Edgeworth expansion for Euler approximation of continuous diffusion processes

Mark Podolskij* Bezirgen Veliyev† Nakahiro Yoshida‡

November 20, 2018

Abstract

In this paper we present the Edgeworth expansion for the Euler approximation scheme of a continuous diffusion process driven by a Brownian motion. Our methodology is based upon a recent work [22], which establishes Edgeworth expansions associated with asymptotic mixed normality using elements of Malliavin calculus. Potential applications of our theoretical results include higher order expansions for weak and strong approximation errors associated to the Euler scheme, and for studentized version of the error process.

Keywords: diffusion processes, Edgeworth expansion, Euler scheme, limit theorems.

AMS 2000 subject classifications. 60F05, 60H10, 65C30.

1 Introduction

In this work we consider a one-dimensional continuous stochastic process $(X_t)_{t\in[0,1]}$ that satisfies the stochastic differential equation

$$dX_t = a(X_t)dt + b(X_t)dW_t \qquad \text{with} \qquad X_0 = x_0, \tag{1.1}$$

where $(W_t)_{t\in[0,1]}$ is a Brownian motion, defined on a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, \mathbb{P}\right)$. A simple and effective numerical scheme for the solution of (1.1) is the Euler approximation scheme, which is given as follows. Let $\varphi_n : \mathbb{R}_+ \to \mathbb{R}_+$ be the function defined by $\varphi_n(t) = i/n$ when $t \in [i/n, (i+1)/n)$. The continuous Euler approximation scheme is described by

$$dX_t^n = a\left(X^n_{\varphi_n(t)}\right)dt + b\left(X^n_{\varphi_n(t)}\right)dW_t \qquad \text{with} \qquad X_0^n = x_0. \tag{1.2}$$

*Department of Mathematics, Aarhus University, Ny Munkegade 118, 8000 Aarhus, Denmark, Email: mpodolskij@math.au.dk.

†CREATES, Department of Economics and Business Economics, Aarhus University, Fuglesangs Alle 4, 8210 Aarhus V, Denmark, Email: bveliyev@econ.au.dk

‡Graduate School of Mathematical Science, 3-8-1 Komaba, Meguro-ku, Tokyo 153, Japan, Email: nakahiro@ms.u-tokyo.ac.jp

The probabilistic properties of the Euler approximation scheme have been investigated in numerous papers. We refer to the classical work [3, 4, 9, 11, 12, 13] among many others. Asymptotic results in the framework of non-regular coefficients can be found in e.g. [2, 6, 8, 19].

In this paper we are aiming to derive an Edgeworth expansion for the error process

$$U^n = X^n - X. \tag{1.3}$$

Let us recall the classical convergence result for $(U_t^n)_{t\in[0,1]}$ from [9].

Theorem 1.1. *[9, Theorem 1.2] Assume that the functions a, b are globally Lipschitz and $a, b \in C^1(\mathbb{R})$. Then we obtain the stable convergence*

$$V^n := \sqrt{n} U^n \xrightarrow{d_{st}} V \qquad on\ C([0,1]) \tag{1.4}$$

equipped with the uniform topology, where $V = (V_t)_{t\in[0,1]}$ is the unique solution of the stochastic differential equation

$$dV_t = a'(X_t)V_t dt + b'(X_t)V_t dW_t - \frac{1}{\sqrt{2}} bb'(X_t) dB_t \qquad with \qquad V_0 = 0, \tag{1.5}$$

and $(B_t)_{t\in[0,1]}$ is a new Brownian motion defined on an extension of the probability space $\big(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, \mathbb{P}\big)$ and independent of the σ-field $\mathcal{F}$.

We will see later that the limiting process V is an $\mathcal{F}$-conditional Gaussian martingale with $\mathcal{F}$-conditional zero mean. In particular, for each $t > 0$, V_t has a mixed normal distribution. The aim of this work is to derive an Edgeworth expansion associated with Theorem 1.1. More specifically, for any regular q-dimensional random variable F and any given times $0 < T_1 < \ldots < T_k \le 1$, we would like to determine the function $p_n : \mathbb{R}^k \times \mathbb{R}^q \to \mathbb{R}$ such that it holds

$$\sup_{f\in\mathcal{C}_{q,k}} \left| \mathbb{E}[f(V_{T_1}^n, \ldots, V_{T_k}^n, F)] - \int_{\mathbb{R}^k\times\mathbb{R}^q} f(z,x) p_n(z,x) dz dx \right| = o(1/\sqrt{n}) \tag{1.6}$$

for a large class of functions $\mathcal{C}_{q,k}$. The methodology is based upon the work of Yoshida [22], which applies Malliavin calculus and stable convergence to obtain the Edgeworth expansion associated with mixed normal limits. Another key ingredient in the derivation of (1.6) is the stochastic expansion of the error process U^n and a non-degeneracy condition, which turns out to be rather complex in the case $k > 1$. Related articles include [7, 15, 16], which have studied Edgeworth expansions associated to covariance estimators, power variations and the pre-averaging estimator.

The paper is structured as follows. Section 2.1 presents various definitions and notation. Sections 2.2 and 2.3 are devoted to derivation of Edgeworth expansion for multivariate weighted quadratic functionals, which plays a crucial role in the asymptotic analysis of the Euler scheme. In Section 3 we investigate the second order stochastic expansion of the standardised error process associated with the Euler approximation scheme. The Edgeworth expansion for the error process is investigated in Section 4. Section 5 is devoted to several applications of our theoretical results, including asymptotic expansion of the weak and strong approximation errors, and density expansion for the studentized version of the error process. Some proofs are presented in Section 6.

2 Background

2.1 Definitions and notation

In this subsection we introduce basic notation, some elements of Malliavin calculus and the definition of stable convergence in law.

All vectors $x \in \mathbb{R}^k$ are understood as column vectors; $\|x\|$ stands for Euclidean norm of x and $x^\star$ denotes the transpose of x. For $x \in \mathbb{R}^k$ and $m \in \mathbb{Z}_+^k$ we set $x^m := \prod_{j=1}^k x_j^{m_j}$ and $|m| = \sum_{j=1}^k m_j$. For any function $f : \mathbb{R} \to \mathbb{R}$ we denote by $f^{(l)}$ its lth derivative; for a function $f : \mathbb{R}^k \times \mathbb{R}^q \to \mathbb{R}$ and $\alpha = (\alpha_1, \alpha_2) \in \mathbb{Z}_+^k \times \mathbb{Z}_+^q$ the operator d^α is defined via $d^\alpha = d_{x_1}^{\alpha_1} d_{x_2}^{\alpha_2}$. The set $C_p^l(\mathbb{R}^k)$ (resp. $C_b^l(\mathbb{R}^k)$) denotes the space of l times continuously differentiable functions $f : \mathbb{R}^k \to \mathbb{R}$ such that all derivatives up to order l have polynomial growth (resp. are bounded). For a matrix $A \in \mathbb{R}^{k\times k}$ and a vector $x \in \mathbb{R}^k$ we write $A[x^{\otimes 2}]$ to denote the quadratic form $x^\star A x$; similarly, for $x, y \in \mathbb{R}^k$ we write $y[x]$ for the linear form $y^\star x$. Finally, $\mathtt{i} := \sqrt{-1}$.

We now introduce some notions of Malliavin calculus (we refer to the books of Ikeda and Watanabe [18] and Nualart [14] for a detailed exposition of Malliavin calculus). The set $\mathbb{L}^p$ denotes the space of random variables with finite pth moment and we use the notation $\mathbb{L}_{\infty-} = \cap_{p>1}\mathbb{L}^p$; the corresponding $\mathbb{L}^p$-norms are denoted by $\|\cdot\|_{\mathbb{L}^p}$. Define $\mathbb{H} = \mathbb{L}^2([0,1], dx)$ and let $\langle \cdot, \cdot \rangle_{\mathbb{H}}$ denote the usual scalar product on $\mathbb{H}$. We denote by D^l the lth Malliavin derivative operator and by δ^l its unbounded adjoint (also called Skrokhod integral of order l). The space $\mathbb{D}_{l,p}$ is the completion of the set of smooth random variables with respect to the norm

$$\|Y\|_{l,p} := \left(\mathbb{E}[|Y|^p] + \sum_{m=1}^{l} \mathbb{E}[\|D^m Y\|^p_{\mathbb{H}^{\otimes m}}] \right)^{1/p}.$$

For any smooth k-dimensional random variable Y the Malliavin matrix is defined via $\sigma_Y := (\langle DY_i, DY_j \rangle_{\mathbb{H}})_{1\le i,j\le k}$. We write $\Delta_Y := \det \sigma_Y$ for the determinant of the Malliavin matrix. Finally, we set $\mathbb{D}_{l,\infty} = \cap_{p\ge 2}\mathbb{D}_{l,p}$. We sometimes use the notation $\mathbb{D}_{l,p}(\mathbb{R}^k)$ to denote the space of all k-dimensional random variable Y such that $Y_i \in \mathbb{D}_{l,p}$.

We use the notation $Y_n \xrightarrow{d_{st}} Y$ to denote the stable convergence in law. We recall that a sequence of random variables $(Y_n)_{n\in\mathbb{N}}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in a metric space E is said to converge stably with limit Y, written $Y_n \xrightarrow{d_{st}} Y$, where Y is defined on an extension $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ of the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$, iff for any bounded, continuous function g and any bounded $\mathcal{F}$-measurable random variable Z it holds that

$$\mathbb{E}[g(Y_n)Z] \to \overline{\mathbb{E}}[g(Y)Z], \quad n \to \infty. \tag{2.1}$$

The notion of stable convergence is due to Renyi [17]. We also refer to [1] for properties of this mode of convergence.

Finally, for two vector fields V_0 and V_1 we denote by $\mathrm{Lie}[V_0; V_1]$ the Lie algebra generated by V_1 and V_0. That is, $\mathrm{Lie}[V_0; V_1] = \mathrm{span}\left(\bigcup_{j=0}^\infty \Sigma_j\right)$, where $\Sigma_0 = \{V_1\}$ and

$\Sigma_j = \{[V, V_i];\ V \in \Sigma_{j-1},\ i = 0, 1\}$ $(j \geq 1)$ with the Lie bracket $[\cdot, \cdot]$. $\mathrm{Lie}[V_0; V_1](x)$ stands for $\mathrm{Lie}[V_0; V_1]$ evaluated at x.

2.2 Edgeworth expansion associated with mixed normal limits: The quadratic case

In this subsection we will study the (second order) Edgeworth expansion associated with certain quadratic functionals of Brownian motion that will be crucial for the treatment of the error process V^n. Indeed, we will see later that the dominating martingale term in the expansion of V^n has a quadratic form. The results are similar in spirit to [22, Theorem 4], but we will require quite different non-degeneracy arguments.

On a filtered Wiener space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, \mathbb{P})$ we consider a k-dimensional random functional Z_n, which admits the decomposition

$$Z_n = M_n + n^{-1/2} N_n, \tag{2.2}$$

where M_n and N_n are tight sequences of random variables. We assume that M_n, which will have a quadratic form, converges stably in law to a mixed normal variable M:

$$M_n \xrightarrow{d_{st}} M, \tag{2.3}$$

where the random variable M is defined on an extension $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ of the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and, conditionally on $\mathcal{F}$, M has a normal law with mean 0 and conditional covariance matrix $C \in \mathbb{R}^{k\times k}$. In this case we use the notation

$$M \sim MN(0, C).$$

For concrete applications it is often useful to consider the Edgeworth expansion for the pair (Z_n, F_n), where F_n is another q-dimensional random functional satisfying the convergence in probability

$$F_n \xrightarrow{\mathbb{P}} F.$$

Obviously, such a framework is important when the statistic at hand does not only depend on the sequence Z_n, but also on an external random variable F (in this case we may set $F_n = F$). In the statistical context the most useful application is the case where $F_n \xrightarrow{\mathbb{P}} C$. In this situation we obtain by properties of stable convergence that

$$F_n^{-1/2} Z_n \xrightarrow{d} \mathcal{N}_k(0, \mathrm{id}_k)$$

when $F_n \in \mathbb{R}^{k\times k}$ is positive definite and id_k denotes the identity matrix. Thus, the asymptotic expansion of the law of (Z_n, F_n) would imply the Edgeworth expansion for the studentized statistic $F_n^{-1/2} Z_n$.

