&\partial_\alpha \mathbb{Q}_n(\theta) = \sum_{j=1}^n \frac{2}{\beta}(\Delta_j X - \alpha h_n), \quad \partial_\beta \mathbb{Q}_n(\theta) = \sum_{j=1}^n \frac{1}{\beta^2 h_n}\left\{(\Delta_j X - \alpha h_n)^2 - \beta h_n\right\}, \\
&\partial_\alpha^2 \mathbb{Q}_n(\theta) = \frac{-2T_n}{\beta}, \quad \partial_\alpha \partial_\beta \mathbb{Q}_n(\theta) = -\sum_{j=1}^n \frac{2}{\beta^2}(\Delta_j X - \alpha h_n), \\
&\partial_\beta^2 \mathbb{Q}_n(\theta) = -\sum_{j=1}^n \frac{2}{\beta^3 h_n}\left\{(\Delta_j X - \alpha h_n)^2 - \frac{\beta h_n}{2}\right\}.
\end{aligned}
$$

We can deduce the convergences

$$
\frac{1}{T_n}\partial_\alpha^2 \mathbb{Q}_n(\theta_0) \to^p -2\beta_0^{-1}, \qquad \frac{1}{\sqrt{n}\sqrt{T_n}}\partial_\alpha \partial_\beta \mathbb{Q}_n(\theta_0) \to^p 0, \qquad \frac{1}{n}\partial_\beta^2 \mathbb{Q}_n(\theta) \to^p -\beta_0^{-2},
$$

so that the normalized quasi observed-information matrix $-D_n^{-1}\partial_\theta^2 \mathbb{Q}_n(\theta_0) D_n^{-1} \to^p \mathrm{diag}(2\beta_0^{-1}, \beta_0^{-2})$, where $D_n := \mathrm{diag}(\sqrt{T_n}, \sqrt{n})$. In view of the classical Cramér type method for M-estimation, we should then have a central limit theorem for the normalized quasi score $\{T_n^{-1/2}\partial_\alpha \mathbb{Q}_n(\theta_0), n^{-1/2}\partial_\beta \mathbb{Q}_n(\theta_0)\}$ for an asymptotic normality at rate D_n to be valid for the M-estimator associated with $\mathbb{Q}_n$. However, different from the drifted Wiener process, the sequence $\{n^{-1/2}\partial_\beta \mathbb{Q}_n(\theta_0)\}$ does *not* converge, because $(h_n^{-1/2}\Delta_j X)_{j\le n}$ cannot be L^q-bounded for large $q > 2$ as can be seen from the moment structure of Lévy processes; see Luschgy and Pagès [21] for general moment estimates in small time with several concrete examples. Although we only mentioned the Lévy process with diagonal norming, situation remains the same even when X is actually an ergodic solution to (1).

The observation made in the last paragraph says that the situation is different from the case of diffusions, when developing asymptotic theory concerning the Gaussian quasi-likelihood for the model (1) under high-frequency sampling framework; it is also different from the case of time series models, where the usual $\sqrt{n}$-consistency holds in most cases (see the references cited in the Introduction). Earlier attempts to tackle this point have been made by Mancini [23], Shimizu and Yoshida [32], Ogihara and Yoshida [31], where they incorporated jump-detection filters in defining a contrast function. The filter approach has its own advantage such as $\sqrt{n}$-rate estimation of the diffusion parameter even in the presence of jumps, however, we should have in mind that its implementation involves fine-tuning parameters, thereby possibly preventing us from straightforward use of the approach.

In order to prove Theorem 2.7, we will look at not $\theta \mapsto \mathbb{Q}_n(\theta)$ but

$$
\theta \mapsto \mathbb{G}_n(\theta) = \left\{\mathbb{G}_n^\alpha(\theta), \mathbb{G}_n^\beta(\theta)\right\},
$$

where $\mathbb{G}_n^\alpha : \Theta \to \mathbb{R}^{p_\alpha}$ and $\mathbb{G}_n^\beta : \Theta \to \mathbb{R}^{p_\beta}$ are defined by

$$\mathbb{G}_n^\alpha(\theta) = \sum_{j=1}^n \partial_\alpha a_{j-1}(\alpha)\left[V_{j-1}^{-1}(\beta)[\chi_j(\alpha)]\right], \tag{22}$$

$$\mathbb{G}_n^\beta(\theta) = \sum_{j=1}^n \left(\{-\partial_\beta V_{j-1}^{-1}(\beta)\}[\chi_j(\alpha)^{\otimes 2}] - h_n \frac{\partial_\beta |V_{j-1}(\beta)|}{|V_{j-1}(\beta)|}\right). \tag{23}$$

Our contrast function $\mathbb{M}_n(\theta)$ is then defined to be the "squared quasi score" as in (14):

$$\mathbb{M}_n(\theta) = -\frac{1}{T_n}|\mathbb{G}_n(\theta)|^2. \tag{24}$$

Trivially, $\mathbb{G}_n : \Theta \to \mathbb{R}^p$ fulfil that $\mathbb{G}_n(\theta) = \left\{\partial_\alpha \mathbb{Q}_n(\theta), 2h_n\partial_\beta \mathbb{Q}_n(\theta)\right\}$. The difference is that we put the factor "$2h_n$" in front of $\partial_\beta \mathbb{Q}_n(\theta)$; our estimating procedure is formally not the usual M-estimation based on the Taylor expansion of $\theta \mapsto \mathbb{Q}_n(\theta)$ around θ_0, but rather a kind of minimum distance estimation concerning the Gaussian-quasi score function. The optimization with respect to θ is asymptotically the same for both of $\mathbb{Q}_n$ and $\mathbb{M}_n$: if there is no root $\theta \in \Theta$ for $\mathbb{G}_n(\theta) = 0$, then we may assign any value (e.g. any element of Θ) to $\hat{\theta}_n$, upholding the claim of Theorem 2.7.

Remark 4.1. *More general cases than (22) and (23) can be treated, such as*

$$\mathbb{G}_n^\alpha(\theta) = \sum_{j=1}^n W_{j-1}^\alpha(\theta)\{X_{t_j} - m_{j-1}(\theta)\},$$

$$\mathbb{G}_n^\beta(\theta) = \sum_{j=1}^n \left(W_{j-1}^{\beta,1}(\theta)[\{X_{t_j} - m_{j-1}(\theta)\}^{\otimes 2}] - h_n W_{j-1}^{\beta,2}(\theta)\right),$$

for some measurable $W^\alpha : \mathbb{R}^d \times \Theta \to \mathbb{R}^{p_\alpha} \otimes \mathbb{R}^d$, $W^{\beta,1} : \mathbb{R}^d \times \Theta \to \mathbb{R}^{p_\beta} \otimes (\mathbb{R}^d \otimes \mathbb{R}^d)$, and $W^{\beta,2} : \mathbb{R}^d \times \Theta \to \mathbb{R}^{p_\beta}$. This may be called a GQMLE as well, for we are still solely fitting the local mean vectors and covariance matrices. This setting allows us to deal with, for example, the parametric model

$$dX_t = a(X_t,\theta)dt + b(X_t,\theta)dw_t + c(X_{t-},\theta)dJ_t$$

with possibly degenerate b and c, the resulting GQMLE $\hat{\theta}_n$ still being asymptotically normal at rate $\sqrt{T_n}$ under suitable conditions. To avoid unnecessarily messy notation and regularity conditions without losing essence, we have decided to treat (1) in this paper.

For later use, we here introduce some convention and recall a couple of basic facts that we will make use often without notice.

- We will often suppress "(θ_0)" from the notation: $a_{j-1} := a_{j-1}(\alpha_0)$, $\mathbb{G}_n^\alpha = \mathbb{G}_n^\alpha(\theta_0)$, and so forth.
- $\int_j$ denotes a shorthand for $\int_{t_{j-1}}^{t_j}$.
- $M'_{j-1}(\theta) := \partial_\alpha a_{j-1}(\alpha)^\top V_{j-1}^{-1}(\beta) \in \mathbb{R}^{p_\alpha} \otimes \mathbb{R}^d$.
- $M''_{j-1}(\beta) := -\partial_\beta V_{j-1}^{-1}(\beta) = \{V_{j-1}^{-1}(\partial_\beta V_{j-1})\partial_\beta V_{j-1}^{-1}\}(\beta) \in \mathbb{R}^{p_\beta} \otimes \mathbb{R}^d \otimes \mathbb{R}^d$.
- $d_{j-1}(\beta) := |V_{j-1}(\beta)|^{-1}\{\partial_\beta |V_{j-1}(\beta)|\} \in \mathbb{R}^{p_\beta}$.
- Given real sequence a_n and random variables Y_n possibly depending on θ, we write $Y_n = O_p^*(a_n)$ if $\sup_{n,\theta} E_0\left[|a_n^{-1}Y_n|^K\right] < \infty$ for every $K > 0$.
- R denotes a generic function on $\mathbb{R}^d$ possibly depending on n and θ, for which there exists a constant $C \geq 0$ such that $\sup_{n,\theta}|R(x)| \leq C(1+|x|)^C$ for every $x \in \mathbb{R}^d$.

- Burkholder's inequality: for a martingale difference array $(\zeta_{nj})_{j\le n}$ and every $p \ge 1$,

$$E_0\left[\max_{k\le n}\left|\sum_{j\le k}\frac{1}{\sqrt{n}}\zeta_{nj}\right|^p\right] \lesssim E_0\left[\left(\frac{1}{n}\sum_{j\le n}\zeta_{nj}^2\right)^{p/2}\right] \lesssim \frac{1}{n}\sum_{i\le n}E\left[|\zeta_{nj}|^p\right].$$

  Moreover, if c is a sufficiently integrable predictable process, then

$$E_0\left[\left|\int_0^T c_{s-}dJ_s\right|^q\right] \lesssim (1\vee T)^{q/2-1}\int_0^T E_0[|c_s|^q]ds$$

  for every $T > 0$ and $q \ge 2$.

- Sobolev inequality (e.g. Friedman [7, Section 10.2]):

$$E_0\left[\sup_{\theta\in\Theta}|u(\theta)|\right] \lesssim \int_\Theta \{E_0[|u(\theta)|^q] + E_0[|\partial_\theta u(\theta)|^q]\}\,d\theta \lesssim \sup_{\theta\in\Theta}\{E_0[|u(\theta)|^q] + E_0[|\partial_\theta u(\theta)|^q]\}$$

  for $q > p$ and a random field $u \in \mathcal{C}^1(\Theta)$; recall that we are presupposing the boundedness and convexity of Θ. We will make use of this type of inequality to derive some uniform-in-θ moment estimates for martingale terms.

We now turn to the proof of Theorem 2.7 by verifying the conditions of Theorem 3.5.

4.1.2 Verification of the conditions on $\mathbb{G}_n$

We rewrite $\mathbb{G}_n$ as follows:

$$\mathbb{G}_n^\alpha(\theta) = \sum_{j=1}^n M'_{j-1}(\theta)[\chi_j] - h_n\sum_{j=1}^n M'_{j-1}(\theta)\left[a_{j-1}(\alpha) - a_{j-1}\right], \tag{25}$$

$$\begin{aligned}\mathbb{G}_n^\beta(\theta) &= \sum_{j=1}^n\left\{M''_{j-1}(\beta)[\chi_j^{\otimes 2}] - h_n d_{j-1}(\beta)\right\}\\ &\quad + 2h_n\sum_{j=1}^n M''_{j-1}(\beta)[\chi_j, a_{j-1} - a_{j-1}(\alpha)] + h_n^2\sum_{j=1}^n M''_{j-1}(\beta)\left[\{a_{j-1} - a_{j-1}(\alpha)\}^{\otimes 2}\right].\end{aligned} \tag{26}$$

We have $\chi_j = \zeta_j + r_j$, where

$$\zeta_j := \int_j \tilde{a}_{j-1}(s)ds + \int_j b(X_s,\beta_0)dw_s + \int_j c(X_{s-},\beta_0)dJ_s, \tag{27}$$

$$r_j := \int_j\left\{E_0^{j-1}[a(X_s,\alpha_0)] - a_{j-1}\right\}ds, \tag{28}$$

with $\tilde{a}_{j-1}(s) := a(X_s,\alpha_0) - E_0^{j-1}[a(X_s,\alpha_0)]$. Obviously, $(\zeta_j)_{j\le n}$ forms a martingale difference array with respect to the discrete-time filtration $(\mathcal{F}_{t_j})_{j\le n}$.