In the next step we embed the previous static framework into a martingale setting. We assume that the leading term M_n is a terminal value of some continuous $(\mathcal{F}_t)$-martingale

$(M_t^n)_{t\in[0,1]}$, that is $M_n = M_1^n$. We also consider stochastic processes $(M_t)_{t\in[0,1]}$ and $(C_t^n)_{t\in[0,1]}$ with values in $\mathbb{R}^k$ and $\mathbb{R}^{k\times k}$ respectively, such that

$$M = M_1, \qquad C_t = \langle M\rangle_t, \qquad C_t^n = \langle M^n\rangle_t, \qquad C_n = \langle M^n\rangle_1. \tag{2.4}$$

Here the process $(M_t)_{t\in[0,1]}$, defined on extended probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$, represents the stable limit of the continuous $(\mathcal{F}_t)$-martingale $(M_t^n)_{t\in[0,1]}$, while C^n denotes the quadratic covariation process associated with M^n.

Now, we shall introduce a particular type of quadratic functionals. For a sequence of time points $(T_j)_{1\le j\le k}$ not depending on n with $0 < T_1 < \ldots < T_k$, we consider a sequence of partitions $\pi^n = (t_i)_{1\le i\le m_n}$ of $[0,1]$ such that $0 = t_0 < t_1 < \ldots < t_{m_n}$ and that $\{T_j\}_{1\le j\le k} \subset \{t_i\}_{1\le i\le m_n}$ for every $n\in\mathbb{N}$. Here t_j may depend on n though we omit n for notational simplicity. Let $I_i = [t_{i-1}, t_i)$ and $|I_i| = t_i - t_{i-1}$. Suppose that $n^4\sum_{i=1}^{m_n}|I_i|^5 = O(1)$ as $n\to\infty$. Next, we consider a strongly predictable kernel $K^n = (K^{n,j})_{1\le j\le k} : \Omega\times[0,1]\to\mathbb{R}^k$ satisfying

$$K^{n,j}(t) = K^{n,j}(t_{i-1}) \text{ for } t\in I_i \qquad \text{and} \qquad K^{n,j}(t) = 0 \text{ if } t\ge T_j.$$

The aforementioned sequence of quadratic type martingales $M^n = (M^{n,j})_{1\le j\le k}$ is defined by

$$M_t^{n,j} = \sqrt{n}\sum_{i=1}^{m_n} K^{n,j}(t_{i-1})\int_{t_{i-1}\wedge t}^{t_i\wedge t}\int_{t_{i-1}}^{s} dW_r dW_s, \qquad t\in[0,1]. \tag{2.5}$$

Let $K : \Omega\times[0,1]\to\mathbb{R}$ be a continuous adapted process and set

$$\mathbb{I}_s^j = \frac{1}{2}\int_s^{T_j} K(r)^2 dr, \qquad s\in(T_{j-1}, T_j]. \tag{2.6}$$

Our first set of conditions relates the kernel K^n to K and introduces some integrability assumptions, which are similar in spirit to assumptions imposed in [22]. Recall that $F\in\mathbb{R}^q$, set $\ell = k+q+8$ and let $\frac{1}{3} < d < \frac{1}{2}$.

(B1) **(i)** $K^n(t)\in\mathbb{D}_{\ell+1,\infty}(\mathbb{R}^k)$ and there exists a density $D_{r_1,\ldots,r_m}K^n(t)$ representing each derivative such that

$$\sup_{\substack{r_1,\ldots,r_m\in(0,1),\\ t\in[0,1],\, n\in\mathbb{N}}} \left\|D_{r_1,\ldots,r_m}K^n(t)\right\|_{\mathbb{L}^p} \;<\; \infty$$

for every $p>1$ and $m = 0,1,\ldots,\ell+1$.

(ii) For every $p>1$ and $j = 1,\ldots,k$,

$$\sup_{1\le i\le m_n}\sup_{t\in(t_{i-1},t_i)}\left\|K^{n,j}(t) - K(t)1_{\{t<T_j\}}\right\|_{\ell,p} \;=\; O(n^{-d})$$

as $n\to\infty$.

(iii) For every $p > 1$ and $j = 1, ..., k$,

$$\sup_{s\in(T_{j-1},T_j)} \left\| \left[\frac{\mathbb{I}_s^j}{T_j - s} \right]^{-1} \right\|_{\mathbb{L}^p} \quad < \quad \infty.$$

From (B1)(i), (ii) we deduce that

$$C_t^{n,j} = \langle M^{n,j} \rangle_t \xrightarrow{\mathbb{P}} C_t^j = \frac{1}{2} \int_0^{t\wedge T_j} K(s)^2 ds.$$

Furthermore, (B1)(iii) implies

$$\left(C_{T_j}^j - C_{T_{j-1}}^j \right)^{-1} \quad \in \quad \mathbb{L}_{\infty-} \tag{2.7}$$

for $j = 1, \ldots, k$. In particular, $\det C^{-1} \in \mathbb{L}_{\infty-}$ for $C = (C_1^{j_1\wedge j_2})_{1\leq j_1,j_2\leq k}$.

Now, let us set

$$\widehat{C}_n = \sqrt{n}(C_n - C), \qquad \widehat{F}_n = \sqrt{n}(F_n - F), \tag{2.8}$$

where $C_n = C_1^n$ with $C_t^n = (\langle M^{n,j_1}, M^{n,j_2} \rangle)_{1\leq j_1,j_2\leq k}$. In the validation of the asymptotic expansion a truncation functional $s_n : \Omega \to \mathbb{R}^k$ will play an important role; see Section 6.4 for its explicit definition. We set $\ell_* = 2[q/2] + 4$ and present the next set of assumptions that determines the asymptotic distribution of the vector $(M_\cdot^n, N_n, \widehat{C}_n, \widehat{F}_n)$ along with some new integrability conditions.

(B2) **(i)** $F \in \mathbb{D}_{\ell+1,\infty}(\mathbb{R}^q)$, $\sup_{r_1,...,r_m\in(0,1)} \|D_{r_1,...,r_m}F\|_{\mathbb{L}^p} < \infty$ for every $p > 1$ and $m = 1, ..., \ell + 1$. Moreover $r \mapsto D_r F$ and $(r, s) \mapsto D_{r,s}F$ $(r \leq s)$ are continuous a.s.

(ii) $F_n \in \mathbb{D}_{\ell+1,\infty}(\mathbb{R}^q)$, $N_n \in \mathbb{D}_{\ell+1,\infty}(\mathbb{R}^k)$ and $s_n = (s_n^j) \in \mathbb{D}_{\ell,\infty}(\mathbb{R}^k)$. Moreover,

$$\sup_{n\in\mathbb{N}} \left\{ \|\widehat{C}_n\|_{\ell,p} + \|\widehat{F}_n\|_{\ell+1,p} + \|N_n\|_{\ell+1,p} + \|s_n\|_{\ell,p} \right\} \quad < \quad \infty$$

for every $p > 1$.

(iii) $(M^n, N_n, \widehat{C}_n, \widehat{F}_n) \xrightarrow{d_{st}} (M, N, \widehat{C}, \widehat{F})$ for a random vector $(M, N, \widehat{C}, \widehat{F})$ defined on an extension of $(\Omega, \mathcal{F}, P)$.

(iv) For $u \in \mathbb{R}^k$ and $v \in \mathbb{R}^q$, the conditional expectations $\mathbb{E}\big[\widehat{C}|M_1 = z\big][u^{\otimes 2}]$, $\mathbb{E}\big[\widehat{F}|M_1 = z\big][v]$ and $\mathbb{E}\big[N|M_1 = z\big]$ are in the polynomial ring $\mathbb{D}_{\ell_*,\infty}(\mathbb{R})[z]$ (with coefficients in $\mathbb{D}_{\ell_*,\infty}(\mathbb{R})$).

Finally, we will require a non-degeneracy condition on the pair (M_t^n, F). Let us introduce the process

$$\mathbb{X}_t^j \quad = \quad (M_1^{n,1}, ..., M_1^{n,j-1}, M_t^{n,j}, F).$$

(B3) **(i)** For each $j = 1, ..., k$, there exists a sequence $(\tau_n^j)_{n\in\mathbb{N}} \subset (T_{j-1}, T_j)$ such that $\sup_n \tau_n^j < T_j$ and that

$$\sup_{t\in[\tau_n^j, T_j]} \mathbb{P}\big[\det \sigma_{\mathbb{X}_t^j} < s_n^j\big] \quad = \quad O(n^{-\nu})$$

for some $\nu > \ell/6$.

(ii) $\limsup_{n\to\infty} \mathbb{E}\big[(s_n^j)^{-p}\big] < \infty$ for every $p > 1$ and $j = 1, \dots, k$.

2.3 Random symbols $\underline{\sigma}, \overline{\sigma}$ and the main result

In order to present the Edgeworth expansion for the pair (Z_n, F_n) we need to define two *random symbols* $\underline{\sigma}$ and $\overline{\sigma}$, which play a crucial role in what follows. We call $\underline{\sigma}$ the adaptive (or classical) random symbol and $\overline{\sigma}$ the anticipative random symbol. The adaptive random symbol $\underline{\sigma}$ is defined by

$$\underline{\sigma}(z; \mathtt{i}u, \mathtt{i}v) \quad = \quad \frac{1}{2}\mathbb{E}\big[\widehat{C}|M_1 = z\big][(\mathtt{i}u)^{\otimes 2}] + \mathbb{E}\big[N|M_1 = z\big][\mathtt{i}u] + \mathbb{E}\big[\widehat{F}|M_1 = z\big][\mathtt{i}v].$$

Let $\overline{K}(t) = (K(1)1_{\{t<T_j\}})_{1\le j\le k}$. The anticipative random symbol $\overline{\sigma}$ is defined by

$$\overline{\sigma}(iu, iv) \quad = \quad \frac{1}{2}\int_0^1 \overline{K}(t)[\mathtt{i}u]\sigma_t(\mathtt{i}u, \mathtt{i}v)dt \tag{2.9}$$

where

$$\begin{aligned}\sigma_t(\mathtt{i}u, \mathtt{i}v) \quad &= \quad \Big(-\frac{1}{2}D_tC[u^{\otimes 2}] + D_tF[\mathtt{i}v]\Big)^2 \\ &\quad + \Big(-\frac{1}{2}D_tD_tC[u^{\otimes 2}] + D_tD_tF[\mathtt{i}v]\Big).\end{aligned}$$

The derivative D_tD_t stands for $\lim_{s\uparrow t} D_sD_t$. The full random symbol is defined by

$$\sigma \quad = \quad \underline{\sigma} + \overline{\sigma}$$

and has the form

$$\sigma(z; \mathtt{i}u, \mathtt{i}v) \quad = \quad \sum_\alpha c_\alpha(z)(\mathtt{i}u)^{\alpha_1}(\mathtt{i}v)^{\alpha_2} \tag{2.10}$$

where $\alpha = (\alpha_1, \alpha_2) \in \mathbb{Z}_+^k \times \mathbb{Z}_+^q$.

Under conditions of the previous subsection, the non-degeneracy of F is ensured and it has a differentiable density function p^F. Thus, the following function p_n is well defined:

$$\begin{aligned}p_n(z, x) &= \mathbb{E}\big[\phi(z; 0, C)|F = x\big]p^F(x) \\ &\quad + n^{-1/2}\sum_\alpha(-d_z)^{\alpha_1}(-d_x)^{\alpha_2}\Big\{\mathbb{E}\big[c_\alpha(z)\phi(z; 0, C)|F = x\big]p^F(x)\Big\}.\end{aligned} \tag{2.11}$$

For positive numbers R and γ, $\mathcal{E}(R,\gamma)$ denotes the set of measurable functions $f:\mathbb{R}^{k+q}\to\mathbb{R}$ such that $|f(z,x)|\le R(1+|z|+|x|)^\gamma$ for all $z\in\mathbb{R}^k$ and $x\in\mathbb{R}^q$. The error of the approximation of the distribution of (Z_n,F_n) by p_n is evaluated by the quantity

$$\Delta_n(f) \;=\; \left|\mathbb{E}\big[f(Z_n,F_n)\big]-\int f(z,x)p_n(z,x)dzdx\right|$$

for $f\in\mathcal{E}(R,\gamma)$. The main result of this section is the following.