Itô's formula and the present integrability condition lead to

$$E_0^{j-1}[a(X_s,\alpha_0)] - a_{j-1} = \int_j E_0^{j-1}[\mathcal{A}a(X_u,\alpha_0)]du = h_n R_{j-1}, \tag{29}$$

where $\mathcal{A}$ denotes the (extended) generator associated with X under P_0, that is, for $f \in \mathcal{C}^2(\mathbb{R}^d)$

$$\begin{aligned}\mathcal{A}f(x) &= \partial f(x)[a(x,\alpha_0)] + \frac{1}{2}\partial^2 f(x)[b(x,\beta_0)^{\otimes 2}]\\ &\quad + \int\{f(x + c(x,\beta_0)z) - f(x) - \partial f(x)c(x,\beta_0)z\}\nu(dz).\end{aligned}$$

Putting (28) and (29) together gives $r_j = h_n^2 R_{j-1}$, therefore

$$\chi_j = \zeta_j + h_n^2 R_{j-1}. \tag{30}$$

Assumption 3.1 obviously holds under the present differentiability conditions. We begin with verifying Assumption 3.2.

Lemma 4.2. *For every $K>0$, we have*

$$\sup_{n\in\mathbb{N}} E_0\left[\left|\frac{1}{\sqrt{T_n}}\mathbb{G}_n(\theta_0)\right|^K\right] + \sup_{n\in\mathbb{N}} E_0\left[\sup_{\theta\in\Theta}\left|\frac{1}{T_n}\mathbb{G}_n(\theta)\right|^K\right] < \infty.$$

Proof. By substituting (30) in (25) and (26) and then rearranging the resulting terms, we have

$$\mathbb{G}_n^\alpha(\theta) = \sum_{j=1}^n M'_{j-1}(\theta)\zeta_j + h_n\sum_{j=1}^n M'_{j-1}(\theta)\{a_{j-1}-a_{j-1}(\alpha)\} + h_n^2\sum_{j=1}^n M'_{j-1}(\theta)R_{j-1}, \tag{31}$$

$$\mathbb{G}_n^\beta(\theta) = \sum_{j=1}^n \left\{M''_{j-1}(\beta)[\zeta_j^{\otimes 2}] - h_n d_{j-1}(\beta)\right\} + 2h_n\sum_{j=1}^n M''_{j-1}(\beta)[\zeta_j, a_{j-1}-a_{j-1}(\alpha)] + h_n^2\sum_{j=1}^n R_{j-1}. \tag{32}$$

To achieve the proof, we will separately look at $T_n^{-1/2}\mathbb{G}_n^\alpha$, $T_n^{-1/2}\mathbb{G}_n^\beta$, $T_n^{-1}\mathbb{G}_n^\alpha(\theta)$, and $T_n^{-1}\mathbb{G}_n^\beta(\theta)$. Fix any integer $K>(2\vee p)$ in the sequel.

First we prove $T_n^{-1/2}\mathbb{G}_n^\alpha = O_p^*(1)$. Observe that

$$\frac{1}{\sqrt{T_n}}\mathbb{G}_n^\alpha = \sum_{j=1}^n \frac{1}{\sqrt{T_n}}M'_{j-1}\zeta_j + \sqrt{T_n h_n^2}\frac{1}{n}\sum_{j=1}^n M'_{j-1} = \sum_{j=1}^n \frac{1}{\sqrt{T_n}}M'_{j-1}\zeta_j + O_p^*\left(\sqrt{T_n h_n^2}\right).$$

By (27),

$$\begin{aligned}\sum_{j=1}^n \frac{1}{\sqrt{T_n}}M'_{j-1}\zeta_j &= \sum_{j=1}^n \frac{1}{\sqrt{n}}\left(M'_{j-1}\frac{1}{\sqrt{h_n}}\int_j b(X_s,\beta_0)dw_s\right) + \sqrt{h_n}\sum_{j=1}^n\frac{1}{\sqrt{n}}\left(M'_{j-1}\frac{1}{h_n}\int_j \tilde{a}_{j-1}(s)ds\right)\\ &\quad + \sum_{j=1}^n \frac{1}{\sqrt{T_n}}M'_{j-1}\int_j c(X_{s-},\beta_0)dJ_s. \end{aligned}\tag{33}$$

Burkholder's inequality implies that the first and second term on the right-hand side are $O_p^*(1)$ and $O_p^*(\sqrt{h_n})$, respectively. As for the last term, by writing $\mathbf{1}_j:(0,\infty)\to\{0,1\}$ for the identity function of the interval $(t_{j-1},t_j]$,

$$\begin{aligned}E_0\left[\left|\sum_{j=1}^n\frac{1}{\sqrt{T_n}}M'_{j-1}\int_j c(X_{s-},\beta_0)dJ_s\right|^K\right] &\lesssim T_n^{-K/2}E_0\left[\left|\int_0^{T_n}\sum_{j=1}^n \mathbf{1}_j(s)M'_{j-1}c(X_{s-},\beta_0)dJ_s\right|^K\right]\\ &\lesssim T_n^{-K/2}T_n^{K/2-1}\int_0^{T_n}E_0\left[\left(\sum_{j=1}^n\mathbf{1}_j(s)|M'_{j-1}c(X_{s-},\beta_0)|\right)^K\right]ds\\ &= \frac{1}{T_n}\int_0^{T_n}\sum_{j=1}^n\mathbf{1}_j(s)E_0[|M'_{j-1}c(X_{s-},\beta_0)|^K]ds\\ &\lesssim \frac{1}{T_n}\sum_{j=1}^n\int_j ds = 1, \end{aligned}\tag{34}$$

hence we are done.

We now prove $T_n^{-1/2}\mathbb{G}_n^\beta = O_p^*(1)$. In the sequel, we may and do suppose that $d=p_\beta=r'=r''=1$: this reduction is possible because of the the polarization identity

$$[S',S''] = \frac{1}{4}\left([S'+S'']-[S'-S'']\right),$$

which is valid for any two semimartingales S' and S''. Substituting (30) in (26) gives

$$\frac{1}{\sqrt{T_n}}\mathbb{G}_n^\beta = \sum_{j=1}^n\frac{1}{\sqrt{T_n}}\left(M''_{j-1}\zeta_j^2 - h_n d_{j-1}\right) + O_p^*\left(\sqrt{T_n h_n^2}\right),$$

so that it remains to verify

$$\sum_{j=1}^{n} \frac{1}{\sqrt{T_n}} M''_{j-1}(\zeta_j^2 - h_n V_{j-1}) = O_p^*(1). \tag{35}$$

Define $\zeta_j(t)$ for $t \in (t_{j-1}, t_j]$ by

$$\zeta_j(t) = \int_{t_{j-1}}^{t} \tilde{a}_{j-1}(s)ds + \int_{t_{j-1}}^{t} b(X_s, \beta_0)dw_s + \int_{t_{j-1}}^{t} c(X_{s-}, \beta_0)dJ_s.$$

Let $N(ds, dz)$ denote the Poisson random measure associated with J (i.e. $J_t = \int_0^t \int z\tilde{N}(ds, dz)$), and $\tilde{N}$ its compensated version. The quadratic variation at time t is then given as follows (cf. Jacod and Shiryaev [12, I.4.49(d), I.4.55(b)]):

$$\begin{aligned}
[\zeta_j(\cdot)]_t &= \int_{t_{j-1}}^{t} b^2(X_{s-}, \beta_0)ds + \int_{t_{j-1}}^{t} \int c^2(X_{s-}, \beta_0)z^2 N(ds, dz) \\
&= (t - t_{j-1})V_{j-1} + \int_{t_{j-1}}^{t} \int c^2(X_{s-}, \beta_0)\tilde{N}(ds, dz) + \int_{t_{j-1}}^{t} g_{j-1}(s)ds,
\end{aligned}$$

where we used the assumption $\int z^2 \nu(dz) = 1$ (with the temporary assumption $r'' = 1$) and $g_{j-1}(s) := b^2(X_s, \beta_0) - b_{j-1}^2 + c^2(X_{s-}, \beta_0) - c_{j-1}^2$. Applying the integration-by-parts formula, we get

$$\begin{aligned}
\zeta_j^2 - h_n V_{j-1} &= \left\{ 2\int_j \zeta_j(s-)d\zeta_j(s) + \int_j \int c^2(X_{s-}, \beta_0)z^2\tilde{N}(ds, dz) \right. \\
&\qquad \left. + \int_j \left(g_{j-1}(s) - E_0^{j-1}[g_{j-1}(s)] \right) ds \right\} + \int_j E_0^{j-1}[g_{j-1}(s)]ds \\
&=: \zeta_j^{(0)} + \zeta_j^{(1)}, \quad \text{say.}
\end{aligned}$$

We can deduce that $\sum_{j=1}^{n} T_n^{-1/2} M''_{j-1}\zeta_j^{(0)} = O_p^*(1)$, as is the case in the proof of $\sum_{j=1}^{n} T_n^{-1/2} M'_{j-1}\zeta_j = O_p^*(1)$ via the expression (33). Moreover, we can apply Itô's formula to get $\zeta_j^{(1)} = h_n^2 R_{j-1}$ under the $\mathcal{C}^2$ property of $x \mapsto (b(x, \beta_0), c(x, \beta_0))$, from which it follows that $\sup_n E_0[|\sum_{j=1}^{n} T_n^{-1/2} M''_{j-1}\zeta_j^{(1)}|^K] \lesssim \sup_n (T_n h_n^2)^{K/2} < \infty$. We thus get (35).

Let us turn to prove $\sup_\theta |T_n^{-1}\mathbb{G}_n^\alpha(\theta)| = O_p^*(1)$. In the same way as in the proof of $T_n^{-1/2}\mathbb{G}_n^\alpha = O_p^*(1)$, we can prove $\sum_{j=1}^{n} T_n^{-1/2} M'_{j-1}(\theta)\zeta_j = O_p^*(T_n^{-1/2})$ for each $\theta \in \Theta$, since the explicit dependence on θ is only through the predictable parts $M'_{j-1}(\theta)$; similar arguments will apply in some places below. Therefore, it follows from (31) that, for each $\theta \in \Theta$,

$$\begin{aligned}
\frac{1}{T_n}\mathbb{G}_n^\alpha(\theta) &= \frac{1}{\sqrt{T_n}} \left(\sum_{j=1}^{n} \frac{1}{\sqrt{T_n}} M'_{j-1}(\theta)\zeta_j \right) + h_n \left(\frac{1}{n}\sum_{j=1}^{n} M'_{j-1}(\theta) \right) + \frac{1}{n}\sum_{j=1}^{n} M'_{j-1}(\theta)\{a_{j-1} - a_{j-1}(\alpha)\} \\
&= O_p^*\left(\frac{1}{\sqrt{T_n}} \vee h_n \right) + \frac{1}{n}\sum_{j=1}^{n} M'_{j-1}(\theta)\{a_{j-1} - a_{j-1}(\alpha)\} \\
&= O_p^*\left(\frac{1}{\sqrt{T_n}} \right) + \frac{1}{n}\sum_{j=1}^{n} M'_{j-1}(\theta)\{a_{j-1} - a_{j-1}(\alpha)\},
\end{aligned} \tag{36}$$

so that $T_n^{-1}\partial_\theta\mathbb{G}_n^\alpha(\theta) = O_p^*(1)$, and in a quite similar manner we obtain (see (50) and (51) below)

$$\frac{1}{T_n}\partial_\theta\mathbb{G}_n^\alpha(\theta) = O_p^*\left(\frac{1}{\sqrt{T_n}} \right) + \frac{1}{n}\sum_{j=1}^{n} \partial_\theta \left[M'_{j-1}(\theta)\{a_{j-1} - a_{j-1}(\alpha)\} \right] = O_p^*(1). \tag{37}$$

Therefore, we arrive at $\sup_\theta |T_n^{-1}\mathbb{G}_n^\alpha(\theta)| = O_p^*(1)$ by means of the Sobolev inequality.

It remains to prove $\sup_\theta |T_n^{-1}\mathbb{G}_n^\beta(\theta)| = O_p^*(1)$; remind we are supposing that $d = p_\beta = r' = r'' = 1$. As in the proof of (35), we can prove

$$\sum_{j=1}^n \frac{1}{\sqrt{T_n}} \partial_\theta^k M''_{j-1}(\beta)(\zeta_j^2 - h_n V_{j-1}) = O_p^*(1)$$

for each $k = 0, 1$ and β, so that the Sobolev inequality gives $\sum_{j=1}^n T_n^{-1/2} M''_{j-1}(\beta)(\zeta_j^2 - h_n V_{j-1}) = O_p^*(1)$. Therefore, it follows from (32) and simple manipulation that

$$\begin{aligned}
\frac{1}{T_n}\mathbb{G}_n^\beta(\theta) &= \frac{1}{\sqrt{T_n}}\left(\sum_{j=1}^n \frac{1}{\sqrt{T_n}} M''_{j-1}(\beta)(\zeta_j^2 - h_n V_{j-1})\right) + \frac{2\sqrt{T_n}}{n}\sum_{j=1}^n \frac{1}{\sqrt{T_n}} M''_{j-1}(\beta)\left\{a_{j-1} - a_{j-1}(\alpha)\right\}\zeta_j \\
&\quad + \frac{h_n}{n}\sum_{j=1}^n R_{j-1} + \frac{1}{n}\sum_{j=1}^n M''_{j-1}(\beta)\left\{V_{j-1} - V_{j-1}(\beta)\right\} \\
&= O_p^*\left(\frac{1}{\sqrt{T_n}} \vee \frac{\sqrt{T_n}}{n} \vee h_n\right) + \frac{1}{n}\sum_{j=1}^n M''_{j-1}(\beta)\left\{V_{j-1} - V_{j-1}(\beta)\right\} \\
&= O_p^*\left(\frac{1}{\sqrt{T_n}}\right) + \frac{1}{n}\sum_{j=1}^n M''_{j-1}(\beta)\left\{V_{j-1} - V_{j-1}(\beta)\right\}. \qquad (38)
\end{aligned}$$

Thus $T_n^{-1}\mathbb{G}_n^\beta(\theta) = O_p^*(1)$. Similarly, we get $T_n^{-1}\partial_\theta\mathbb{G}_n^\beta(\theta) = O_p^*(1)$ (see (52) and (53) below)

$$\frac{1}{T_n}\partial_\theta\mathbb{G}_n^\beta(\theta) = O_p^*\left(\frac{1}{\sqrt{T_n}}\right) + \frac{1}{n}\sum_{j=1}^n \partial_\theta\left[M''_{j-1}(\beta)\left\{V_{j-1} - V_{j-1}(\beta)\right\}\right] = O_p^*(1), \qquad (39)$$

completing the proof. □

Next we turn to verifying the uniform moment estimates in Assumptions 3.3. To this end, we prove a preliminary lemma.

Lemma 4.3. *Let $f : \mathbb{R}^d \times \Theta \to \mathbb{R}$ and $g : \mathbb{R}^d \to [1, \infty)$ be measurable functions and suppose the following conditions:*

- *$\theta \mapsto f(x, \theta)$ is differentiable for each x, and $\sup_\theta\{|f(x,\theta)| \vee |\partial_\theta f(x,\theta)|\} \le g(x)$ for $k = 0, 1$;*
- *$\sup_t E_0[g(X_t)] < \infty$;*
- *$\|P_t(x, \cdot) - \pi_0(\cdot)\|_g \lesssim e^{-at} g(x)$ for some constant $a > 0$ and a probability measure π_0;*

Then we have

$$\sup_{n\in\mathbb{N}} E_0\left[\sup_{\theta\in\Theta}\left|\sqrt{T_n}\left(\frac{1}{n}\sum_{j=1}^n f_{j-1}(\theta) - \int f(x,\theta)\pi_0(dx)\right)\right|^K\right] < \infty.$$

Proof. Let $\Lambda'_n(f;\theta) := n^{-1}\sum_{j=1}^n\{f_{j-1}(\theta) - E_0[f_{j-1}(\theta)]\}$ and $\Lambda''_n(f;\theta) := n^{-1}\sum_{j=1}^n\{E_0[f_{j-1}(\theta)] - \int f(x,\theta)\pi_0(dx)\}$, so that $n^{-1}\sum_{j=1}^n f_{j-1}(\theta) - \int f(x,\theta)\pi_0(dx) = \Lambda'_n(f;\theta) + \Lambda''_n(f;\theta)$. Under the present assumptions, we can apply Yoshida [41, Lemma 4] to get $E_0[|\partial_\theta^k\Lambda'_n(\theta)|^K] \lesssim T_n^{-K/2} + T_n^{1-K} \lesssim T_n^{-K/2}$ for $k = 0, 1$, yielding that $\max_{k=0,1}\sup_\theta\sup_n E_0[|\sqrt{T_n}\partial_\theta^k\Lambda'_n(f;\theta)|^K] < \infty$. Therefore, the Sobolev inequality gives

$$\sup_{n\in\mathbb{N}} E_0\left[\sup_{\theta\in\Theta}\left|\sqrt{T_n}\Lambda'_n(f;\theta)\right|^K\right] < \infty$$

As for $\Lambda''_n(f;\theta)$, we have for $k = 0, 1$:

$$\left|\sqrt{T_n}\partial_\theta^k\Lambda''_n(f;\theta)\right| = \left|\frac{\sqrt{T_n}}{n}\sum_{j=1}^n\left(\iint \partial_\theta^k f(y,\theta)P_{t_{j-1}}(x,dy)\eta(dx) - \iint \partial_\theta^k f(y,\theta)\pi_0(dy)\eta(dx)\right)\right|$$

$$
\begin{aligned}
&= \left| \frac{\sqrt{T_n}}{n} \sum_{j=1}^{n} \int \left(\int \partial_\theta^k f(y,\theta) \left\{ P_{t_{j-1}}(x,dy) - \pi_0(dx) \right\} \right) \eta(dx) \right| \\
&\leq \frac{\sqrt{T_n}}{n} \sum_{j=1}^{n} \int \left\| P_{t_{j-1}}(x,\cdot) - \pi_0(\cdot) \right\|_g \eta(dx) \\
&\lesssim \frac{\sqrt{T_n}}{n} \sum_{j=1}^{n} \exp(-a t_{j-1}) \lesssim \frac{1}{\sqrt{T_n}}.
\end{aligned}
$$

This ends the proof. □

Corollary 4.4. *Assumption 3.3(a) holds true.*

Proof. Again we may and do suppose that $d = p_\beta = r' = r'' = 1$. Recalling (36), (37), (38), and (39), we apply Lemma 4.3 with $f(x,\theta) = M'(x,\theta)\{a(x,\alpha_0) - a(x,\alpha)\}$ and $f(x,\theta) = M''(x,\beta)\{V(x,\beta_0) - V(x,\beta)\}$ to conclude

$$
\sup_{n\in\mathbb{N}} E_0 \left[\sup_{\theta\in\Theta} \left| \sqrt{T_n} \left(\frac{1}{T_n} \mathbb{G}_n(\theta) - \mathbb{G}_\infty(\theta) \right) \right|^K \right] < \infty
$$

for every $K > 0$, where $\mathbb{G}_\infty(\theta) := (\mathbb{G}_\infty^\alpha(\theta), \mathbb{G}_\infty^\beta(\theta))$ are given by (6) and (7), the integrals in which are finite by the assumptions. Trivially $\mathbb{G}_\infty(\theta_0) = 0$, and Assumption 3.3(a) is verified with $\chi = \chi_\alpha \wedge \chi_\beta$. □

Let us mention the fundamental fact concerning conditional size of X's increments. For the convenience of reference we include a sketch of the proof.

Lemma 4.5. *Let $g(x) := |a(x,\alpha_0)| \vee |b(x,\beta_0)| \vee |c(x,\beta_0)|$, and fix any $q \geq 2$ such that $E[|J_t|^q] < \infty$. Then*

$$
E_0^{j-1}\left[\sup_{s\in[t_{j-1},t_j]} |X_s - X_{t_{j-1}}|^q \right] \lesssim \begin{cases} h_n^{q/2} g^q(X_{t_{j-1}}) & \text{if } c \equiv 0, \\ h_n g^q(X_{t_{j-1}}) & \text{otherwise}. \end{cases}
$$

In particular, the left-hand side is essentially bounded if so is g.

Proof. We only mention the case of $c \not\equiv 0$. Given an $M > 0$, we let $\tau_{j-1,M} := \inf\{s \geq t_{j-1} : |X_s| \geq M\}$ and $\xi_{j-1,M}(s) := E_0^{j-1}\left[\sup\{|X_u - X_{t_{j-1}}|^q : u \in [t_{j-1}, s \wedge \tau_{j-1,M}]\}\right]$. We can make use of the Lipschitz property of the coefficients and Masuda [24, Lemma E.1] to derive $\xi_{j-1,M}(t_j) \lesssim \int_{t_{j-1}}^{t_j} \xi_{j-1,M}(s)ds + h_n g^q(X_{t_{j-1}})$, the upper bound being P_0-a.s. finite according to the definition of $\tau_{j-1,M}$. Hence the claim follows on applying Gronwall's inequality and then letting $M \uparrow \infty$. The case of $c \equiv 0$ is similar. □

We now prove the central limit theorem required in Assumption 3.4.

Lemma 4.6. *We have*

$$
\frac{1}{\sqrt{T_n}} \mathbb{G}_n(\theta_0) \to^{\mathcal{L}} \mathcal{N}_p\left(0, \mathbb{V}(\theta_0)\right),
$$

where $\mathbb{V}(\theta_0)$ is given by (12).

Proof. We begin with extracting the leading martingale terms of the sequences $T_n^{-1/2}\mathbb{G}_n^\alpha$ and $T_n^{-1/2}\mathbb{G}_n^\beta$; recall the expressions (31) and (32). Let us rewrite (27) as $\zeta_j = m_j + r_j'$, where

$$
\begin{aligned}
m_j &:= b_{j-1}\Delta_j w + c_{j-1}\Delta_j J, \\
r_j' &:= \int_j \tilde{a}_{j-1}(s)ds + \int_j (b(X_s,\beta_0) - b_{j-1})dw_s + \int_j (c(X_{s-},\beta_0) - c_{j-1})dJ_s.
\end{aligned}
$$

We claim that it suffices to prove that

$$
\sum_{j=1}^{n} \frac{1}{\sqrt{T_n}} \begin{pmatrix} \tilde{\gamma}_j^\alpha \\ \tilde{\gamma}_j^\beta \end{pmatrix} \to^{\mathcal{L}} \mathcal{N}_p\left(0, \mathbb{V}(\theta_0)\right), \tag{40}
$$

where $\tilde{\gamma}_j^\alpha := M'_{j-1} m_j$ and $\tilde{\gamma}_j^\beta := M''_{j-1}[m_j^{\otimes 2}] - h_n d_{j-1}$, both of which form martingale difference arrays with respect to $(\mathcal{F}_{t_j})_{j\le n}$; we can verify that $E_0^{j-1}[\tilde{\gamma}_j^\beta[u]] = 0$ for each $u \in \mathbb{R}^{p_\beta}$, making use of the identity $\mathrm{trace}\{A(x)^{-1}\partial_x A(x)\} = \partial_x |A(x)|/|A(x)|$ for a differentiable square-matrix function A. In fact, recalling what we have seen in the proof of Lemma 4.2, we observe the following.