Theorem 2.1. *Suppose that Z_n is given by (2.2) with M_n defined by (2.5). Suppose that $(B1)$, $(B2)$ and $(B3)$ are satisfied. Then*

$$\sup_{f\in\mathcal{E}(R,\gamma)}\Delta_n(f)=o(n^{-1/2}) \tag{2.12}$$

as $n\to\infty$ for any positive numbers R and γ.

3 Stochastic expansion of the error process

In this section we derive explicit expressions for the first and second order approximation of the normalised error process V^n. The following well known lemma, which presents an explicit solution of an affine stochastic differential equation, will be a helpful tool.

Lemma 3.1. *Assume that $(Y_t)_{t\in[0,1]}$ is the unique strong solution of the stochastic differential equation*

$$dY_t=(c_tY_t+\tilde{c}_t)dt+(d_tY_t+\tilde{d}_t)dW_t \qquad \text{with} \qquad Y_0=y_0, \tag{3.1}$$

where $(c_t)_{t\in[0,1]},(\tilde{c}_t)_{t\in[0,1]},(d_t)_{t\in[0,1]},(\tilde{d}_t)_{t\in[0,1]}$ are predictable stochastic processes. Then the process $(Y_t)_{t\in[0,1]}$ exhibits an explicit solution given by

$$Y_t=\Sigma_t\left[y_0+\int_0^t\Sigma_s^{-1}\Big((\tilde{c}_s-d_s\tilde{d}_s)ds+\tilde{d}_sdW_s\Big)\right], \tag{3.2}$$

$$\Sigma_t=\exp\left(\int_0^t d_sdW_s+\int_0^t\left(c_s-\frac{1}{2}d_s^2\right)ds\right).$$

Proof. The proof follows a classical route for solutions of inhomogeneous differential equations. First, we recall that the process Σ satisfies the stochastic differential equation $d\Sigma_t=c_t\Sigma_tdt+d_t\Sigma_tdW_t$, which is shown by applying Itô's formula to the function $f(x,y)=\exp(x+y)$. Now, setting $Z_t=y_0+\int_0^t\Sigma_s^{-1}(\tilde{c}_s-d_s\tilde{d}_s)ds+\int_0^t\Sigma_s^{-1}\tilde{d}_sdW_s$, we conclude by the product formula that

$$\begin{aligned}
Y_t&=y_0+\int_0^t\Sigma_sdZ_s+\int_0^tZ_sd\Sigma_s+\langle Z,\Sigma\rangle_t\\
&=y_0+\int_0^t(\tilde{c}_s-d_s\tilde{d}_s)ds+\int_0^t\tilde{d}_sdW_s+\int_0^tc_sZ_s\Sigma_sds+\int_0^td_sZ_s\Sigma_sdW_s+\int_0^td_s\tilde{d}_sds\\
&=y_0+\int_0^t(c_sY_s+\tilde{c}_s)ds+\int_0^t(d_sY_s+\tilde{d}_s)ds
\end{aligned}$$

and the proof is complete. □

Applying the same type of proof as in Lemma 3.1, we deduce that the limiting process V introduced at (1.5) can be written explicitly as

$$V_t = -\frac{1}{\sqrt{2}}\Sigma_t \int_0^t \Sigma_s^{-1} bb'(X_s) dB_s,$$

where the process $(\Sigma_t)_{t\geq 0}$ is defined by

$$\Sigma_t = \exp\left(\int_0^t b'(X_s)dW_s + \int_0^t \left(a' - \frac{1}{2}(b')^2\right)(X_s)ds\right). \tag{3.3}$$

Since the process Σ is $\mathcal{F}$-measurable, we see that V is an $\mathcal{F}$-conditional Gaussian martingale with $\mathcal{F}$-conditional mean zero.

In the first step we will obtain an explicit representation of the leading term of the normalised error process V^n defined at (1.4). This stochastic expansion can be also found in the proof of [9, Theorem 1.2]. Nevertheless, we will prove this result for the sake of completeness.

Theorem 3.2. *Let us consider the process*

$$\overline{V}^n_t = -\sqrt{n}\Sigma_t \int_0^t \Sigma^{-1}_{\varphi_n(s)} bb'(X^n_{\varphi_n(s)})(W_s - W_{\varphi_n(s)})dW_s, \tag{3.4}$$

where Σ is defined in (3.3)*. Then it holds that*

$$\sup_{t\in[0,1]} |V^n_t - \overline{V}^n_t| \xrightarrow{\mathbb{P}} 0.$$

We remark at this stage that the process $(\Sigma_t^{-1}\overline{V}^n_t)_{t\in[0,1]}$ is a continuous martingale of quadratic form with random weights. Thus, second order Edgeworth expansion for the functional $\overline{V}^n_t$ can be deduced from the corresponding expansion for the pair $(\Sigma_t, \Sigma_t^{-1}\overline{V}^n_t)$.

In the next step we need to determine the second order stochastic expansion for the standardised error process $(V^n_t)_{t\in[0,1]}$. Apart from rather complex approximation techniques, the result of Lemma 3.1 is crucial for the next theorem. We remark that this statement has an interest in its own right.

Theorem 3.3. *Assume that the functions a, b are globally Lipschitz and $a, b \in C^2(\mathbb{R})$. Define the process $(R^n_t)_{t\geq 0}$ via*

$$\begin{aligned} dR^n_t = &\bigg(\frac{1}{2\sqrt{n}} a''(X_t)(V^n_t)^2 + \sqrt{n} b\left((b')^2 - a'\right)(X^n_{\varphi_n(t)})(W_t - W_{\varphi_n(t)}) \\ &- \sqrt{n} aa'(X^n_{\varphi_n(t)})(t - \varphi_n(t)) - \frac{\sqrt{n}}{2} b^2 a''(X^n_{\varphi_n(t)})(W_t - W_{\varphi_n(t)})^2\bigg) dt \\ &+ \bigg(\frac{1}{2\sqrt{n}} b''(X_t)(V^n_t)^2 + \sqrt{n}\Big(b(b')^2 - \frac{b^2 b''}{2}\Big)(X^n_{\varphi_n(t)})(W_t - W_{\varphi_n(t)})^2 \\ &- \sqrt{n} ab'(X^n_{\varphi_n(t)})(t - \varphi_n(t))\bigg) dW_t = R^n_t(1)dt + R^n_t(2)dW_t \end{aligned} \tag{3.5}$$

Then the process $\sqrt{n}R^n$ is tight and we have that

$$\sqrt{n}\sup_{t\in[0,1]}\left|V_t^n-\left(\overline{V}_t^n+\Sigma_t\int_0^t\Sigma_s^{-1}\left(dR_s^n-b'(X_s)R_s^n(2)ds\right)\right)\right|\xrightarrow{\mathbb{P}}0.$$

Theorem 3.3 implies that, for any fixed $t\in[0,1]$, we have the stochastic expansion $V_t^n=\Sigma_t(M_t^n+n^{-1/2}N_t^n)$ with

$$M_t^n=\Sigma_t^{-1}\overline{V}_t^n,\qquad N_t^n=\sqrt{n}\int_0^t\Sigma_s^{-1}\left(dR_s^n-b'(X_s)R_s^n(2)ds\right)+o_{\mathbb{P}}(1).\tag{3.6}$$

In the next section we will determine the stable central limit theorem for the triplet $(M_t^n,\sqrt{n}(C_t^n-C_t),N_t^n)_{t\in[0,1]}$.

4 Stable central limit theorems and Edgeworth expansion

4.1 Central limit theorems

Having derived the stochastic expansion for the standardised error process $(V_t^n)_{t\in[0,1]}$ in the previous section, we now need to prove the stable central limit theorem required in assumpion (B2)(iii). For this purpose we introduce the following auxiliary processes:

$$\begin{aligned}A_t^n(1)&=n\int_0^t\Sigma_s^{-1}b\left((b')^2-a'\right)(X_{\varphi_n(s)}^n)\left(\varphi_n(s+n^{-1})-s\right)dW_s\\&\quad-n\int_0^t\Sigma_s^{-1}ab'(X_{\varphi_n(s)}^n)(s-\varphi_n(s))dW_s\\&\quad+n\int_0^t\Sigma_s^{-1}\left(b(b')^2-\frac{b^2b''}{2}\right)(X_{\varphi_n(s)}^n)(W_s-W_{\varphi_n(s)})^2dW_s,\end{aligned}\tag{4.1}$$

$$A_t^n(2)=2n^{3/2}\int_0^t\left(\Sigma_s^{-1}bb'(X_{\varphi_n(s)}^n)\right)^2\left(\varphi_n(s+n^{-1})-s\right)\left(W_s-W_{\varphi_n(s)}\right)dW_s,\tag{4.2}$$

$$\begin{aligned}A_t^n(3)&=n\int_0^t\Sigma_s^{-1}a\left((b')^2-a'\right)(X_{\varphi_n(s)}^n)\left(s-\varphi_n(s)\right)ds\\&\quad-\frac{n}{2}\int_0^t\Sigma_s^{-1}\left(b^2a''+b^2b'b''-2b(b')^3\right)(X_{\varphi_n(s)}^n)\left(W_s-W_{\varphi_n(s)}\right)^2ds.\end{aligned}\tag{4.3}$$

Our first asymptotic result is the following stable central limit theorem.

Proposition 4.1. *Assume that conditions of Theorem 3.3 are satisfied. Then it holds that*

$$L^n:=(M^n,A^n(1),A^n(2))\xrightarrow{d_{st}}L=\int_0^\cdot v_sdW_s+\int_0^\cdot(u_s-v_s^\star v_s)^{1/2}dB_s\quad on\ C([0,1])^3,\tag{4.4}$$

where $(B_t)_{t\in[0,1]}$ *is a* 3*-dimensional Brownian motion defined on an extension* $(\overline{\Omega},\overline{\mathcal{F}},\overline{\mathbb{P}})$ *of the original probability space and independent of* $\mathcal{F}$*, and the processes* $v_s=(v_s^1,v_s^2,v_s^3)$*,* $u_s=(u_s^{ij})_{1\le i,j\le 3}$ *are defined by*

$$v_s^1=v_s^3=0,\qquad v_s^2=\Sigma_s^{-1}\left(b(b')^2-\frac{ab'+a'b}{2}-\frac{b^2b''}{4}\right)(X_s),$$

$$u_s^{12}=u_s^{21}=u_s^{23}=u_s^{32}=0,$$

$$u_s^{11}=\frac{1}{2}\left(\Sigma_s^{-1}bb'(X_s)\right)^2,\qquad u_s^{33}=\frac{1}{3}\left(\Sigma_s^{-1}bb'(X_s)\right)^4,\qquad u_s^{13}=u_s^{31}=-\frac{1}{3}\left(\Sigma_s^{-1}bb'(X_s)\right)^3,$$

$$u_s^{22}=\frac{1}{3}\Sigma_s^{-2}\Big(b^2\left[\left((b')^2-a'\right)^2+\left((b')^2-\frac{bb''}{2}\right)\left(4(b')^2-\frac{3bb''}{2}-a'\right)\right]$$

$$+(ab')^2-abb'\left(3(b')^2-a'-bb''\right)\Big)(X_s).$$

Proof. Note that L^n is a continuous martingale with mean zero. According to [10, Theorem IX.7.3], it is sufficient to prove that

$$\langle L^n\rangle_t\xrightarrow{\mathbb{P}}\int_0^t u_s ds,\qquad \langle L^n,W\rangle_t\xrightarrow{\mathbb{P}}\int_0^t v_s ds,\qquad \langle L^n,Q\rangle_t\xrightarrow{\mathbb{P}}0,\qquad \forall t\in[0,1],$$

where the last statement should hold for any bounded continuous martingale Q with $\langle W,Q\rangle=0$. The first two statements follow by a straightforward but tedious computation taking into account that $X_s^n\xrightarrow{\mathbb{P}}X_s$ for any $s\in[0,1]$, $\sup_{s\in[0,1]}|\varphi_n(s)-s|\to 0$ and the continuity of involved processes/functions. The third condition is a consequence of the formula $\langle\int_0^\cdot w_s dW_s,Q\rangle_t=\int_0^t w_s d\langle W,Q\rangle_s=0$ for any predictable process $(w_s)_{s\in[0,1]}$. □

As a consequence of the previous result we deduce the joint stable central limit theorem for the triplet $(M_t^n,N_t^n,\sqrt{n}(C_t^n-C_t))_{t\in[0,1]}$.

Proposition 4.2. *Assume that conditions of Theorem 3.3 are satisfied. Then it holds that*

$$(M^n,N^n,\sqrt{n}(C^n-C))\xrightarrow{d_{st}}\left(L^1,\frac{1}{2}\int_0^\cdot\Sigma_s\left(a''+b''-b'b''\right)(X_s)(L_s^1)^2ds+L^2+A(3),L^3\right)$$

$$:=(M,N,\widehat{C})\qquad on\ C([0,1])^3, \tag{4.5}$$

where the process $L=(L^1,L^2,L^3)$ *has been introduced in Proposition 4.1 and the process* $(A_t(3))_{t\in[0,1]}$ *is defined as*

$$A_t(3)=\int_0^t\Sigma_s^{-1}\left(\frac{1}{2}a(b')^2+\frac{1}{2}b(b')^3-\frac{1}{2}aa'-\frac{1}{4}a''b^2-\frac{1}{4}b^2b'b''\right)(X_s)ds.$$

Proof. First of all, it holds that $\sup_{t\in[0,1]}|A_t^n(3)-A_t(3)|\xrightarrow{\mathbb{P}}0$, which is due to [9, Theorem 7.2.2]. Secondly, using the identities $(W_b-W_a)^2-(b-a)=2\int_a^b(W_s-W_a)dW_s$ and

$\int_a^b (Y_s - Y_a)ds = \int_a^b (b-s)dY_s$, which hold for any $b > a$ and any continuous semimartingale Y, we obtain that

$$\sqrt{n}(C_t^n - C_t) = A_t^n(2).$$

Furthermore, observing the definition (3.5) of the process R^n, we deduce the identity

$$\begin{aligned} N_t^n &= \sqrt{n} \int_0^t \Sigma_s^{-1} \left(dR_s^n - b'(X_s) R_s^n(2) ds \right) + o_{\mathbb{P}}(1) \\ &= A_t^n(1) + A_t^n(3) + \frac{1}{2} \int_0^t \Sigma_s^{-1} (V_s^n)^2 \left(a'' + b'' - b'b'' \right)(X_s) ds + o_{\mathbb{P}}(1). \end{aligned}$$

Now, due to convergence (4.4) in Proposition 4.1 and the properties of stable convergence we deduce that $(M^n, A^n(1), A^n(2), A^n(3), \Sigma, X) \xrightarrow{d_{st}} (L^1, L^2, L^3, A(3), \Sigma, X)$ on $C([0,1])^6$. Hence, by [10, Theorem VI.6.22] and continuous mapping theorem for stable convergence applied to the function $H : C([0,1])^6 \to C([0,1])^3$

$$H(y) := \left(y_1, y_2 + y_4 + \frac{1}{2} \int_0^{\cdot} y_5(s)^{-1} (y_1(s) y_5(s))^2 \left(a'' + b'' - b'b'' \right)(y_6(s)) ds, y_3 \right)$$

we obtain that

$$(M^n, N^n, \sqrt{n}(C^n - C)) \xrightarrow{d_{st}} (M, N, \widehat{C}) \qquad \text{on } C([0,1])^3.$$

This completes the proof of Proposition 4.2. □

We remark that the 3-dimensional limiting process $(M, N, \widehat{C})$ is an $\mathcal{F}$-conditional Gaussian martingale. This property will help us to compute the classical random symbol $\underline{\sigma}(z, \mathtt{i}u, \mathtt{i}v)$ in the next section.

4.2 Multivariate Edgeworth expansion associated with the Euler scheme

Let us now consider fixed time points $0 = T_0 < T_1 < \ldots < T_k \le 1$. In this section we will investigate the multivariate Edgeworth expansion for the vector $(V_{T_1}^n, \ldots, V_{T_k}^n)$. We recall the representation introduced at (3.6):

$$V_{T_j}^n = \Sigma_{T_j} (M_{T_j}^n + n^{-1/2} N_{T_j}^n) \qquad \text{with}$$

$$M_{T_j}^n = \Sigma_{T_j}^{-1} \overline{V}_{T_j}^n, \qquad N_{T_j}^n = \sqrt{n} \int_0^{T_j} \Sigma_s^{-1} \left(dR_s^n - b'(X_s) R_s^n(2) ds \right) + o_{\mathbb{P}}(1).$$

According to the Edgeworth expansion theory demonstrated in Section 2, we will first derive the density expansion for the vector $(\Sigma_{T_j}, M_{T_j}^n + n^{-1/2} N_{T_j}^n)_{1 \le j \le k}$.

We define the k-dimensional $(\mathcal{F}_t)$-martingale with components $M^{n,j} := (M_{\min(t,T_j)}^n)_{t \in [0,1]}$, which obviously satisfies the terminal condition $M_1^{j,n} = M_{T_j}^n$ for $j = 1, \ldots, k$. Similarly, we set $N^{n,j} = N_{T_j}^n$. We introduce the set of increasing numbers $(t_i)_{0 \le i \le m_n}$ via

$\{t_i\} = \{j/n:\ j = 0,\ldots,n\} \cup \{T_1,\ldots,T_k\}$. In the notation of Section 2.2 the martingale $M^{n,j}$ satisfies the representation (2.5) with

$$K^{n,j}(s) = -\Sigma^{-1}_{\varphi_n(s)} bb'(X^n_{\varphi_n(s)}) 1_{[0,T_j)}(\varphi_n(s)) \quad \text{and} \quad K(s) = -\Sigma_s^{-1} bb'(X_s). \tag{4.6}$$

The anticipative random symbol $\overline{\sigma}$ is then defined through the identity (2.9). Now, we turn our attention to the adaptive random symbol $\underline{\sigma}$.

We consider a $\overline{q} := (k+q)$-dimensional random variable

$$G = (\Sigma_{T_1},\ldots,\Sigma_{T_k}, F),$$

where F is a q-dimensional random functional. From Proposition 4.2 we readily deduce the stable convergence

$$\left(M^n, N^n, \sqrt{n}(C^n - C)\right) \xrightarrow{d_{st}} \left(M, N, \widehat{C}\right) \tag{4.7}$$

where $M^n = (M^{n,1},\ldots,M^{n,k})$ and $N^n = (N^{n,1},\ldots,N^{n,k})$. Now, we need to determine the mixed normal representation of the vector $(M,N,\widehat{C})$. Note that the $\mathcal{F}$-conditional mean of the first and the third component is zero, which is due to Proposition 4.2. On the other hand, the $\mathcal{F}$-conditional mean of N is not vanishing. Observing the representation of N in Proposition 4.2 and applying Itô's formula we conclude that

$$\mu_j := \overline{\mathbb{E}}[N^j|\mathcal{F}] = \int_0^{T_j} v_s^2 dW_s + A_{T_j}(3) + \frac{1}{2}\int_0^{T_j} \Sigma_s \left(a'' + b'' - b'b''\right)(X_s) \left(\int_0^s u_r^{11} dr\right) ds,$$

where the processes v^2, u^{11} and $A(3)$ have been introduced in Propositions 4.1 and 4.2. Furthermore, the $\mathcal{F}$-conditional covariance structure of the vector $(M,N,\widehat{C})$ is fully determined by Proposition 4.2. Thus, setting $\mu = (\mu_1,\ldots,\mu_k)$, we may write

$$(M_1, N, \widehat{C}) \sim MN\left((0,\mu,0), \begin{pmatrix} \Theta_{11} & \Theta_{12} & \Theta_{13} \\ \Theta_{21} & \Theta_{22} & \Theta_{13} \\ \Theta_{31} & \Theta_{32} & \Theta_{33} \end{pmatrix}\right).$$

Due to $\mathcal{F}$-conditional Gaussianity of the limit $(M_1, N, \widehat{C})$, the adaptive symbol $\underline{\sigma}$ has the following form:

$$\underline{\sigma}(z, \mathtt{i}u, \mathtt{i}v) = \frac{(\Theta_{31}\Theta_{11}^{-1}z)[(\mathtt{i}u)^{\otimes 2}]}{2} + (\mu + \Theta_{21}\Theta_{11}^{-1}z)[\mathtt{i}u] \qquad z, u \in \mathbb{R}^k. \tag{4.8}$$

Combining two random symbols, we end up with the approximative density

$$\begin{aligned} &p_n^{(Z_n,G)}(z,y,x) = \mathbb{E}[\phi(z;0,C)|G=(y,x)]p^G(y,x) \\ &+n^{-1/2}\sum_j (-d_z)^{m_j}(-d_x)^{n_j}\Big(\mathbb{E}\left[c_j(z)\phi(z;0,C)|G=(y,x)\right]p^G(y,x)\Big), \end{aligned} \tag{4.9}$$

with $(z,x,y) \in \mathbb{R}^k \times \mathbb{R}^k \times \mathbb{R}^q$, as in (2.11).

In this setting, however, the kernels $K^{n,j}$ and K are defined by (4.6), and the functionals $c_\alpha(z)$ in the representation (2.10) of the full random symbol σ and also in (4.9) are associated with $\underline{\sigma}$ of (4.8) and $\overline{\sigma}$ of (2.9).

In the following we will assume the following condition:

(A) The functions a and b are in $C^\infty(\mathbb{R})$ and all their derivatives of positive order are bounded.

Under (A) conditions of Theorem 2.1 can be slimed down. Recall that the variables $\mathbb{I}_s^j$ are defined by (2.6).

(C1) For every $p > 1$ and $j = 1, ..., k$,

$$\sup_{s\in(T_{j-1},T_j)} \left\| \left[\frac{\mathbb{I}_s^j}{T_j - s} \right]^{-1} \right\|_{\mathbb{L}^p} \quad < \quad \infty.$$

Recall that $\ell = k + \overline{q} + 8$.

(C2) $F \in \mathbb{D}_{\ell+1,\infty}(\mathbb{R}^q)$, $\sup_{r_1,...,r_m\in(0,1)} \|D_{r_1,...,r_m}F\|_{\mathbb{L}^p} < \infty$ for every $p > 1$ and $m = 1, ..., \ell + 1$. Moreover, $r \mapsto D_rF$ and $(r, s) \mapsto D_{r,s}F$ $(r \le s)$ are continuous a.s.

(C3) $\det \sigma_G \in \mathbb{L}_{\infty-}$.

Under the aforementioned conditions we obtain the following theorem, which is proved in Section 6.4.

Theorem 4.3. *Suppose that conditions (A), $(C1)$, $(C2)$ and $(C3)$ are fulfilled. Then, for every pair of positive numbers (K, γ),*

$$\sup_{h\in\mathcal{E}(K,\gamma)} \left| \mathbb{E}[h(Z_n, G)] - \int_{\mathbb{R}^k\times\mathbb{R}^k\times\mathbb{R}^q} h(z, y, x) p_n^{(Z_n,G)}(z, y, x) dz dy dx \right| \quad = \quad o(n^{-1/2})$$

as $n \to \infty$.

As a consequence of Theorem 4.3 we finally obtain the approximative density of the pair (V_n, F) for $V_n = (V_{T_1}^n, \ldots, V_{T_k}^n)$ and an external q-dimensional random variable F.

Corollary 4.4. *We set*

$$p_n^{(V_n,F)}(z, x) = \int_{\mathbb{R}_+^k} \frac{1}{y_1 \cdots y_k} p_n^{(Z_n,G)} \left(z_1/y_1, \ldots, z_k/y_k, y_1, \ldots, y_k, x\right) dy. \tag{4.10}$$

Under the conditions of Theorem 4.3 we obtain that

$$\sup_{h\in\mathcal{E}(K,\gamma)} \left| \mathbb{E}[h(V_n, F)] - \int_{\mathbb{R}^k\times\mathbb{R}^q} h(z, x) p_n^{(V_n,F)}(z, x) dz dx \right| = o(n^{-1/2}).$$

Theorem 4.3 relies on the non-degeneracy of G. We will discuss some sufficient conditions in the following subsections. When $k \ge 2$, the non-degeneracy becomes a global problem and it is not so straightforward to consider the question in full generality. However, a localization method provides a practical solution.