- We have

$$\begin{aligned}
\frac{1}{\sqrt{T_n}}\mathbb{G}_n^\alpha &= \sum_{j=1}^n \frac{1}{\sqrt{T_n}} M'_{j-1}\left(\int_j b(X_s,\beta_0)dw_s + \int_j c(X_{s-},\beta_0)dJ_s\right) + o_p(1) \\
&= \sum_{j=1}^n \frac{1}{\sqrt{T_n}}\tilde{\gamma}_j^\alpha + \sum_{j=1}^n \frac{1}{\sqrt{T_n}} M'_{j-1}\int_j (b(X_s,\beta_0) - b_{j-1})dw_s \\
&\quad + \sum_{j=1}^n \frac{1}{\sqrt{T_n}} M'_{j-1}\int_j (c(X_{s-},\beta_0) - c_{j-1})dJ_s + o_p(1).
\end{aligned}$$

  By means of Burkholder's inequality and Lemma 4.5 combined with the conditioning argument,

$$E_0\left[\left|\sum_{j=1}^n \frac{1}{\sqrt{T_n}} M'_{j-1}\int_j (b(X_s,\beta_0) - b_{j-1})dw_s\right|^2\right] \lesssim E_0\left[\sum_{j=1}^n \frac{1}{T_n}|M'_{j-1}|^2 |R_{j-1}| \int_j h_n ds\right] \lesssim h_n.$$

  Following the same line as in (34), we also get $E_0[|\sum_{j=1}^n T_n^{-1/2} M'_{j-1}\int_j (c(X_s,\beta_0) - c_{j-1})dJ_s|^2] \lesssim h_n$. Therefore, it follows that

$$\frac{1}{\sqrt{T_n}}\mathbb{G}_n^\alpha = \sum_{j=1}^n \frac{1}{\sqrt{T_n}}\tilde{\gamma}_j^\alpha + o_p(1). \tag{41}$$

- Put $B'_n = 2\sum_{j=1}^n T_n^{-1/2} M''_{j-1}[m_j, r'_j]$ and $B''_n = \sum_{j=1}^n T_n^{-1/2} M''_{j-1}[r'_j, r'_j]$, then we see that

$$\frac{1}{\sqrt{T_n}}\mathbb{G}_n^\beta = \sum_{j=1}^n \frac{1}{\sqrt{T_n}}(M''_{j-1}[\zeta_j^{\otimes 2}] - h_n d_{j-1}) + o_p(1) = \sum_{j=1}^n \frac{1}{\sqrt{T_n}}\tilde{\gamma}_j^\beta + B'_n + B''_n + o_p(1).$$

  Since $\sup_{j\le n} E_0[|r'_j|^q] \lesssim h_n^2$ for every $q \ge 2$ and $E_0^{j-1}[|m_j|^2] \lesssim |R_{j-1}|^2 h_n$, the Cauchy-Schwarz inequality leads to

$$E_0[|B'_n|] \lesssim \frac{1}{n}\sum_{j=1}^n \sqrt{\frac{n}{h_n}} E_0\left[|R_{j-1}|^2 E_0^{j-1}[|m_j|^2]\right]^{1/2} E_0\left[|r'_j|^2\right]^{1/2} \lesssim \sqrt{n h_n^2} \to 0.$$

  Moreover, for any $\epsilon \in (0, 1/3)$ Hölder's inequality gives

$$\begin{aligned}
E_0[|B''_n|] &\lesssim \frac{1}{n}\sum_{j=1}^n \sqrt{\frac{n}{h_n}} E_0\left[|R_{j-1}||r'_j|^2\right] \\
&\lesssim \frac{1}{n}\sum_{j=1}^n \sqrt{\frac{n}{h_n}} E_0\left[|R_{j-1}|^{(1+\epsilon)/\epsilon}\right]^{\epsilon/(1+\epsilon)} E_0\left[|r'_j|^{2(1+\epsilon)}\right]^{1/(1+\epsilon)} \\
&\lesssim \frac{1}{n}\sum_{j=1}^n \sqrt{\frac{n}{h_n}} E_0\left[|r'_j|^{2(1+\epsilon)}\right]^{1/(1+\epsilon)} \lesssim \sqrt{n h_n^{4/(1+\epsilon)-1}} \lesssim \sqrt{n h_n^2} \to 0.
\end{aligned}$$

  Hence we have derived

$$\frac{1}{\sqrt{T_n}}\mathbb{G}_n^\beta = \sum_{j=1}^n \frac{1}{\sqrt{T_n}}\tilde{\gamma}_j^\beta + o_p(1). \tag{42}$$

Having (41) and (42) in hand, it remains to verify (40). We are going to apply the classical martingale central limit theorem (e.g. Dvoretzky [4]).

Put $\tilde{\gamma}_j = (\tilde{\gamma}_j^\alpha, \tilde{\gamma}_j^\beta)$. It is easy to verify the Lyapunov condition: in fact, we have $E_0^{j-1}[|\tilde{\gamma}_j|^K] \lesssim h_n|R_{j-1}|$ for any $K > 2$, so that $\sum_{j=1}^n E_0[|T_n^{-1/2}\tilde{\gamma}_j|^K] \lesssim T_n^{1-K/2} \to 0$. It remains to compute the convergence of the quadratic characteristics: $\sum_{j=1}^n E_0^{j-1}[\tilde{\gamma}_j^{\otimes 2}] \to^p \mathbb{V}(\theta_0)$. By means of the Cramér-Wold device, it suffices to prove that for each $v_1', v_2' \in \mathbb{R}^{p_\alpha}$ and $v_1'', v_2'' \in \mathbb{R}^{p_\beta}$,

$$\sum_{j=1}^n \frac{1}{T_n} E_0^{j-1}\left[(\tilde{\gamma}_j^\alpha)^{\otimes 2}\right][v_1', v_2'] \to^p \mathbb{G}_\infty'^\alpha[v_1', v_2'], \tag{43}$$

$$\mathbb{V}_{\alpha\beta,n}[v_1', v_1''] := \sum_{j=1}^n \frac{1}{T_n} E_0^{j-1}\left[\tilde{\gamma}_j^\alpha \otimes \tilde{\gamma}_j^\beta\right][v_1', v_1''] \to^p \mathbb{V}_{\alpha\beta}[v_1', v_1''], \tag{44}$$

$$\mathbb{V}_{\beta\beta,n}[v_1'', v_2''] := \sum_{j=1}^n \frac{1}{T_n} E_0^{j-1}\left[(\tilde{\gamma}_j^\beta)^{\otimes 2}\right][v_1'', v_2''] \to^p \mathbb{V}_{\beta\beta}[v_1'', v_2'']. \tag{45}$$

First, (43) readily follows on noting $E_0^{j-1}[m_j^{\otimes 2}] = h_n V_{j-1}$ and applying the ergodic theorem (5). Next,

$$\begin{aligned}
\mathbb{V}_{\alpha\beta,n}[v_1', v_1''] &= \frac{1}{n}\sum_{j=1}^n \frac{1}{h_n} E_0^{j-1}\left[M_{j-1}'[m_j] \otimes M_{j-1}''[m_j^{\otimes 2}]\right][v_1', v_1''] \\
&= \frac{1}{n}\sum_{j=1}^n \frac{1}{h_n} \sum_{k,l,s} E_0^{j-1}\left[m_j^{(k)} m_j^{(l)} m_j^{(s)}\right]\left\{M_{j-1}'^{(\cdot s)} \otimes M_{j-1}''^{(\cdot kl)}\right\}[v_1', v_1''].
\end{aligned}\tag{46}$$

For later use, we here note that, as $h \to 0$,

$$E\left[J_h^{(i_1)} \cdots J_h^{(i_m)}\right] = \begin{cases} h\nu_{i_1 i_2 i_3}(3) & m = 3, \\ h\nu_{i_1 i_2 i_3 i_4}(4) + O(h^2) & m = 4; \end{cases}$$

this can be easily seen through the relation between the mixed moments and cumulants of J_h, where the latter can be computed as the values at 0 of the partial derivatives of the cumulant function $u \mapsto \log E[\exp(iJ_h[u])] = h\int\{\exp(iu[z]) - 1 - iu[z]\}\nu(dz)$. In view of the expression

$$m_j^{(k)} = \sum_{k'} b_{j-1}^{(kk')} \Delta_j w^{(k')} + \sum_{k''} c_{j-1}^{(kk'')} \Delta_j J^{(k'')}$$

together with the orthogonalities between the increments of w and J, we get

$$\begin{aligned}
E_0^{j-1}\left[m_j^{(k)} m_j^{(l)} m_j^{(s)}\right] &= \sum_{k',l',s'} c_{j-1}^{(kk')} c_{j-1}^{(ll')} c_{j-1}^{(ss')} E\left[\Delta_j J^{(k')} \Delta_j J^{(l')} \Delta_j J^{(s')}\right] \\
&= \sum_{k',l',s'} c_{j-1}^{(kk')} c_{j-1}^{(ll')} c_{j-1}^{(ss')} E\left[J_{h_n}^{(k')} J_{h_n}^{(l')} J_{h_n}^{(s')}\right] \\
&= h_n \sum_{k',l',s'} c_{j-1}^{(kk')} c_{j-1}^{(ll')} c_{j-1}^{(ss')} \nu_{k'l's'}(3).
\end{aligned}\tag{47}$$

(Since $E[J_1] = 0$, the 3rd mixed cumulants and the 3rd mixed moments of J_{h_n} coincides.) Substituting (47) in (46), we get (44):

$$\begin{aligned}
\mathbb{V}_{\alpha\beta,n}[v_1', v_1''] &= \frac{1}{n}\sum_{j=1}^n \sum_{k,l,s} \sum_{k',l',s'} c_{j-1}^{(kk')} c_{j-1}^{(ll')} c_{j-1}^{(ss')} \nu_{k'l's'}(3) \left\{M_{j-1}'^{(\cdot s)} \otimes M_{j-1}''^{(\cdot kl)}\right\}[v_1', v_1''] \\
&= \frac{1}{n}\sum_{j=1}^n \sum_{k',l',s'} \nu_{k'l's'}(3) \left\{M_{j-1}'[v_1', c_{j-1}^{(\cdot s')}]\right\}\left\{M_{j-1}''[v_1'', c_{j-1}^{(\cdot k')}, c_{j-1}^{(\cdot l')}]\right\} \\
&\to^p \mathbb{V}_{\alpha\beta}[v_1', v_1''].
\end{aligned}$$

Finally, we look at $\mathbb{V}_{\beta\beta,n}$. Direct computation gives

$$\begin{aligned}
\mathbb{V}_{\beta\beta,n}[v_1'', v_2''] &= \frac{1}{n}\sum_{j=1}^{n}\frac{1}{h_n}E_0^{j-1}\left[\left(M_{j-1}''\otimes M_{j-1}''\right)[v_1'', m_j^{\otimes 2}, v_2'', m_j^{\otimes 2}]\right] \\
&\quad - \frac{1}{n}\sum_{j=1}^{n}E_0^{j-1}\left[\left(d_{j-1}\otimes M_{j-1}''\right)[v_1'', v_2'', m_j^{\otimes 2}]\right] \\
&\quad - \frac{1}{n}\sum_{j=1}^{n}E_0^{j-1}\left[\left(d_{j-1}\otimes M_{j-1}''\right)[v_2'', v_1'', m_j^{\otimes 2}]\right] + h_n\left(\frac{1}{n}\sum_{j=1}^{n}d_{j-1}^{\otimes 2}[v_1'', v_2'']\right) \\
&= \frac{1}{n}\sum_{j=1}^{n}\frac{1}{h_n}E_0^{j-1}\left[\{M_{j-1}''[v_1'', m_j^{\otimes 2}]\}\{M_{j-1}''[v_2'', m_j^{\otimes 2}]\}\right] + O_p(h_n) \\
&= \frac{1}{n}\sum_{j=1}^{n}\frac{1}{h_n}\sum_{k,l,k',l'}M_{j-1}''^{(\cdot kl)}[v_1'']M_{j-1}''^{(\cdot k'l')}[v_2'']E_0^{j-1}\left[m_j^{(k)}m_j^{(l)}m_j^{(k')}m_j^{(l')}\right] + O_p(h_n). \qquad (48)
\end{aligned}$$