4.3 On condition $(C1)$

In this section we will give a sufficient condition for $(C1)$. We are working in the setting of Section 4.2 imposing assumption (A). We consider the following condition:

(C1$^\sharp$) **(i)** $\inf_{x\in\mathbb{R}} |b(x)| > 0$.

(ii) There exists a compact set $B \subseteq \mathbb{R}$ such that

(a) $\inf_{x\in B^c} |b'(x)| > 0$,

(b) $\sum_{j=1}^{\infty} |b^{(j)}(x)| \neq 0$ for each $x \in B$.

For example, in the setting of null drift, if X_t visits the set $\{x : b'(x) = 0\}$ after some time, then Σ_t does not diffuse there and we never get non-degeneracy of Σ_t thereafter. This explains the necessity of a global condition like $(C1^\sharp)$(ii)(a). As a matter of fact, such a degenerate case is essentially in the scope of the classical expansion for a martingale with an exactly normal limit (cf. [20]). Now we have the following result.

Proposition 4.5. *Condition $(C1)$ holds under (A) and $(C1^\sharp)$.*

Proof. We need to show that

$$\sup_{s\in(T_{j-1},T_j)} \left\| \left[\frac{\mathbb{I}_s^j}{T_j - s}\right]^{-1} \right\|_{\mathbb{L}^p} \quad < \quad \infty \tag{4.11}$$

for every $p > 1$ and $j = 1, \dots, k$. Let $s \in (T_{j-1}, T_j)$. Recalling (2.6), we have

$$\begin{aligned} \frac{\mathbb{I}_s^j}{T_j - s} &= \frac{1}{2}\frac{1}{T_j - s}\int_s^{T_j} \Sigma_r^{-2}\{bb'(X_r)\}^2 dr \\ &\geq \frac{1}{2}\inf_{r\in[s,T_j]} \Sigma_r^{-2} \times \frac{1}{T_j - s}\int_s^{T_j} \{bb'(X_r)\}^2 dr. \end{aligned}$$

By $(C1^\sharp)$ and the compactness of B, there exist a finite set $\mathcal{N} \subset B$, a positive constant c and an integer $m \geq 2$ such that

$$\{bb'(x)\}^2 \quad \geq \quad \min_{z\in\mathcal{N}} c^{m/2}(1 \wedge |x - z|^m) \tag{4.12}$$

for all $x \in \mathbb{R}$. Indeed, by $(C1^\sharp)$(i) and (ii)(a), there exists a positive constant c' such that $\inf_{x\in B^c}\{bb'(x)\}^2 \geq c'$. For each $z \in B$, by $(C1^\sharp)$(ii)(b), there exists an integer $j_z \geq 1$ such that $b^{(j_z)}(z) \neq 0$ and $b'(x) = ((j_z-1)!)^{-1}b^{(j_z)}(z)(x-z)^{j_z-1}+\cdots$ for all x near z. Therefore, from $(C1^\sharp)$(i), for each $z \in B$, there exists a positive constant c_z and a neighborhood B_z such that $\{bb'(x)\}^2 \geq c_z(1 \wedge |x - z|^{m_z})$ for all $x \in B_z$, with $m_z = (j_z - 1)^2 \geq 0$. Since B is compact, one can find a finite set $\mathcal{N} \subset B$ such that $B \subset \cup_{z\in\mathcal{N}} B_z$, and hence

$$\{bb'(x)\}^2 \geq \min_{z\in\mathcal{N}} \big(\min_{z'\in\mathcal{N}} c_{z'}\big)\big(1 \wedge |x - z|^{\max_{z'\in\mathcal{N}} m_{z'}}\big)$$

for all $x \in B$ since there exists z for each $x \in B$ such that $x \in B_z$. If we set $c = \big(\min\{c', \min_{z\in\mathcal{N}} c_z\}\big)^{2/m}$ for $m = \max\{2, \max_{z\in\mathcal{N}} m_z\}$ we obtain (4.12).

Let $\delta > 0$ and $B_0 := \{x : \text{ dist } (x, \mathcal{N}) < 2\delta\}$. Let $s_i = s + i(T_j - s)/n$. Then, there exists $n_0 \in \mathbb{N}$ independent of s such that for $n \geq n_0$,

$$
\begin{aligned}
& \mathbb{P}\bigg[\frac{1}{T_j - s}\int_s^{T_j} \{bb'(X_r)\}^2 dr \leq \frac{1}{n^{3m/2}}\bigg] \\
\leq\ & \mathbb{P}\bigg[\frac{c^{m/2}}{T_j - s}\int_s^{T_j} \min_{z\in\mathcal{N}}(1 \wedge |X_r - z|^m) dr \leq \frac{1}{n^{3m/2}}\bigg] \\
\leq\ & \mathbb{P}\bigg[\frac{c}{T_j - s}\int_s^{T_j} \min_{z\in\mathcal{N}}(1 \wedge |X_r - z|^2) dr \leq \frac{1}{n^{3}}\bigg] \\
\leq\ & \sum_{i=1}^n \mathbb{P}\bigg[\frac{c}{T_j - s}\int_{s_{i-1}}^{s_i} \min_{z\in\mathcal{N}}(1 \wedge |X_r - z|^2) dr \leq \frac{1}{n^{4}}\bigg] \\
=\ & \sum_{i=1}^n \mathbb{P}\bigg[\frac{c}{T_j - s}\int_{s_{i-1}}^{s_i} \min_{z\in\mathcal{N}}(1 \wedge |X_r - z|^2) dr \leq \frac{1}{n^{4}},\ \inf_{r\in[s_{i-1},s_i]} \min_{z\in\mathcal{N}} |X_r - z| < n^{-1/2}\bigg] \\
\leq\ & \sum_{z\in\mathcal{N}}\sum_{i=1}^n \mathbb{P}\bigg[\frac{c}{T_j - s}\int_{s_{i-1}}^{s_i} (1 \wedge |X_r - z|^2) dr \leq \frac{1}{n^{4}},\ \sup_{u\in[s_{i-1},s_i]} |X_r - z| < n^{-1/3}\bigg] + O(n^{-L})
\end{aligned}
$$

where L is any positive number independent of s; in fact, on the event $\big\{\inf_{r\in[s_{i-1},s_i]} |X_r - z| < n^{-1/2}\big\}$ for $z \in \mathcal{N}$, the process X keeps $\sup_{r'\in[s_{i-1},s_i]} |X_{r'} - z| < n^{-1/3}$ with probability $1 - O(n^{-L-1})$, and $\min_{z'\in\mathcal{N}} |X_{r'} - z'| = |X_{r'} - z|$ for $n \geq n_0$ since the points in $\mathcal{N}$ are isolated. The first term of the right-hand side of the above inequality is bounded by

$$
\begin{aligned}
& \sum_{z\in\mathcal{N}}\sum_{i=1}^n \mathbb{P}\bigg[\frac{c}{T_j - s}\int_{s_{i-1}}^{s_i} |X_r - z|^2 dr \leq \frac{1}{n^4},\ X_r \in B_0 \text{ for all } r \in [s_{i-1}, s_i]\bigg] \\
=\ & \sum_{z\in\mathcal{N}}\sum_{i=1}^n \mathbb{P}\bigg[\frac{1}{s_i - s_{i-1}}\int_{s_{i-1}}^{s_i} |X_r - z|^2 dr \leq \frac{1}{cn^3},\ X_r \in B_0 \text{ for all } r \in [s_{i-1}, s_i]\bigg]
\end{aligned}
$$

for large n. Since on the bounded set B_0, the process X_r behaves like a Brownian motion, the last probability is bounded by $c_1^{-1} n \exp(-c_1 n)$ for some positive constant c_1 independent of $s \in (T_{j-1}, T_j)$, which follows from a similar inequality to [18, Lemma 10.6]. Consequently, we obtain (4.11) by using the estimate

$$
\begin{aligned}
\sup_{s\in(T_{j-1},T_j)} \mathbb{E}[\Gamma_s^{-p}] &= \sup_{s\in(T_{j-1},T_j)} \int_0^\infty p t^{p-1} \mathbb{P}[\Gamma_s < t^{-1}] dt \\
&\leq \sum_{n=0}^\infty p(n+1)^{3mp/2} \sup_{s\in(T_{j-1},T_j)} \mathbb{P}[\Gamma_s < n^{-3m/2}] < \infty
\end{aligned}
$$

for $\Gamma_s = (T_j - s)^{-1} \int_s^{T_j} \{bb'(X_r)\}^2 dr$ and $p > 1$. $\square$

4.4 On condition $(C3)$ for non-degeneracy of G in the case $k = 1$

The problem of non-degeneracy of σ_G can be reduced to local properties of the stochastic differential equations in the case $k = 1$. Consider a system of stochastic differential equations in Stratonovich form

$$d\overline{X}_t \;=\; \overline{V}_0(\overline{X}_t)dt + \overline{V}_1(\overline{X}_t) \circ dW_t, \qquad \overline{X}_0 \;=\; (x_0, 1, f) \tag{4.13}$$

for a $(2+q)$-dimensional process $\overline{X}_t = (\overline{X}_t^{(j)})_{j=1,2,3}$, where $\overline{V}_i = (\overline{V}_i^{(j)})_{j=1,2,3}$ $(i = 0, 1)$ are vector fields. The elements of $\overline{V}_i$'s are specified as follows:

$$\begin{aligned}
\overline{V}_0^{(1)}(\overline{x}) &= \tilde{a}(x_1) := a(x_1) - \frac{1}{2}b(x_1)b'(x_1), \qquad \overline{V}_1^{(1)}(\overline{x}) \;=\; b(x_1), \\
\overline{V}_0^{(2)}(\overline{x}) &= \tilde{a}'(x_1)x_2 = \left\{a'(x_1) - \frac{1}{2}\big(b''(x_1)b(x_1) - (b'(x_1))^2\big)\right\}x_2, \\
\overline{V}_1^{(2)}(\overline{x}) &= b'(x_1)x_2.
\end{aligned}$$

Suppose that the vector fields $\overline{V}_i^{(3)}$ $(i = 0, 1)$ are smooth and their derivatives of positive order are bounded, and that the q-dimensional random variable F is represented by the third element of $\overline{X}_T$ as $F = \overline{X}_T^{(3)}$, $T \in (0, 1]$. In the case $F = 0$, $\overline{X}_t$ is $(\overline{X}_t^{(1)}, \overline{X}_t^{(2)})$ and $\overline{V}_i$ are $(\overline{V}_i^{(1)}, \overline{V}_i^{(2)})$ $(i = 0, 1)$ respectively. By definition, $\overline{X}_T^{(1)} = X_T$ and $\overline{X}_T^{(2)} = \Sigma_T$.

The Lie algebra generated by $\overline{V}_i$ $(i = 0, 1)$ at $\overline{x} \in \mathbb{R}^{2+q}$ is denoted by $\text{Lie}[\overline{V}_1, \overline{V}_0](\overline{x})$, namely, it is the linear span of the vectors in $\cup_{i=0}^{\infty}\mathcal{V}_i$ with $\mathcal{V}_0 = \{\overline{V}_1(\overline{x})\}$, $\mathcal{V}_i = \{[\overline{V}_j(\overline{x}), \mathsf{V}]; \mathsf{V} \in \mathcal{V}_{i-1}\}$ $(i \in \mathbb{N})$, where $[V, W](x) = \mathcal{D}V(x)W(x) - \mathcal{D}W(x)V(x)$ with $\mathcal{D}V(x)$ being the derivative of V at x. A simple criterion for non-degeneracy of σ_G is provided by the Hörmander condition (see Section 2.3.2 in [14] for details).

Proposition 4.6. *Let $k = 1$. For a constant $\overline{X}_0$, if* $\text{span}\,\text{Lie}[\overline{V}_1, \overline{V}_0](\overline{X}_0) = \mathbb{R}^{2+q}$, *then $(C3)$ holds.*

A variation is the case where F has a component X_T, that is, $F = (X_T, F_1)$; F_1 may be empty. If we have a representation $F_1 = \overline{X}_T^{(3)}$, then Proposition 4.6 remains valid.

The non-degeneracy problem for σ_G becomes a global one when $k > 1$ since we need non-degeneracy of $\Sigma_{T_2} - \Sigma_{T_1}$, but the support of Σ_{T_1} is no longer compact. Though we could assume some strong condition that gives uniform non-degeneracy over the whole space, it would be a quite restrictive solution. Instead, in Section 4.5, we will consider a different way by slightly modifying Theorem 4.3, but such modification keeps the error bound of the approximation meaningful in practice.