Using the orthgonality as before and noting the fact that $E[|w_{h_n}|^4] = O(h_n^2)$, we get

$$\begin{aligned}
E_0^{j-1}\left[m_j^{(k)}m_j^{(l)}m_j^{(k')}m_j^{(l')}\right] &= \sum_{s,t,s',t'}c_{j-1}^{(ks)}c_{j-1}^{(lt)}c_{j-1}^{(k's')}c_{j-1}^{(l't')}E\left[J_{h_n}^{(s)}J_{h_n}^{(t)}J_{h_n}^{(s')}J_{h_n}^{(t')}\right] + R_{j-1}h_n^2 \\
&= h_n\sum_{s,t,s',t'}c_{j-1}^{(ks)}c_{j-1}^{(lt)}c_{j-1}^{(k's')}c_{j-1}^{(l't')}\{\nu_{sts't'}(4) + O(h_n)\} + R_{j-1}h_n^2 \\
&= h_n\sum_{s,t,s',t'}c_{j-1}^{(ks)}c_{j-1}^{(lt)}c_{j-1}^{(k's')}c_{j-1}^{(l't')}\nu_{sts't'}(4) + R_{j-1}h_n^2. \qquad (49)
\end{aligned}$$

By putting (48) and (49) together, we get (45):

$$\begin{aligned}
\mathbb{V}_{\beta\beta,n}[v_1'', v_2''] &= \frac{1}{n}\sum_{j=1}^{n}\sum_{s,t,s',t'}\nu_{sts't'}(4)\left\{M_{j-1}''[v_1'', c_{j-1}^{(\cdot s)}, c_{j-1}^{(\cdot t)}]\right\}\left\{M_{j-1}''[v_2'', c_{j-1}^{(\cdot s')}, c_{j-1}^{(\cdot t')}]\right\} + O_p(h_n) \\
&\to^p \mathbb{V}_{\beta\beta}[v_1'', v_2''].
\end{aligned}$$

The proof is thus complete. □

4.1.3 Verification of the conditions on $\partial_\theta^k\mathbb{G}_n$

Based on (25) and (26), we derive the following bilinear forms:

$$\partial_\alpha\mathbb{G}_n^\alpha(\theta) = \sum_{j=1}^{n}\partial_\alpha M_{j-1}'(\theta)[\chi_j] - h_n\sum_{j=1}^{n}\partial_\alpha M_{j-1}'(\theta)[a_{j-1}(\alpha) - a_{j-1}] - h_n\sum_{j=1}^{n}M_{j-1}'(\theta)\partial_\alpha a_{j-1}(\alpha), \qquad (50)$$

$$\partial_\beta\mathbb{G}_n^\alpha(\theta) = \sum_{j=1}^{n}\partial_\beta M_{j-1}'(\theta)[\chi_j] - h_n\sum_{j=1}^{n}\partial_\beta M_{j-1}'(\theta)[a_{j-1}(\alpha) - a_{j-1}], \qquad (51)$$

$$\partial_\alpha\mathbb{G}_n^\beta(\theta) = -2h_n\sum_{j=1}^{n}\left\{M_{j-1}''(\beta)\partial_\alpha a_{j-1}(\alpha)\right\}[\chi_j - h_n\{a_{j-1}(\alpha) - a_{j-1}\}], \qquad (52)$$

$$\begin{aligned}
\partial_\beta\mathbb{G}_n^\beta(\theta) &= \sum_{j=1}^{n}\left\{\partial_\beta M_{j-1}''(\beta)[\chi_j^{\otimes 2}] - h_n\partial_\beta d_{j-1}(\beta)\right\} - 2h_n\sum_{j=1}^{n}\partial_\beta M_{j-1}''(\beta)[\chi_j, a_{j-1}(\alpha) - a_{j-1}] \\
&\quad + h_n^2\sum_{j=1}^{n}\partial_\beta M_{j-1}''(\beta)\left[\{a_{j-1}(\alpha) - a_{j-1}\}^{\otimes 2}\right]. \qquad (53)
\end{aligned}$$

We can prove the following lemma in a similar way to the proof of Lemma 4.2.

Lemma 4.7. *For every $K > 0$,*

$$\sup_n E_0\left[\sup_\theta\left|\frac{1}{T_n}\partial_\theta^k \mathbb{G}_n(\theta)\right|^K\right] < \infty, \quad k = 1, 2, 3.$$

Recall that the matrix $\mathbb{G}'_\infty(\theta_0) = \mathrm{diag}\{\mathbb{G}'^{\alpha}_\infty(\theta_0), \mathbb{G}'^{\beta}_\infty(\theta_0)\}$ is given by (8) and (9).

Lemma 4.8. *For every $K > 0$,*

$$\sup_{n\in\mathbb{N}} E_0\left[\left|\sqrt{T_n}\left(\frac{1}{T_n}\partial_\theta \mathbb{G}_n(\theta_0) - \mathbb{G}'_\infty(\theta_0)\right)\right|^K\right] < \infty.$$

Proof. First, concerning the off-diagonal parts we have

$$\begin{aligned}
\frac{1}{T_n}\partial_\beta \mathbb{G}_n^\alpha &= \frac{1}{\sqrt{T_n}}\sum_{j=1}^n \frac{1}{\sqrt{T_n}}\partial_\beta M'_{j-1}[\chi_j] = O_p^*\left(\frac{1}{\sqrt{T_n}}\right),\\
\frac{1}{T_n}\partial_\alpha \mathbb{G}_n^\beta &= -2\frac{h_n}{\sqrt{T_n}}\sum_{j=1}^n \frac{1}{\sqrt{T_n}} M''_{j-1}\left[\partial_\alpha a_{j-1}, \chi_j\right] = O_p^*\left(\frac{h_n}{\sqrt{T_n}}\right),
\end{aligned}$$

where the moment estimates for the martingale terms will be proved in an analogous way to the proof of Lemma 4.2. Next, we observe

$$\begin{aligned}
\frac{1}{T_n}\partial_\alpha \mathbb{G}_n^\alpha - \mathbb{G}'^{\alpha}_\infty &= \frac{1}{\sqrt{T_n}}\sum_{j=1}^n \frac{1}{\sqrt{T_n}}\partial_\alpha M'_{j-1}[\chi_j] - \frac{1}{n}\sum_{j=1}^n M'_{j-1}\partial_\alpha a_{j-1} - \mathbb{G}'^{\alpha}_\infty(\theta_0)\\
&= O_p^*\left(\frac{1}{\sqrt{T_n}}\right) + \frac{1}{\sqrt{T_n}}\left\{\sqrt{T_n}\left(-\frac{1}{n}\sum_{j=1}^n M'_{j-1}\partial_\alpha a_{j-1} - \mathbb{G}'^{\alpha}_\infty(\theta_0)\right)\right\}\\
&= O_p^*\left(\frac{1}{\sqrt{T_n}}\right),
\end{aligned}$$

where we used Lemma 4.3 for the last equality. It remains to look at $T_n^{-1}\partial_\beta \mathbb{G}_n^\beta$. Plugging in the identity $\chi_j = m_j + r'_j + h_n^2 R_{j-1}$ and making use of what we have seen in the first half of the proof of Lemma 4.6, we proceed as follows:

$$\begin{aligned}
\frac{1}{T_n}\partial_\beta \mathbb{G}_n^\beta &= \frac{1}{T_n}\sum_{j=1}^n \left(\partial_\beta M''_{j-1}[(m_j + r'_j)^{\otimes 2}] - h_n\partial_\beta d_{j-1}\right) + O_p^*(h_n)\\
&= \frac{1}{T_n}\sum_{j=1}^n \left(\partial_\beta M''_{j-1}[m_j^{\otimes 2}] - h_n\partial_\beta d_{j-1}\right) + O_p^*\left(\sqrt{h_n}\right)\\
&= \frac{1}{\sqrt{T_n}}\left\{\sum_{j=1}^n \frac{1}{\sqrt{T_n}}\left(\partial_\beta M''_{j-1}[m_j^{\otimes 2}] - E_0^{j-1}\left[\partial_\beta M''_{j-1}[m_j^{\otimes 2}]\right]\right)\right\}\\
&\quad + \frac{1}{T_n}\sum_{j=1}^n \left(E_0^{j-1}\left[\partial_\beta M''_{j-1}[m_j^{\otimes 2}]\right] - h_n\partial_\beta d_{j-1}\right) + O_p^*\left(\sqrt{h_n}\right)\\
&= \frac{1}{T_n}\sum_{j=1}^n \left(E_0^{j-1}\left[\partial_\beta M''_{j-1}[m_j^{\otimes 2}]\right] - h_n\partial_\beta d_{j-1}\right) + O_p^*\left(\frac{1}{\sqrt{T_n}}\right)\\
&= \frac{1}{n}\sum_{j=1}^n \left[\mathrm{trace}\left\{\left(-\partial_{\beta_l}\partial_{\beta_{l'}} V_{j-1}^{-1}\right) V_{j-1}\right\} - \partial_{\beta_l}\partial_{\beta_{l'}}\log|V_{j-1}|\right]_{l,l'=1}^{p_\beta} + O_p^*\left(\frac{1}{\sqrt{T_n}}\right).
\end{aligned} \tag{54}$$

The (l, l')th component of the first term in (54) tends in probability to

$$\int \left[\mathrm{trace}\left\{-\partial_{\beta_l}\partial_{\beta_{l'}} V^{-1} V(x, \beta_0)\right\} - \partial_{\beta_l}\partial_{\beta_{l'}}\log|V|(x, \beta_0)\right]\pi_0(dx)$$

$$= -\int \mathrm{trace}\left\{\left(V^{-1}(\partial_{\beta_l} V)V^{-1}(\partial_{\beta_{l'}} V)\right)(x,\beta_0)\right\}\pi_0(dx),$$

which equals the (l,l')th component of $\mathbb{G}'^{\beta}_{\infty}(\theta_0)$. Accordingly, a reduced version of Lemma 4.3 with $\Theta = \{\theta_0\}$ applies to conclude that $T_n^{-1}\partial_\beta \mathbb{G}^{\beta}_n(\theta_0) - \mathbb{G}'^{\beta}_{\infty}(\theta_0) = O_p^*(T_n^{-1/2})$. The proof is complete. □

4.2 Proof of Corollary 2.8

Since $\bar{\alpha}_n := \sqrt{T_n}(\hat{\alpha}_n - \alpha_0) = O_p(1)$ and $\bar{\beta}_n := \sqrt{T_n}(\hat{\beta}_n - \beta_0) = O_p(1)$, it is easy to see from Taylor expansion that $\hat{\mathbb{G}}'^{\alpha}_n \to^p \mathbb{G}'^{\alpha}_{\infty}(\theta_0)$ and $\hat{\mathbb{G}}'^{\beta}_n \to^p \mathbb{G}'^{\beta}_{\infty}(\theta_0)$. As for $\hat{\mathbb{V}}_{\alpha\beta,n}$ and $\hat{\mathbb{V}}_{\beta\beta,n}$, we can substitute $\chi_j(\hat{\alpha}_n) = \chi_j + \sqrt{h_n/n}R_{j-1}[\bar{\alpha}_n]$ in their definitions and apply Taylor expansion as before to derive

$$\hat{\mathbb{V}}_{\alpha\beta,n}[v_1', v_1''] = -\sum_{j=1}^{n}\frac{1}{T_n}\left\{V_{j-1}^{-1}[\partial_\alpha a_{j-1}[v_1'], \chi_j]\right\}\left\{\partial_\beta V_{j-1}^{-1}[v_1'', \chi_j^{\otimes 2}]\right\} + O_p\left(\frac{1}{\sqrt{T_n}}\right), \tag{55}$$

$$\hat{\mathbb{V}}_{\beta\beta,n}[v_1'', v_2''] = \sum_{j=1}^{n}\frac{1}{T_n}\left(\partial_\beta V_{j-1}^{-1}\otimes\partial_\beta V_{j-1}^{-1}\right)[v_1'', \chi_j^{\otimes 2}, v_2'', \chi_j^{\otimes 2}] + O_p\left(\frac{1}{\sqrt{T_n}}\right). \tag{56}$$

We only show that $\hat{\mathbb{V}}_{\alpha\beta,n}[v_1', v_2''] \to^p \mathbb{V}_{\alpha\beta}[v_1', v_1'']$, for the case of $\hat{\mathbb{V}}_{\beta\beta,n}$ is similar.

Write $\sum_{j=1}^n \eta_j$ for the first term in the right-hand side of (55). In view of the Lenglart domination property for martingale $\sum_{j=1}^n(\eta_j - E_0^{j-1}[\eta_j])$ (cf. Jacod and Shiryaev [12, I.3.30]), it suffices to show that $\sum_{j=1}^n E_0^{j-1}[\eta_j] \to^p \mathbb{V}_{\alpha\beta}[v_1', v_1'']$ and $\sum_{j=1}^n E_0[\eta_j^2] \to 0$. But the former can be similarly derived as in what we have seen in the proof of Lemma 4.6. Likewise, noting that $E_0^{j-1}[|\chi_j|^q] \le h_n R_{j-1}$ for every $q \ge 2$, we get $\sum_{j=1}^n E_0[\eta_j^2] \lesssim T_n^{-1} \to 0$, whence $\hat{\mathbb{V}}_{\alpha\beta,n}[v_1', v_1''] \to^p \mathbb{V}_{\alpha\beta}[v_1', v_1'']$.

5 Appendix: A criterion for the exponential ergodicity in dimension one

Here we set $d = r' = r'' = 1$, and suppress dependence on the parameter from the notation:

$$dX_t = a(X_t)dt + b(X_t)dw_t + c(X_{t-})dJ_t. \tag{57}$$

We introduce the following set of conditions.

E1. *(a,b,c) is of class $\mathcal{C}^1(\mathbb{R})$ and globally Lipschitz, and (b,c) is bounded.*

E2. *Either one of the following conditions holds true:*

(i) *$b(x') \neq 0$ for some x', and there exists a constant $\bar{\epsilon} > 0$ such that $\nu(-\epsilon, 0) \wedge \nu(0,\epsilon) > 0$ for every $\epsilon \in (0, \bar{\epsilon})$;*

(ii) *$b \equiv 0$, $c(x'') \neq 0$ for every x'', and we have the decomposition*

$$\nu = \nu_\star + \nu_\natural$$

for two Lévy measures $\nu_\star$ and $\nu_\natural$, where the restriction of $\nu_\star$ to some open set of the form $(-\bar{\epsilon}, 0) \cup (0, \bar{\epsilon})$ admits a continuously differentiable positive density $g_\star$.

E3. (i) *$E[J_1] = 0$ and $\int_{|z|>1}|z|^q\nu(dz) < \infty$ for some $q \ge 1$, and*

$$\limsup_{|x|\to\infty}\frac{a(x)}{x} < 0.$$

(ii) *$E[J_1] = 0$ and $\int_{|z|>1}\exp(q|z|)\nu(dz) < \infty$ for some $q > 0$, and*

$$\limsup_{|x|\to\infty}\mathrm{sgn}(x)a(x) < 0.$$

The next proposition gives a pretty simple criterion for Assumption 2.3.

Proposition 5.1. *The following holds true.*

(a) *Suppose E1, E2, and E3(i). Then, there exist a probability measure π and a constant $a > 0$ such that (3) and (4) hold true for a $\mathcal{C}^2$-function g such that $g(x) = 1 + |x|^q$ outside a neighborhood of the origin.*

(b) *Suppose E1, E2, and E3(ii). Then, there exist a probability measure π and constants $a, \epsilon > 0$ such that (3) and (4) hold true for a $\mathcal{C}^2$-function g such that $g(x) = 1 + \exp(\epsilon|x|)$ outside a neighborhood of the origin.*

Proof. The Lipschitz continuity implies that the SDE (57) admits a unique strong solution. We consider the following conditions:

(I) Every compact sets are petite for some skeleton chain of X (that is, the Markov chain $(X_{j\Delta})_{j\in\mathbb{Z}_+}$);

(II) The exponential Lyapunov-drift criterion (to be fulfilled for some φ)

$$\mathcal{A}\varphi \le -c\varphi + d, \tag{58}$$

where $\mathcal{A}$ denotes the extended generator of X, the function $\varphi : \mathbb{R} \to \mathbb{R}_+$ belonging to the domain of $\mathcal{A}$ satisfies that $\lim_{|x|\to\infty} \varphi(x) = \infty$, and $c, d > 0$ are constants.

As in the proof of Masuda [25, The proof of Theorem 2.2], in each of (a) and (b) the exponential ergodicity (3) follows from (I) and (II), and the moment bound (4) from (II). In order to prove (I), we will here first verify the *Local Doeblin (LD) condition* (see Kulik [16, Appendix A.1] for details); we note that the LD condition implies (I) for any $\Delta > 0$. Then, we will derive the drift condition (58) in (II) with different choices of φ under E3(i) and E3(ii).

Verification of (I): the LD condition.

First, we verify the LD condition when E2(i) is in force. Let $\Pi_x(A) := \nu(\{z \in \mathbb{R} : c(x)z \in A\})$. By Kulik [16, Proposition A.2 and Proposition 4.7]), it suffices to verify the following condition:

$$\forall x \in \mathbb{R}\ \forall v \in \{-1, 1\}\ \exists\rho \in (0, 1)\ \forall\delta > 0 :\ \Pi_x(\{y \in \mathbb{R} : yv \ge \rho|y|\} \cap \{y \in \mathbb{R} : |y| \le \delta\}) > 0.$$

Simple manipulation shows that this conditions is equivalent to the following: for every $x \in \mathbb{R}$ and $\delta > 0$ we have $\nu(\{z \in \mathbb{R} : 0 \le c(x)z \le \delta\}) \wedge \nu(\{z \in \mathbb{R} : -\delta \le c(x)z \le 0\}) > 0$. Since $\nu(\mathbb{R}) > 0$, it suffices to look at x such that $c(x) \not\equiv 0$. However, for such x, the condition obviously holds true under E2(i).

Next we verify the LD condition under E2(ii). If c is constant, then we can apply Kulik [16, Proposition 0.1] to verify the LD condition. Hence we suppose that $\partial_x c \not\equiv 0$ in what follows. First, we smoothly truncate the support of $\nu_\star$ as follows: pick any $\underline{\epsilon} \in (0, \overline{\epsilon})$, let $\psi : \mathbb{R} \to [0, 1]$ be given by [1]

$$\psi(z) := \begin{cases} \exp\left\{-(z-\underline{\epsilon})^{-1} - (\overline{\epsilon} - z)^{-1}\right\} & (\underline{\epsilon} < z < \overline{\epsilon}), \\ 0 & (\text{otherwise}), \end{cases}$$

and set

$$\nu_1(dz) := \{\psi(z) + \psi(-z)\}g_\star(z)dz.$$

Then we have the decomposition $\nu = \nu_1 + \nu_2$, where $\nu_2(dz) := [1 - \{\psi(z) + \psi(-z)\}]\nu_\star(dz) + \nu_\natural(dz)$ defines a Lévy measure. The function $z \mapsto \{\psi(z) + \psi(-z)\}g_\star(z)$ is smooth and supported by $[-\overline{\epsilon}, -\underline{\epsilon}] \cup [\underline{\epsilon}, \overline{\epsilon}]$. With this truncation, we can apply Kulik [16, Proposition A.1]: we have already verified Kulik's condition **S** in the previous paragraph, and it suffices to prove that

$$\exists x'' \in \mathbb{R}\ \exists t'' > 0 :\ P_{x''}\left[\hat{S}_{t''} = \mathbb{R}\right] > 0,$$

where $\hat{S}_t := \left\{u\mathcal{E}^t_\tau c(X_{\tau-});\ u \in \mathbb{R},\ \tau \in \mathcal{D}_1 \cap (0, t)\right\}$, with $\mathcal{D}_1$ and $(\mathcal{E}^t_s)_{0\le s\le t}$ respectively denoting the domain of the point process N_1 associated with ν_1 and the right-continuous solution to

$$\mathcal{E}^t_s = 1 + \int_s^t \partial_x a(X_u)\mathcal{E}^u_s du + \int_s^t \partial_x c(X_{u-})\mathcal{E}^{u-}_s dJ_u.$$

[1] The author owes Professor A. M. Kulik for this clear-cut choice of ψ.

The stochastic-exponential formula leads to

$$\mathcal{E}_s^t = \exp(Y_t - Y_s) \prod_{s<u\le t} (1+\Delta Y_u)\exp(-\Delta Y_u), \quad s \le t,$$

where $Y_u := \int_0^u \partial_x a(X_v)dv + \int_0^u \partial_x c(X_{v-})dJ_v$. We now introduce the two auxiliary sets:

$$A'(t) := \{\omega \in \Omega : \mathcal{D}_1 \cap (0,t) \neq \emptyset\},$$
$$A''(t) := \big\{\omega \in \Omega : \mu\big((0,t], \{z \in \mathbb{R}; |z| \ge 1/\|\partial_x c\|_\infty\}\big) = 0\big\},$$

where $\mu(dt,dz)$ denotes the Poisson random measure associated with J. According to the implications

$$\begin{aligned}\big\{|\Delta J_u| < \|\partial_x c\|_\infty^{-1},\ u \in (0,t]\big\} &\subset \{|\partial_x c(X_{u-})\Delta J_u| < 1,\ u \in (0,t]\} \\ &= \{|\Delta Y_u| < 1,\ u \in (0,t]\} \\ &\subset \big\{\mathcal{E}_s^t \neq 0,\ s \in [0,t]\big\},\end{aligned}$$

the process $(|\mathcal{E}_s^t|)_{0\le s\le t}$ stays positive a.s. on $A''(t)$. Since $P[A'(t) \cap A''(t)] > 0$ for every $t > 0$ and c is supposed to be non-vanishing on $\mathbb{R}$, we observe that for every $x \in \mathbb{R}$ and $t > 0$,

$$\begin{aligned}P_x[\hat{S}_t = \mathbb{R}] &\ge P_x\left[\{\hat{S}_t = \mathbb{R}\} \cap A'(t) \cap A''(t)\right] \\ &\ge P_x\left[\{\mathcal{E}_s^t c(X_{s-}) \neq 0 \text{ for some } s \in (0,t)\} \cap A'(t) \cap A''(t)\right] \\ &= P_x\left[\{c(X_{s-}) \neq 0 \text{ for some } s \in (0,t)\} \cap A'(t) \cap A''(t)\right] \\ &= P_x\left[A'(t) \cap A''(t)\right] > 0,\end{aligned}$$

whence we have verified the LD condition.

Verification of (II): the drift condition.

Now we turn to the verification of (58). For verification under E3(i), one can refer to Kulik [16] and Masuda [25, 26]; in this case, we may set $\varphi(x) = |x|^q$ outside a sufficiently large neighborhood of the origin. It remains to prove (58) under E3(ii), and we will achieve this in a somewhat similar manner to the proof of Masuda [26, Theorem 1.2].