4.5 Localization

To convey the idea simply, we shall only treat the case $F = (X_{T_j})_{j=1,\dots,k}$, while more general cases can be formulated in a similar manner.

Let us consider the situation of Section 4.4 with the system (4.13) of stochastic differential equations for $\overline{X}_t = (\overline{X}_t^{(1)}, \overline{X}_t^{(2)}) = (X_t, \Sigma_t)$.

(D) $\mathrm{Lie}[\overline{V}_1, \overline{V}_0](x,1) = \mathbb{R}^2$ for $x \in I$.

For positive numbers K and γ, let $\mathcal{E}(K,\gamma,I)$ be the set of measurable functions $h : \mathbb{R}^{3k} \to \mathbb{R}$ such that $h(z,y,x) = 0$ when $x_j \in I^c$ for some $j \in \{1,...,k-1\}$, $x = (x_j)_{j=1,\dots,k}$, and that $|h(z,y,x)| \le M(1+|z|+|y|+|x|)^\gamma$ for all $(z,y,x) \in \mathbb{R}^{3k}$.

Denote by $(X_t(s,x), \Sigma_t(s,(x,y)))$ the stochastic flow defined by

$$\begin{cases} dX_t(s,x) &= \tilde{a}(X_t(s,x))dt + b(X_t(s,x)) \circ dW_t, \\ d\Sigma_t(s,(x,y)) &= \tilde{a}'(X_t(s,x))\Sigma_t(s,(x,y))dt + b'(X_t(s,x))\Sigma_t(s,(x,y)) \circ dW_t \end{cases}$$

with $(X_s(s,x), \Sigma_s(s,(x,y))) = (x,y)$, $0 \le s \le t \le 1$. Assume conditions (A), $(C1)$ and (D). Then by Proposition 4.6 and Theorem 4.3, for each $x_{j-1} \in I$ and $y_{j-1} > 0$, there exists a density

$$q_n^{(j)}\big(\zeta_j, \eta_j, x_j | y_{j-1}, x_{j-1}\big) \;=\; p_n^{\big(V^n_{T_j} - V^n_{T_{j-1}},\, y_{j-1}^{-1}\Sigma_{T_j}(T_{j-1},(x_{j-1},y_{j-1})),\, X_{T_j}(T_{j-1},x_{j-1})\big)}\big(\zeta_j, \eta_j, x_j\big)$$

with initial value $(\Sigma_{T_{j-1}}, X_{T_{j-1}}) = (y_{j-1}, x_{j-1})$ of the system starting at time T_{j-1} that gives the asymptotic expansion

$$\begin{aligned} & \mathbb{E}\big[h_j\big(V^n_{T_j} - V^n_{T_{j-1}}, y_{j-1}^{-1}\Sigma_{T_j}, X_{T_j}\big)\big|\Sigma_{T_{j-1}} = y_{j-1},\ X_{T_{j-1}} = x_{j-1}\big] \\ & - \int_{\mathbb{R}^3} h_j(\zeta_j,\eta_j,x_j) q_n^{(j)}\big(\zeta_j,\eta_j,x_j|y_{j-1},x_{j-1}\big) d\zeta_j d\eta_j dx_j \\ = \;& o(n^{-1/2}) \end{aligned}$$

uniformly in $h_j \in \mathcal{E}(K,\gamma)$ for every $(K,\gamma) \in (0,\infty)^2$. Indeed, $q_n^{(j)}\big(\zeta_j,\eta_j,x_j|y_{j-1},x_{j-1}\big)$ is the density $p_n(\zeta_j,\eta_j,x_j)$ in the one-step case starting from time T_{j-1} and the initial values $X_0 = x_{j-1} \in I$ and $\Sigma_0 = 1$. Then we obtain a function $q_n^{(Z_n,G)}(z,y,x)$ that approximates the distribution of (Z_n, G) with $G = ((\Sigma_{T_j})_{j=1,\dots,k}, (X_{T_j})_{j=1,\dots,k})$:

$$q_n^{(Z_n,G)}(z,y,x) \;=\; \prod_{j=1}^k q_n^{(j)}\big(z_j - z_{j-1}, y_{j-1}^{-1}y_j, x_j | y_{j-1}, x_{j-1}\big) y_{j-1}^{-1}$$

for $(z,y,x) = \big((z_j)_{j=1,\dots,k}, (y_j)_{j=1,\dots,k}, (x_j)_{j=1,\dots,k}\big)$, $(z_0,y_0) = (0,1)$. We should remark that this function is defined only when $x_{j-1} \in I$ for $j = 1,...,k$. Now we give a localized version of Theorem 4.3.

Theorem 4.7. *Suppose that Conditions (A), $(C1)$ and (D) are fulfilled for some finite closed interval I. Let $G = ((\Sigma_{T_j})_{j=1,\dots,k}, (X_{T_j})_{j=1,\dots,k})$. Then, for every pair of positive numbers (K,γ),*

$$\sup_{h \in \mathcal{E}(K,\gamma,I)} \left| \mathbb{E}[h(Z_n,G)] - \int_{\mathbb{R}^k \times \mathbb{R}^k \times \mathbb{R}^k} h(z,y,x) q_n^{(Z_n,G)}(z,y,x) dz dy dx \right| \;=\; o(n^{-1/2})$$

as $n \to \infty$.

For a sketch of the proof of Theorem 4.7, we notice that the function h admits the estimate

$$|h(z,y,x)| \quad \leq \quad M_1 \prod_{j=1}^{k} \left(1+|z_j|+|y_j|+|x_j|\right)^{\gamma_1}$$

for some $(M_1,\gamma_1) \in (0,\infty)^2$. Then repeated use of the approximation yields the desired error bound.

The asymptotic expansion for $(V_n,(X_{T_j})_{j=1,\dots,k})$ as in Corollary 4.4 also follows under conditions of Theorem 4.7.

5 Applications

5.1 Strong and weak error expansions

As the first application of the density expansion introduced in (4.10) we study the strong and the weak approximation error associated with the Euler approximation scheme.

Proposition 5.1. *(Weak and strong approximation errors) Suppose that conditions of Theorem 4.3 are satisfied.*

(i) *(Strong approximation error) Let $p_n^{V_n}(z)$ be the marginal density obtained from $p_n^{(V_n,F)}(z,x)$, defined at (4.10), by projection onto the first component and let $U_n = (X^n_{T_1},\dots,X^n_{T_k}) - (X_{T_1},\dots,X_{T_k})$. Then we obtain the following expansion for the L^p-norm of the approximation error*

$$\mathbb{E}[\|U_n\|^p]^{1/p} = n^{-1/2}\left(\int_{\mathbb{R}^k} \|z\|^p p_n^{V_n}(z)dz\right)^{1/p} + o(n^{-1/2}).$$

(ii) *(Weak approximation error) Consider a function $f \in C^2(\mathbb{R}^k)$ such that the second derivative of f has polynomial growth. Setting $p_n^{(V_n,F)}(z,x) = p_1(z,x) + n^{-1/2}p_2(z,x)$ we deduce the asymptotic expansion*

$$\begin{aligned}
&\mathbb{E}[f(X^n_{T_1},\dots,X^n_{T_k}) - f(X_{T_1},\dots,X_{T_k})] \\
&= n^{-1}\int_{\mathbb{R}^k\times\mathbb{R}^k}\left(\langle \nabla f(x), z\rangle \cdot p_2(z,x) + \frac{1}{2} z^\star \mathit{Hess} f(x) z \cdot p_1(z,x)\right) dz dx + o(n^{-1}).
\end{aligned}$$

Proof. Part (i) of the statement is a direct consequence of Corollary 4.4 applied to the function $h(z) = \|z\|^p$. Now, we set $\mathbf{X}^n = (X^n_{T_1},\dots,X^n_{T_k})$ and $\mathbf{X} = (X_{T_1},\dots,X_{T_k})$. To obtain part (ii) of Proposition 5.1 we apply Taylor expansion to conclude that

$$\begin{aligned}
f(\mathbf{X}^n) - f(\mathbf{X}) &= \langle \nabla f(\mathbf{X}), \mathbf{X}^n - \mathbf{X}\rangle + \frac{1}{2}(\mathbf{X}^n - \mathbf{X})^\star \mathrm{Hess} f(\mathbf{X})(\mathbf{X}^n - \mathbf{X}) \\
&\quad + \frac{1}{2}(\mathbf{X}^n - \mathbf{X})^\star \left(\mathrm{Hess} f(\mathbf{Y}^n) - \mathrm{Hess} f(\mathbf{X})\right)(\mathbf{X}^n - \mathbf{X}),
\end{aligned}$$

for some random vector $\mathbf{Y}^n \in \mathbb{R}^k$ with $\|\mathbf{Y}^n - \mathbf{X}\| \le \|\mathbf{X}^n - \mathbf{X}\|$. In particular, $\mathbf{Y}^n \xrightarrow{\mathbb{P}} \mathbf{X}$. Observe that

$$\mathbb{E}\left[(\mathbf{X}^n - \mathbf{X})^\star \left(\mathrm{Hess} f(\mathbf{Y}^n) - \mathrm{Hess} f(\mathbf{X})\right)(\mathbf{X}^n - \mathbf{X})\right] \to 0 \qquad \text{as } n \to \infty,$$

which is due to $f \in C^2(\mathbb{R}^k)$. We deduce the expansion

$$\begin{aligned}&\mathbb{E}[f(X^n_{T_1}, \ldots, X^n_{T_k}) - f(X_{T_1}, \ldots, X_{T_k})] \\ &= n^{-1} \int_{\mathbb{R}^k \times \mathbb{R}^k} \left(\langle \nabla f(x), z \rangle \cdot p_2(z,x) + \frac{1}{2} z^\star \mathrm{Hess} f(x) z \cdot p_1(z,x) \right) dz dx + o(n^{-1})\end{aligned}$$

since, according to Theorem 4.3 and Corollary 4.4 applied to $F = (X_{T_1}, \ldots, X_{T_k})$, it holds that

$$\int_{\mathbb{R}^k \times \mathbb{R}^k} \langle \nabla f(x), z \rangle \cdot p_1(z,x) dz dx = 0,$$

because the dz-integral is taking over an odd function in z. This completes the proof of Proposition 5.1. □

We remark that the weak error expansion of Proposition 5.1(ii) has been obtained in [3, 4] for $k = 1$ and the discrete Euler scheme. Furthermore, the authors proved that the error of the expansion in Proposition 5.1(ii) is $O(n^{-2})$, which is more precise than $o(n^{-1})$. We note however that the theory developed in [3, 4] is not sufficient to obtain the density expansion (4.10) of Corollary 4.4.

5.2 Studentized statistics

In this part we will apply results of Section 4.2 to derive the density of the studentized statistic. To avoid complex notations, we restrict our attention to the case $k = 1$.

To this end, let $T \in [0,1]$. We note that $V^n_T = \Sigma_T Z^n_T$ and $V_T \sim MN(0, S_T)$ with $S_T = \Sigma_T^2 C_T$. Then, the studentized statistic is

$$\frac{V^n_T}{\sqrt{S_T}} = \frac{Z^n_T}{\sqrt{C_T}} \tag{5.1}$$

Hence, it suffices to derive the density of the studentized statistic $Z^n_T/\sqrt{C_T}$.

We write (4.8) in the form

$$\underline{\sigma}(z, \mathtt{i}u, \mathtt{i}v) = \mathcal{H}_1 z (\mathtt{i}u)^2 + (\mathcal{H}_2 + \mathcal{H}_3 z)(\mathtt{i}u)$$

Moreover, we restructure (2.9) as

$$\overline{\sigma}(\mathtt{i}u, \mathtt{i}v) = \mathcal{H}_4(\mathtt{i}u)^3 + \mathcal{H}_5(\mathtt{i}u)(\mathtt{i}v) + \mathcal{H}_6(\mathtt{i}u)^5 + \mathcal{H}_7(\mathtt{i}u)(\mathtt{i}v)^2 + \mathcal{H}_8(\mathtt{i}u)^3(\mathtt{i}v). \tag{5.2}$$

Adding these two random symbols, we obtain the full random symbol

$$\sigma(z, \mathtt{i}u, \mathtt{i}v) = \sum_{j=1}^{7} c_j(z)(\mathtt{i}u)^{m_j}(\mathtt{i}v)^{n_j}. \tag{5.3}$$

where the components of $(m,n) = ((m_j, n_j))_{1\le j\le 7}$ and $c(z) = (c_j(z))_{1\le j\le 7}$ are given by

$$(m,n) = ((1,0),(2,0),(1,1),(3,0),(1,2),(3,1),(5,0))$$

and

$$c(z) = \left(\mathcal{H}_2 + \mathcal{H}_3 z, \mathcal{H}_1 z, \mathcal{H}_5, \mathcal{H}_4, \mathcal{H}_7, \mathcal{H}_8, \mathcal{H}_6\right).$$

In view of Theorem 4.3 and denoting $C = C_T, Z_n = Z_T^n$, we obtain that

$$p_n^{(Z_n,C)}(z,x) = \phi(z;0,x)p^C(x) + n^{-1/2}\sum_{j=1}^{8} p_j(z,x)$$

where for each j we have

$$p_j(z,x) = (-d_z)^{m_j}(-d_x)^{n_j}\Big(\phi(z;0,x)p^C(x)\mathbb{E}\left[c_j(z)|C=x\right]\Big).$$

Note that, in this case, most of the terms are the same as in [15, Section 6]. Hence, adopting their derivations, we easily obtain the following identities:

$$\begin{aligned}
\int_{\mathbb{R}^2} g(z/\sqrt{x})p_1(z,x)dzdx &= \mathbb{E}[\mathcal{H}_2 C^{-1/2}]\int_{\mathbb{R}} g(y)y\phi(y;0,1)dy \\
&\quad + \mathbb{E}[\mathcal{H}_3]\int_{\mathbb{R}} g(y)(y^2-1)\phi(y;0,1)dy, \\
\int_{\mathbb{R}^2} g(z/\sqrt{x})p_2(z,x)dzdx &= \mathbb{E}[\mathcal{H}_1 C^{-1/2}]\int_{\mathbb{R}} g(y)H_3(y)\phi(y;0,1)dy, \\
\int_{\mathbb{R}^2} g(z/\sqrt{x})p_3(z,x)dzdx &= \frac{-1}{2}\mathbb{E}[\mathcal{H}_5 C^{-3/2}]\int_{\mathbb{R}} g(y)(y^3-2y)\phi(y;0,1)dy, \\
\int_{\mathbb{R}^2} g(z/\sqrt{x})p_4(z,x)dzdx &= \mathbb{E}[\mathcal{H}_4 C^{-3/2}]\int_{\mathbb{R}} g(y)H_3(y)\phi(y;0,1)dy, \\
\int_{\mathbb{R}^2} g(z/\sqrt{x})p_5(z,x)dzdx &= \frac{1}{4}\mathbb{E}[\mathcal{H}_7 C^{-5/2}]\int_{\mathbb{R}} g(y)(y^5-4y^3)\phi(y;0,1)dy, \\
\int_{\mathbb{R}^2} g(z/\sqrt{x})p_6(z,x)dzdx &= \frac{-1}{2}\mathbb{E}[\mathcal{H}_8 C^{-5/2}]\int_{\mathbb{R}} g(y)(y^5-7y^3+6y)\phi(y;0,1)dy, \\
\int_{\mathbb{R}^2} g(z/\sqrt{x})p_7(z,x)dzdx &= \mathbb{E}[\mathcal{H}_6 C^{-5/2}]\int_{\mathbb{R}} g(y)H_5(y)\phi(y;0,1)dy,
\end{aligned}$$

where

$$H_3(y) = y^3 - 3y, \quad H_5(y) = y^5 - 10y^3 + 15y.$$

Due to $F = C$ in $\sigma_t(\mathtt{i}u, \mathtt{i}v)$ of (2.9), we notice the equalities

$$\mathcal{H}_4 = \frac{\mathcal{H}_5}{2} \text{ and } 4\mathcal{H}_6 = \mathcal{H}_7 = \mathcal{H}_8,$$

which leads to the following result.

Corollary 5.2. *Under conditions of Theorem 4.3, the Edgeworth expansion is*

$$p^{Z_n/\sqrt{C}}(y) = \phi(y;0,1) + n^{-1/2}\phi(y;0,1)(a_1 y + a_2(y^2-1) + a_3 y^3) \tag{5.4}$$

where

$$\begin{aligned}
a_1 &= \mathbb{E}[\mathcal{H}_2 C^{-1/2}] - 3\mathbb{E}[\mathcal{H}_1 C^{-1/2}] + \frac{1}{2}\mathbb{E}[\mathcal{H}_5 C^{-3/2}] + 3\mathbb{E}[\mathcal{H}_6 C^{-5/2}],\\
a_2 &= \mathbb{E}[\mathcal{H}_3],\\
a_3 &= \mathbb{E}[\mathcal{H}_1 C^{-1/2}].
\end{aligned}$$

6 Proofs

Throughout this section all positive constants are denoted by C although they may change from line to line. Furthermore, due to a standard localisation procedure (see e.g. [5]) all continuous stochastic processes $(Y_t)_{t\in[0,1]}$ can be assumed to be uniformly bounded in (ω, t) when proving Theorems 3.2 and 3.3. In particular, it applies to stochastic processes $Y_t = a^{(l)}(X_t)$ and $Y_t = b^{(l)}(X_t)$ for $l = 0,1,2$. For a generic diffusion process $(Y_t)_{t\in[0,1]}$ of the form (1.1) with bounded coefficients we obtain the inequality

$$\mathbb{E}[|Y_t - Y_s|^p] \le C_p |t-s|^{p/2} \qquad \text{for any } p>0 \text{ and } t,s\in[0,1], \tag{6.1}$$

which holds due to Burkholder-Davis-Gundy inequality. We will use the notation $Y^n \stackrel{u.c.p.}{\longrightarrow} Y$ to denote the uniform convergence in probability $\sup_{t\in[0,1]} |Y_t^n - Y_t| \stackrel{\mathbb{P}}{\longrightarrow} 0$. In the proofs we will deal with sequences of stochastic processes of the form

$$Y_t^n = \sum_{i=1}^{[nt]} \xi_i^n,$$

where ξ_i^n, $i=1,\dots,n$, are $\mathcal{F}_{i/n}$-measurable random variables with $\mathbb{E}[|\xi_i^n|^p] < \infty$ for any $p>0$. The following statements trivially hold:

$$\sum_{i=1}^{[nt]} \mathbb{E}[|\xi_i^n|] \to 0 \qquad \Rightarrow \qquad Y^n \stackrel{u.c.p.}{\longrightarrow} 0 \tag{6.2}$$

$$\sum_{i=1}^{[nt]} \mathbb{E}[\xi_i^n |\, \mathcal{F}_{(i-1)/n}] \stackrel{u.c.p.}{\longrightarrow} Y_t \quad \text{and} \quad \sum_{i=1}^{[nt]} \mathbb{E}[(\xi_i^n)^2 |\, \mathcal{F}_{(i-1)/n}] \stackrel{\mathbb{P}}{\longrightarrow} 0$$
$$\Rightarrow Y^n \stackrel{u.c.p.}{\longrightarrow} Y. \tag{6.3}$$

6.1 Proof of Theorem 2.1

We will sketch the proof, basically following the ideas of [22, Theorem 4], but outlining the difference caused by the multiple stopping in the present situation. Note that as in [22, Theorem 4], it suffices to verify assumptions of [22, Theorem 1].

6.1.1 Construction of the truncation functional ψ_n from s_n and other variables

Let $\bar{d}$ satisfy the inequality $1/3 < \bar{d} < d < 1/2$, where the constant d has been introduced before assumption (B1), and define ξ_n by

$$\xi_n = 10^{-1} n^{2\bar{d}} |C_n - C|^2 + 2\big[1 + 4 \det \sigma_{(M_n,F)}(s_n^k)^{-1}\big]^{-1} + \int_{[0,1]^2} \left(\frac{|C_t^n - C_t - C_s^n + C_s| n^{\bar{d}}}{|t-s|^{3/8}} \right)^8 dtds.$$

We define $Q_n = (M_n, F)$, $R_n = (N_n, \widehat{F}_n)$ and set

$$R'_n = \sigma_{Q_n}^{-1} \left(n^{-1/2} \langle DQ_n, DR_n \rangle_{\mathbb{H}} + n^{-1/2} \langle DR_n, DQ_n \rangle_{\mathbb{H}} + n^{-1} \langle DR_n, DR_n \rangle_{\mathbb{H}} \right).$$

Let $\psi \in C^\infty(\mathbb{R}; [0,1])$ be a function such that $\psi(x) = 1$ if $|x| \le 1/2$ and $\psi(x) = 0$ if $|x| \ge 1$. We introduce the random truncation

$$\psi_n = \psi(\xi_n)\psi(n^{1/2}|R'_n|^2).$$

Remark that ψ_n is well defined because so is $\sigma_{Q_n}^{-1}$ under the truncation by ξ_n. In fact, if $\xi_n \le 1$, then $\det \sigma_{Q_n} \ge s_n^k/4$, that is nondegenerate thanks to $(B3)$(ii). Therefore $\sigma_{Q_n}^{-1}$ makes sense on the event $\{\xi_n \le 1\}$. We are defining $\psi_n = 0$ on the event $\{\xi_n > 1\}$ since $\psi(\xi_n) = 0$ there. Thus, ψ_n is well-defined.

6.1.2 Characteristic function and its decomposition

Let $\check{Z}_n = (Z_n, F_n)$ and let $\check{Z}_n^\alpha = Z_n^{\alpha_1} F_n^{\alpha_2}$ for $\alpha = (\alpha_1, \alpha_2) \in \mathbb{Z}_+^k \times \mathbb{Z}_+^q$. Define

$$\hat{g}_n^\alpha(u,v) = \mathbb{E}\big[\psi_n \check{Z}_n^\alpha \exp\big(Z_n[\mathtt{i}u] + F_n[\mathtt{i}v]\big)\big]$$

for $u \in \mathbb{R}^k$ and $v \in \mathbb{R}^q$ and let

$$g_n^\alpha(z,x) = (2\pi)^{-(k+q)} \int_{\mathbb{R}^{k+q}} \exp\big(-z[\mathtt{i}u] - x[\mathtt{i}v]\big) \hat{g}_n^\alpha(u,v) dudv. \tag{6.4}$$

The existence of the integral (6.4) can be verified by the nondegeneracy of the Malliavin covariance matrix of (Z_n, F_n) under the truncation by ψ_n. We define the quantities

$$\Psi(u,v) = \exp\left(-\frac{1}{2} C[u^{\otimes 2}] + \mathtt{i}F[v] \right),$$

$$\varepsilon_n(u,v) = -\frac{1}{2}(C_n - C)[u^{\otimes 2}] + \mathtt{i}(F_n[v] - F[v]) + \mathtt{i}n^{-1/2} N_n[u],$$

$$e_t^n(u) = \exp\left(\mathtt{i}M_t^n[u] + \frac{1}{2} C_t^n[u^{\otimes 2}] \right),$$

$$L_t^n(u) = e_t^n(u) - 1 \qquad \text{and} \qquad \overset{\circ}{e}(x) = \int_0^1 e^{sx} ds.$$

Finally, we introduce the functions

$$\Phi_n^{1,\alpha}(u,v) = \partial^\alpha \mathbb{E}\Big[e_1^n(u)\Psi(u,v)\varepsilon_n(u,v) \overset{\circ}{e} (\varepsilon_n(u,v))\psi_n\Big],$$

$$\Phi_n^{2,\alpha}(u,v) = \partial^\alpha \mathbb{E}\Big[L_1^n(u)\Psi(u,v)\psi_n\Big].$$

The existence of $\Phi_n^{1,\alpha}(u,v)$ and $\Phi_n^{2,\alpha}(u,v)$ involving $e_1^n(u)\Psi(u,v)$ is ensured by the truncation ψ_n. Let us set

$$\Phi_n^{0,\alpha}(u,v) \quad = \quad \partial^\alpha \mathbb{E}\big[\Psi(u,v)\psi_n\big].$$

Then $\hat{g}_n^\alpha(u,v)$ possess the decomposition

$$\hat{g}_n^\alpha(u,v) \quad = \quad \Phi_n^{0,\alpha}(u,v) + \Phi_n^{1,\alpha}(u,v) + \Phi_n^{2,\alpha}(u,v).$$

6.1.3 Error bound

We apply [22, Theorem 1] by verifying conditions $[B1]$, $[B2]_\ell$, $[B3]$ and $[B4]_{\ell,\mathfrak{m},\mathfrak{n}}$ therein under our assumptions (B1), (B2), (B3). Remark that "ℓ" therein corresponds to $\check{\mathsf{d}}+3$, where $\check{\mathsf{d}} = k+q$. Condition $[B1]$ follows from (B2)(iii) and a standard central limit theorem with a mixed normal limit. Condition $[B2]_\ell$ is verified by (B2)(i), (B1)(i)-(ii), (B2)(ii) and the definition of ξ_n.