Fix any $\epsilon \in (0, q\|c\|_\infty^{-1} \wedge 1)$ and pick a $\varphi \in \mathcal{C}^2(\mathbb{R})$ such that the following three conditions are in force: (i) $\varphi(x) = \exp(\epsilon|x|)$ for $|x| \ge \epsilon^{-1}$; (ii) $\varphi(x) \le \exp(\epsilon|x|)$ for every x; and (iii) $|\partial_x^2\varphi(x)| \le C\epsilon^2\varphi(x)$ for every x. We can write $\mathcal{A}\varphi = \mathcal{G}\varphi + \mathcal{J}\varphi$, where

$$\mathcal{G}\varphi(x) := \partial_x\varphi(x)a(x) + \frac{1}{2}\partial_x^2\varphi(x)b^2(x),$$
$$\mathcal{J}\varphi(x) := \int \{\varphi(x + c(x)z) - \varphi(x) - \partial_x\varphi(x)c(x)z\}\, \nu(dz)$$

According to the local boundedness of $x \mapsto \mathcal{A}\varphi(x)$, we may and do concentrate on x with $|x|$ large enough. Direct algebra gives

$$\mathcal{G}\varphi(x) \le \epsilon\varphi(x)\,\{\mathrm{sgn}(x)a(x) + C\epsilon\}. \tag{59}$$

Also, by means of Taylor's theorem and the property of φ,

$$\begin{aligned}|\mathcal{J}\varphi(x)| &\lesssim |c(x)|^2 \int |z|^2 \left(\sup_{0\le s\le 1} \left|\partial_x^2\varphi(x + sc(x)z)\right|\right) \nu(dz) \\ &\lesssim \epsilon^2 \exp(\epsilon|x|) \int |z|^2 \exp\left(\epsilon\|c\|_\infty|z|\right) \nu(dz) \\ &\lesssim \epsilon^2\varphi(x).\end{aligned} \tag{60}$$

By putting (59) and (60) together and by taking ϵ small enough, we can find a constant $c_0 > 0$ for which $\mathcal{A}\varphi(x) \le -c_0\varphi(x)$ for every $|x|$ large enough. The proof of Proposition 5.1 is complete. □

Remark 5.2. *If the condition on ν in E2(i) fails to hold, then J is necessarily a compound-Poisson process. In this case, we can utilize the criteria given in Masuda [26].*

Remark 5.3. *Comparing E3(ii) with E3(i), we may say that (58) follows from a weaker condition on the drift function a in compensation for a stronger moment condition on ν.*

Remark 5.4. *By combining the results of the LD-condition argument and general stability theory for Markov processes, it is possible to formulate subexponential- and polynomial-ergodicity versions, as well as the ergodicity version (without rate specification): see e.g. Meyn and Tweedie [29] and Fort and Roberts [6]. Especially, as in Masuda [26], the conditions on (a, b, c) in Proposition 5.1 can be considerably relaxed in case of the ergodicity version.*

Acknowledgement. The author is grateful to Professor A. M. Kulik for his helpful comments concerning the verification of the LD condition appearing in the proof of Proposition 5.1. The present work was partly supported by Grant-in-Aid for Young Scientists (B) 20740061 and 23740082, Japan.

References

[1] Bardet, J. M. and Wintenberger, O. (2009), Asymptotic normality of the quasi-maximum likelihood estimator for multidimensional causal processes. *Ann. Statist.* **37**, 2730–2759.

[2] Bhansali, R. J. and Papangelou, F. (1991), Convergence of moments of least squares estimators for the coefficients of an autoregressive process of unknown order. *Ann. Statist.* **19**, 1155–1162.

[3] Chan, N. H. and Ing, C. K. (2011) Uniform moment bounds of Fisher's information with applications to time series. *Ann. Statist.* **39**, 1526–1550.

[4] Dvoretzky, A. (1977), Asymptotic normality of sums of dependent random vectors. *Multivariate analysis, IV* (Proc. Fourth Internat. Sympos., Dayton, Ohio, 1975), pp. 23–34. North-Holland, Amsterdam.

[5] Findley, D. F. and Wei, C. Z. (2002), AIC, overfitting principles, and the boundedness of moments of inverse matrices for vector autoregressions and related models. *J. Multivariate Anal.* **83**, 415–450.

[6] Fort, G. and Roberts, G. O. (2005), Subgeometric ergodicity of strong Markov processes. *Ann. Appl. Probab.* **15**, 1565–1589.

[7] Friedman, A. (2006), *Stochastic Differential Equations and Applications.* Two volumes bound as one. Dover Publications, Inc., Mineola, NY.

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

[9] Heyde, C. C. (1997), *Quasi-Likelihood and its Application. A General Approach to Optimal Parameter Estimation.* Springer-Verlag, New York.

[10] Ibragimov, I. A. and Has'minski, R. Z. (1981), *Statistical Estimation. Asymptotic Theory.* New York-Berlin.

[11] Inagaki, N. and Ogata, Y. (1975), The weak convergence of likelihood ratio random fields and its applications. *Ann. Inst. Statist. Math.* **27**, 391–419.

[12] Jacod, J. and Shiryaev, A. N. (2003), *Limit Theorems for Stochastic Processes, Second Edition.* Springer-Verlag, Berlin Heidelberg New York.

[13] Jeganathan, P. (1982), On the convergence of moments of statistical estimators. Sankhyā Ser. A **44**, 213–232.

[14] Jeganathan, P. (1989), A note on inequalities for probabilities of large deviations of estimators in nonlinear regression models. *J. Multivariate Anal.* **30**, 227–240.

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

[16] Kulik, A. M., (2009), Exponential ergodicity of the solutions to SDE's with a jump noise. *Stochastic Process. Appl.* **119**, 602–632.

[17] Kulik, A. M. (2011), Asymptotic and spectral properties of exponentially ϕ-ergodic Markov processes. *Stochastic Process. Appl.* **121**, 1044–1075.

[18] Kunita, H. (1990), *Stochastic Flows and Stochastic Differential Equations.* Cambridge University Press, Cambridge.

[19] Kutoyants, Y. A. (1984), *Parameter Estimation for Stochastic Processes.* Heldermann Verlag, Berlin.

[20] Kutoyants, Y. A. (2004), *Statistical Inference for Ergodic Diffusion Processes.* Springer-Verlag London, Ltd., London.

[21] Luschgy, H. and Pagès, G. (2008), Moment estimates for Levy processes. *Electron. Commun. Probab.* **13**, 422–434.

[22] Maruyama, G. and Tanaka, H. (1959), Ergodic property of N-dimensional recurrent Markov processes. *Mem. Fac. Sci. Kyushu Univ. Ser. A* **13**, 157–172.

[23] Mancini, C. (2004), Estimation of the characteristics of the jumps of a general Poisson-diffusion model. *Scand. Actuar. J.* **1**, 42–52.

[24] Masuda, H. (2005), Simple estimators for parametric Markovian trend of ergodic processes based on sampled data. *J. Japan Statist. Soc.* **35**, 147–170.

[25] Masuda, H. (2007), Ergodicity and exponential β-mixing bound for multidimensional diffusions with jumps. *Stochastic Processes Appl.* **117**, 35–56. [Erratum to this paper: (2009) *Stochastic Processes Appl.* **119**, 676–678.]

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

[27] Masuda, H. (2011), Approximate quadratic estimating function for discretely observed Lévy driven SDEs with application to a noise normality test. *RIMS Kôkyûroku* **1752**, 113–131. (The preliminary version is available at MI Preprint Series, Kyushu University, 2010-21)

[28] Menaldi, J. L. and Robin, M. (1999), Invariant measure for diffusions with jumps. *Appl. Math. Optim.* **40**, 105–140.

[29] Meyn, S. P. and Tweedie, R. L. (1993b), Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.* **25**, 518–548.

[30] Ogata, Y. and Inagaki, N. (1977), The weak convergence of the likelihood ratio random fields for Markov observations. *Ann. Inst. Statist. Math.* **29**, 165–187.

[31] Ogihara, T. and Yoshida, N., Quasi-likelihood analysis for the stochastic differential equation with jumps. To appear in *Stat. Inference Stoch. Process.*

[32] Shimizu, Y. and Yoshida, N. (2006), Estimation of parameters for diffusion processes with jumps from discrete observations. *Stat. Inference Stoch. Process.* **9**, 227–277.

[33] Sieders, A. and Dzhaparidze, K. A. (1987), A large deviation result for parameter estimators and its application to nonlinear regression analysis. *Ann. Statist.* **15**, 1031–1049.

[34] Sørensen, M. (2008), Efficient estimation for ergodic diffusions sampled at high frequency. Preprint.

[35] Straumann, D. and Mikosch, T. (2006), Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: a stochastic recurrence equations approach. *Ann. Statist.* **34**, 2449–2495.

[36] Uchida, M. (2010), Contrast-based information criterion for ergodic diffusion processes from discrete observations. *Ann. Inst. Statist. Math.* **62**, 161–187.

[37] Uchida, M. and Yoshida, N. (2010), Adaptive estimation of an ergodic diffusion process based on sampled data. Preprint.

[38] Wang, J. (2008), Criteria for ergodicity of Lévy type operators in dimension one. *Stochastic Process. Appl.* **118**, 1909–1928.

[39] Wedderburn, R. W. M. (1974), Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika* **61**, 439–447.

[40] Yoshida, N. (1992), Estimation for diffusion processes from discrete observation. *J. Multivariate Anal.* **41**, 220–242.

[41] Yoshida, N. (2011), Polynomial type large deviation inequalities and quasi-likelihood analysis for stochastic differential equations. *Ann. Inst. Statist. Math.* **63**, 431–479.

List of MI Preprint Series, Kyushu University

The Global COE Program
Math-for-Industry Education & Research Hub

MI

MI2008-1 Takahiro ITO, Shuichi INOKUCHI & Yoshihiro MIZOGUCHI
Abstract collision systems simulated by cellular automata

MI2008-2 Eiji ONODERA
The intial value problem for a third-order dispersive flow into compact almost Hermitian manifolds

MI2008-3 Hiroaki KIDO
On isosceles sets in the 4-dimensional Euclidean space

MI2008-4 Hirofumi NOTSU
Numerical computations of cavity flow problems by a pressure stabilized characteristic-curve finite element scheme

MI2008-5 Yoshiyasu OZEKI
Torsion points of abelian varieties with values in nfinite extensions over a p-adic field

MI2008-6 Yoshiyuki TOMIYAMA
Lifting Galois representations over arbitrary number fields

MI2008-7 Takehiro HIROTSU & Setsuo TANIGUCHI
The random walk model revisited

MI2008-8 Silvia GANDY, Masaaki KANNO, Hirokazu ANAI & Kazuhiro YOKOYAMA
Optimizing a particular real root of a polynomial by a special cylindrical algebraic decomposition

MI2008-9 Kazufumi KIMOTO, Sho MATSUMOTO & Masato WAKAYAMA
Alpha-determinant cyclic modules and Jacobi polynomials

MI2008-10 Sangyeol LEE & Hiroki MASUDA
Jarque-Bera Normality Test for the Driving Lévy Process of a Discretely Observed Univariate SDE

MI2008-11 Hiroyuki CHIHARA & Eiji ONODERA
A third order dispersive flow for closed curves into almost Hermitian manifolds

MI2008-12 Takehiko KINOSHITA, Kouji HASHIMOTO and Mitsuhiro T. NAKAO
On the L^2 a priori error estimates to the finite element solution of elliptic problems with singular adjoint operator

MI2008-13 Jacques FARAUT and Masato WAKAYAMA
Hermitian symmetric spaces of tube type and multivariate Meixner-Pollaczek polynomials

MI2008-14 Takashi NAKAMURA
Riemann zeta-values, Euler polynomials and the best constant of Sobolev inequality

MI2008-15 Takashi NAKAMURA
Some topics related to Hurwitz-Lerch zeta functions

MI2009-1 Yasuhide FUKUMOTO
Global time evolution of viscous vortex rings

MI2009-2 Hidetoshi MATSUI & Sadanori KONISHI
Regularized functional regression modeling for functional response and predictors

MI2009-3 Hidetoshi MATSUI & Sadanori KONISHI
Variable selection for functional regression model via the L_1 regularization

MI2009-4 Shuichi KAWANO & Sadanori KONISHI
Nonlinear logistic discrimination via regularized Gaussian basis expansions

MI2009-5 Toshiro HIRANOUCHI & Yuichiro TAGUCHII
Flat modules and Groebner bases over truncated discrete valuation rings

MI2009-6 Kenji KAJIWARA & Yasuhiro OHTA
Bilinearization and Casorati determinant solutions to non-autonomous 1+1 dimensional discrete soliton equations

MI2009-7 Yoshiyuki KAGEI
Asymptotic behavior of solutions of the compressible Navier-Stokes equation around the plane Couette flow

MI2009-8 Shohei TATEISHI, Hidetoshi MATSUI & Sadanori KONISHI
Nonlinear regression modeling via the lasso-type regularization

MI2009-9 Takeshi TAKAISHI & Masato KIMURA
Phase field model for mode III crack growth in two dimensional elasticity

MI2009-10 Shingo SAITO
Generalisation of Mack's formula for claims reserving with arbitrary exponents for the variance assumption

MI2009-11 Kenji KAJIWARA, Masanobu KANEKO, Atsushi NOBE & Teruhisa TSUDA
Ultradiscretization of a solvable two-dimensional chaotic map associated with the Hesse cubic curve

MI2009-12 Tetsu MASUDA
Hypergeometric τ-functions of the q-Painlevé system of type $E_8^{(1)}$

MI2009-13 Hidenao IWANE, Hitoshi YANAMI, Hirokazu ANAI & Kazuhiro YOKOYAMA
A Practical Implementation of a Symbolic-Numeric Cylindrical Algebraic Decomposition for Quantifier Elimination

MI2009-14 Yasunori MAEKAWA
On Gaussian decay estimates of solutions to some linear elliptic equations and its applications

MI2009-15 Yuya ISHIHARA & Yoshiyuki KAGEI
Large time behavior of the semigroup on L^p spaces associated with the linearized compressible Navier-Stokes equation in a cylindrical domain

MI2009-16 Chikashi ARITA, Atsuo KUNIBA, Kazumitsu SAKAI & Tsuyoshi SAWABE
Spectrum in multi-species asymmetric simple exclusion process on a ring

MI2009-17 Masato WAKAYAMA & Keitaro YAMAMOTO
Non-linear algebraic differential equations satisfied by certain family of elliptic functions

MI2009-18 Me Me NAING & Yasuhide FUKUMOTO
Local Instability of an Elliptical Flow Subjected to a Coriolis Force

MI2009-19 Mitsunori KAYANO & Sadanori KONISHI
Sparse functional principal component analysis via regularized basis expansions and its application

MI2009-20 Shuichi KAWANO & Sadanori KONISHI
Semi-supervised logistic discrimination via regularized Gaussian basis expansions

MI2009-21 Hiroshi YOSHIDA, Yoshihiro MIWA & Masanobu KANEKO
Elliptic curves and Fibonacci numbers arising from Lindenmayer system with symbolic computations

MI2009-22 Eiji ONODERA
A remark on the global existence of a third order dispersive flow into locally Hermitian symmetric spaces

MI2009-23 Stjepan LUGOMER & Yasuhide FUKUMOTO
Generation of ribbons, helicoids and complex scherk surface in laser-matter Interactions

MI2009-24 Yu KAWAKAMI
Recent progress in value distribution of the hyperbolic Gauss map

MI2009-25 Takehiko KINOSHITA & Mitsuhiro T. NAKAO
On very accurate enclosure of the optimal constant in the a priori error estimates for H_0^2-projection

MI2009-26 Manabu YOSHIDA
Ramification of local fields and Fontaine's property (Pm)

MI2009-27 Yu KAWAKAMI
Value distribution of the hyperbolic Gauss maps for flat fronts in hyperbolic three-space

MI2009-28 Masahisa TABATA
Numerical simulation of fluid movement in an hourglass by an energy-stable finite element scheme

MI2009-29 Yoshiyuki KAGEI & Yasunori MAEKAWA
Asymptotic behaviors of solutions to evolution equations in the presence of translation and scaling invariance

MI2009-30 Yoshiyuki KAGEI & Yasunori MAEKAWA
On asymptotic behaviors of solutions to parabolic systems modelling chemotaxis

MI2009-31 Masato WAKAYAMA & Yoshinori YAMASAKI
Hecke's zeros and higher depth determinants

MI2009-32 Olivier PIRONNEAU & Masahisa TABATA
Stability and convergence of a Galerkin-characteristics finite element scheme of lumped mass type

MI2009-33 Chikashi ARITA
Queueing process with excluded-volume effect

MI2009-34 Kenji KAJIWARA, Nobutaka NAKAZONO & Teruhisa TSUDA
Projective reduction of the discrete Painlevé system of type$(A_2 + A_1)^{(1)}$

MI2009-35 Yosuke MIZUYAMA, Takamasa SHINDE, Masahisa TABATA & Daisuke TAGAMI
Finite element computation for scattering problems of micro-hologram using DtN map

MI2009-36 Reiichiro KAWAI & Hiroki MASUDA
Exact simulation of finite variation tempered stable Ornstein-Uhlenbeck processes

MI2009-37 Hiroki MASUDA
On statistical aspects in calibrating a geometric skewed stable asset price model

MI2010-1 Hiroki MASUDA
Approximate self-weighted LAD estimation of discretely observed ergodic Ornstein-Uhlenbeck processes

MI2010-2 Reiichiro KAWAI & Hiroki MASUDA
Infinite variation tempered stable Ornstein-Uhlenbeck processes with discrete observations

MI2010-3 Kei HIROSE, Shuichi KAWANO, Daisuke MIIKE & Sadanori KONISHI
Hyper-parameter selection in Bayesian structural equation models

MI2010-4 Nobuyuki IKEDA & Setsuo TANIGUCHI
The Itô-Nisio theorem, quadratic Wiener functionals, and 1-solitons

MI2010-5 Shohei TATEISHI & Sadanori KONISHI
Nonlinear regression modeling and detecting change point via the relevance vector machine

MI2010-6 Shuichi KAWANO, Toshihiro MISUMI & Sadanori KONISHI
Semi-supervised logistic discrimination via graph-based regularization

MI2010-7 Teruhisa TSUDA
UC hierarchy and monodromy preserving deformation

MI2010-8 Takahiro ITO
Abstract collision systems on groups

MI2010-9 Hiroshi YOSHIDA, Kinji KIMURA, Naoki YOSHIDA, Junko TANAKA & Yoshihiro MIWA
An algebraic approach to underdetermined experiments

MI2010-10 Kei HIROSE & Sadanori KONISHI
Variable selection via the grouped weighted lasso for factor analysis models

MI2010-11 Katsusuke NABESHIMA & Hiroshi YOSHIDA
Derivation of specific conditions with Comprehensive Groebner Systems

MI2010-12 Yoshiyuki KAGEI, Yu NAGAFUCHI & Takeshi SUDOU
Decay estimates on solutions of the linearized compressible Navier-Stokes equation around a Poiseuille type flow

MI2010-13 Reiichiro KAWAI & Hiroki MASUDA
On simulation of tempered stable random variates

MI2010-14 Yoshiyasu OZEKI
Non-existence of certain Galois representations with a uniform tame inertia weight

MI2010-15 Me Me NAING & Yasuhide FUKUMOTO
Local Instability of a Rotating Flow Driven by Precession of Arbitrary Frequency

MI2010-16 Yu KAWAKAMI & Daisuke NAKAJO
The value distribution of the Gauss map of improper affine spheres

MI2010-17 Kazunori YASUTAKE
On the classification of rank 2 almost Fano bundles on projective space

MI2010-18 Toshimitsu TAKAESU
Scaling limits for the system of semi-relativistic particles coupled to a scalar bose field

MI2010-19 Reiichiro KAWAI & Hiroki MASUDA
Local asymptotic normality for normal inverse Gaussian Lévy processes with high-frequency sampling

MI2010-20 Yasuhide FUKUMOTO, Makoto HIROTA & Youichi MIE
Lagrangian approach to weakly nonlinear stability of an elliptical flow

MI2010-21 Hiroki MASUDA
Approximate quadratic estimating function for discretely observed Lévy driven SDEs with application to a noise normality test

MI2010-22 Toshimitsu TAKAESU
A Generalized Scaling Limit and its Application to the Semi-Relativistic Particles System Coupled to a Bose Field with Removing Ultraviolet Cutoffs

MI2010-23 Takahiro ITO, Mitsuhiko FUJIO, Shuichi INOKUCHI & Yoshihiro MIZOGUCHI
Composition, union and division of cellular automata on groups

MI2010-24 Toshimitsu TAKAESU
A Hardy's Uncertainty Principle Lemma in Weak Commutation Relations of Heisenberg-Lie Algebra

MI2010-25 Toshimitsu TAKAESU
On the Essential Self-Adjointness of Anti-Commutative Operators

MI2010-26 Reiichiro KAWAI & Hiroki MASUDA
On the local asymptotic behavior of the likelihood function for Meixner Lévy processes under high-frequency sampling

MI2010-27 Chikashi ARITA & Daichi YANAGISAWA
Exclusive Queueing Process with Discrete Time

MI2010-28 Jun-ichi INOGUCHI, Kenji KAJIWARA, Nozomu MATSUURA & Yasuhiro OHTA
Motion and Bäcklund transformations of discrete plane curves

MI2010-29 Takanori YASUDA, Masaya YASUDA, Takeshi SHIMOYAMA & Jun KOGURE
On the Number of the Pairing-friendly Curves

MI2010-30 Chikashi ARITA & Kohei MOTEGI
Spin-spin correlation functions of the q-VBS state of an integer spin model

MI2010-31 Shohei TATEISHI & Sadanori KONISHI
Nonlinear regression modeling and spike detection via Gaussian basis expansions

MI2010-32 Nobutaka NAKAZONO
Hypergeometric τ functions of the q-Painlevé systems of type $(A_2 + A_1)^{(1)}$

MI2010-33 Yoshiyuki KAGEI
Global existence of solutions to the compressible Navier-Stokes equation around parallel flows

MI2010-34 Nobushige KUROKAWA, Masato WAKAYAMA & Yoshinori YAMASAKI
Milnor-Selberg zeta functions and zeta regularizations

MI2010-35 Kissani PERERA & Yoshihiro MIZOGUCHI
Laplacian energy of directed graphs and minimizing maximum outdegree algorithms

MI2010-36 Takanori YASUDA
CAP representations of inner forms of $Sp(4)$ with respect to Klingen parabolic subgroup

MI2010-37 Chikashi ARITA & Andreas SCHADSCHNEIDER
Dynamical analysis of the exclusive queueing process

MI2011-1 Yasuhide FUKUMOTO& Alexander B. SAMOKHIN
Singular electromagnetic modes in an anisotropic medium

MI2011-2 Hiroki KONDO, Shingo SAITO & Setsuo TANIGUCHI
Asymptotic tail dependence of the normal copula

MI2011-3 Takehiro HIROTSU, Hiroki KONDO, Shingo SAITO, Takuya SATO, Tatsushi TANAKA & Setsuo TANIGUCHI
Anderson-Darling test and the Malliavin calculus

MI2011-4 Hiroshi INOUE, Shohei TATEISHI & Sadanori KONISHI
Nonlinear regression modeling via Compressed Sensing

MI2011-5 Hiroshi INOUE
Implications in Compressed Sensing and the Restricted Isometry Property

MI2011-6 Daeju KIM & Sadanori KONISHI
Predictive information criterion for nonlinear regression model based on basis expansion methods

MI2011-7 Shohei TATEISHI, Chiaki KINJYO & Sadanori KONISHI
Group variable selection via relevance vector machine

MI2011-8 Jan BREZINA & Yoshiyuki KAGEI
Decay properties of solutions to the linearized compressible Navier-Stokes equation around time-periodic parallel flow
Group variable selection via relevance vector machine

MI2011-9 Chikashi ARITA, Arvind AYYER, Kirone MALLICK & Sylvain PROLHAC
Recursive structures in the multispecies TASEP

MI2011-10 Kazunori YASUTAKE
On projective space bundle with nef normalized tautological line bundle

MI2011-11 Hisashi ANDO, Mike HAY, Kenji KAJIWARA & Tetsu MASUDA
An explicit formula for the discrete power function associated with circle patterns of Schramm type

MI2011-12 Yoshiyuki KAGEI
Asymptotic behavior of solutions to the compressible Navier-Stokes equation around a parallel flow

MI2011-13 Vladimír CHALUPECKÝ & Adrian MUNTEAN
Semi-discrete finite difference multiscale scheme for a concrete corrosion model: approximation estimates and convergence

MI2011-14 Jun-ichi INOGUCHI, Kenji KAJIWARA, Nozomu MATSUURA & Yasuhiro OHTA
Explicit solutions to the semi-discrete modified KdV equation and motion of discrete plane curves

MI2011-15 Hiroshi INOUE
A generalization of restricted isometry property and applications to compressed sensing

MI2011-16 Yu KAWAKAMI
A ramification theorem for the ratio of canonical forms of flat surfaces in hyperbolic three-space

MI2011-17 Naoyuki KAMIYAMA
Matroid intersection with priority constraints

MI2012-1 Kazufumi KIMOTO & Masato WAKAYAMA
Spectrum of non-commutative harmonic oscillators and residual modular forms

MI2012-2 Hiroki MASUDA
Mighty convergence of the Gaussian quasi-likelihood random fields for ergodic Levy driven SDE observed at high frequency