Condition $[B3]$ is verified as follows. $C^{n,j}$ and C^j are expressed as

$$C_t^{n,j} \quad = \quad n\sum_{i=1}^{m_n} \int_{i=1}^{m_n} K^{n,j}(t_{j-1})^2 \int_{t_{i-1}\wedge t}^{t_i\wedge t} \left(\int_{t_{i-1}}^{s} dW_r\right)^2 ds$$

and

$$C_t^j \quad = \quad \frac{1}{2}\int_0^{t\wedge T_j} K(s)^2 ds.$$

Routinely, we have

$$\sup_{n\in\mathbb{N}} \left\| n^{\bar{d}} \sup_{t\in[0,1]} \big|C_t^{n,j} - C_t^j\big| \right\|_p \quad < \quad \infty$$

for every $p > 1$ from $(B1)$(i) and (ii). Therefore, $[B3]$(i) follows as

$$\begin{aligned}\mathbb{P}[|\xi_n| > 1/2] \quad &\leq \quad \mathbb{P}[n^{\bar{d}}|C_1^n - C_1| \geq 1] + \mathbb{P}[\det \sigma_{\mathbb{X}_1^k} \leq s_n^k] \\ &\quad +\mathbb{P}\left[\int_{[0,1]^2} \left(\frac{|C_t^n - C_t - C_s^n + C_s| n^{\bar{d}}}{|t-s|^{3/8}}\right)^8 dtds \geq \frac{1}{10}\right] \\ &\to \quad 0\end{aligned}$$

as $n \to \infty$ thanks to $(B3)$(i) and $(B1)$(ii). By the definition of ξ_n, on the event $\{|\xi_n| < 1\}$, $n^{(1-\mathsf{a})/2}|C_n - C| \le 1$ for large n, which is $[B3]$(ii). Moreover, $[B3]$(iii) follows from $(B3)$(ii) since $\limsup_{n\to\infty} \mathbb{E}\big[1_{\{|\xi_n|\le 1\}} \det \sigma_{\mathbb{X}_1^k}^{-p}\big] \le \limsup_{n\to\infty} \mathbb{E}\big[4^p (s_n^k)^{-p}\big] < \infty$.

Condition $[B4]_{\ell,\mathfrak{m},\mathfrak{n}}$(i) is rephrased as (B2)(iv). The present $\overline{\sigma}$ is in $\mathcal{S}(\check{\mathsf{d}}+3,5,2)$ in particular; see [22, p. 892] for the relevant definitions. Thus, [22, Theorem 1] gives the error bound

$$\sup_{f\in\mathcal{E}(R,\gamma)} \Delta_n(f) \;=\; o(n^{-1/2})$$

if the following two conditions are fulfilled:

$$\lim_{n\to\infty} n^{1/2}\Phi_n^{2,\alpha}(u,v) \;=\; \partial^\alpha \mathbb{E}\big[\Psi(u,v)\overline{\sigma}(\mathtt{i}u,\mathtt{i}v)\big] \tag{6.5}$$

for $u \in \mathbb{R}^k$, $v \in \mathbb{R}^q$ and $\alpha \in \mathbb{Z}_+^{\check{\mathsf{d}}}$, and

$$\sup_n \sup_{(u,v)\in\Lambda_n^0(\check{\mathsf{d}},d)} n^{1/2}\big|(u,v)\big|^{\check{\mathsf{d}}+1-\varepsilon}\big|\Phi_n^{2,\alpha}(u,v)\big| \;<\; \infty \tag{6.6}$$

for some $\varepsilon = \varepsilon(\alpha) \in (0,1)$ for every $\alpha \in \mathbb{Z}_+^{\check{\mathsf{d}}}$, where $\Lambda_n^0(\check{\mathsf{d}},\bar{d}) = \{(u,v) \in \mathbb{R}^{\check{\mathsf{d}}}; |(u,v)| \le n^{\bar{d}/2}\}$.

We obtain (6.5) as in [22, Eq. (41)], except for the parts concerning the derivation of [22, Eqs. (38) and (43)] by a non-degeneracy argument. We shall show (6.6). By using duality twice for the double stochastic integrals, we have

$$\begin{aligned}
& n^{1/2}\Phi_n^{2,\alpha}(u,v) \\
&= n \sum_{\mathsf{a}_0,\mathsf{a}_1:\mathsf{a}_0+\mathsf{a}_1=\alpha} c_{\mathsf{a}_0,\mathsf{a}_1} \sum_{i=1}^{m_n} \int_{t_{i-1}}^{t_i}\int_r^{t_i} \mathbb{E}\bigg[\partial^{\mathsf{a}_0}K^n(t_{i-1})[\mathtt{i}u]\partial^{\mathsf{a}_1}D_r\Big\{e_s^n(u)D_s\big(\Psi(u,v)\psi_n\big)\Big\}\bigg]dsdr
\end{aligned}$$

for some constants $c_{\mathsf{a}_0,\mathsf{a}_1}$. We have

$$\begin{aligned}
D_r\Big\{e_s^n(u)D_s\big(\Psi(u,v)\psi_n\big)\Big\} &= e_s^n(u)\Psi(u,v)\sigma(n,r,s;\mathtt{i}u,\mathtt{i}v) \\
&= \mathbb{F}_s^n\mathbb{G}_s\mathbb{H}_s^n\sigma(n,r,s;\mathtt{i}u,\mathtt{i}v),
\end{aligned}$$

where

$$\begin{aligned}
\mathbb{F}_s^n &= \exp\Big(M_s^n[\mathtt{i}u] + F[\mathtt{i}v]\Big), \\
\mathbb{G}_s &= \exp\Big(-\frac{1}{2}\big(C_1 - C_s\big)[u^{\otimes 2}]\Big), \\
\mathbb{H}_s^n &= \exp\Big(\frac{1}{2}\big(C_s^n - C_s\big)[u^{\otimes 2}]\Big)
\end{aligned}$$

and $\sigma(n,r,s;\mathtt{i}u,\mathtt{i}v)$ is a polynomial random symbol of fourth order in (u,v) with coefficients in $\mathbb{D}_{\ell-2,\infty}(\mathbb{R})$.

First, we will consider the case $\alpha = 0$, and estimate $n^{1/2}\Phi_n^{2,0}(u,v)$. Let $s \in (T_{j-1}, T_j)$. Then $M_s^n = \big(M_{T_1}^{n,1}, ..., M_{T_{j-1}}^{n,j-1}, M_s^{n,j}, \dots, M_s^{n,k}\big)$. We will estimate the speed of the decay of the expectations of the components of $n^{1/2}\Phi_n^{2,\alpha}(u,v)$ for $(u,v) \in \Lambda_n^0(\check{\mathsf{d}}, \bar{d})$. Our strategy is as follows. For $s \in (\tau_n^j, T_j)$, we apply the integration-by-parts formula for $\big(M_{T_1}^{n,1}, ..., M_{T_{j-1}}^{n,j-1}, M_s^{n,j}, F\big)$ to obtain the decay $|(u_1, ..., u_j, v_1, ..., v_q)|^{-(\check{\mathsf{d}}+1-\varepsilon)}$, where $u = (u_1, ..., u_k)$ and $v = (v_1, ..., v_q)$. For that, we need to show that the D-derivatives of $\mathbb{G}_s$ and $\mathbb{H}_s^n$ up to ℓ-times are $\mathbb{L}_p$-bounded uniformly in $(u,v) \in \Lambda_n^0(\check{\mathsf{d}}, \bar{d})$ and $n \in \mathbb{N}$, under the truncation by ψ_n. We see that this property holds for $\mathbb{H}_s^n$ by (B1)(ii). For $\mathbb{G}_s$, we verify the property as follows. The multiple D-derivative of $\mathbb{G}_s$ is a linear combination of terms of the form

$$\bigg\{\prod_{\alpha=1}^{\bar{\alpha}} D_{\mathsf{A}_\alpha}\bigg(\sum_{i_1,i_2=j}^{k} \mathbb{I}_s^{i_1\wedge i_2} u_{i_1} u_{i_2}\bigg)\bigg\}\mathbb{G}_s \quad \Big(\mathsf{A}_\alpha = r_{a(\alpha-1)+1}, ..., r_{a(\alpha)}, \quad 1 \le a(1) \le a(2) \le \cdots\Big)$$

that is bounded by $\mathbb{G}_s$ times a polynomial $\mathfrak{p}$ of random variables

$$\max_{i=j,\dots,k}\left|\frac{D_{\mathsf{A}_\alpha}\mathbb{I}_s^i}{T_i - s}\right|, \quad \max_{i=j,\dots,k}\left[\frac{\mathbb{I}_{s\vee T_{i-1}}^i}{T_i - (s\vee T_{i-1})}\right]^{-1}, \quad \max_{i=j,\dots,k}\left|\frac{\mathbb{I}_s^i}{T_i - s}\right|, \quad |(u_i)_{i=j,\dots,k}|.$$

Indeed,

$$\begin{aligned}
&\left|D_{\mathsf{A}_\alpha}\bigg(\sum_{i_1,i_2=j}^{k} \mathbb{I}_s^{i_1\wedge i_2} u_{i_1} u_{i_2}\bigg)\right| \\
&= \left|\frac{\sum_{i_1,i_2=j}^{k} D_{\mathsf{A}_\alpha}\mathbb{I}_s^{i_1\wedge i_2}/(T_{i_1\wedge i_2} - s)u_{i_1}u_{i_2}}{\sum_{i_1,i_2=j}^{k} \mathbb{I}_s^{i_1\wedge i_2}/(T_{i_1\wedge i_2} - s)u_{i_1}u_{i_2}}\right|\left|\sum_{i_1,i_2=j}^{k} \mathbb{I}_s^{i_1\wedge i_2}/(T_{i_1\wedge i_2} - s)u_{i_1}u_{i_2}\right| \\
&\le \left|\Big(D_{\mathsf{A}_\alpha}\mathbb{I}_s^{i_1\wedge i_2}/(T_{i_1\wedge i_2} - s)\Big)_{i_1,i_2=j,\dots,k}\right|\left|\Big(\mathbb{I}_s^{i_1\wedge i_2}/(T_{i_1\wedge i_2} - s)\Big)_{i_1,i_2=j,\dots,k}^{-1}\right| \\
&\quad \times\left|\sum_{i_1,i_2=j}^{k} \mathbb{I}_s^{i_1\wedge i_2}/(T_{i_1\wedge i_2} - s)u_{i_1}u_{i_2}\right|,
\end{aligned}$$

where we used

$$|S^{-1/2}| \quad \lesssim \quad \|S^{-1/2}\|_{op} = \bigg(\sup_{v:|v|=1} S^{-1}[v^{\otimes 2}]\bigg)^{1/2} \le |S^{-1}|^{1/2}$$

for any non-degenerate symmetric matrix S. Moreover, the identity

$$\det\Big(\mathbb{I}_s^{i_1\wedge i_2}/(T_{i_1\wedge i_2} - s)\Big)_{i_1,i_2=j,\dots,k} = \prod_{i=j,\dots,k}\frac{\mathbb{I}_{s\vee T_{i-1}}^i}{T_i - (s\vee T_{i-1})}$$

can be used to estimate the inverse matrix in the above expression.

The term $\mathfrak{p}\mathbb{G}_s$ is $\mathbb{L}_p$-bounded due to (B1)(i), (iii) and

$$\sup_{\substack{u\in\mathbb{R}^k,\omega,\\ s\in(T_{j-1},T_j)}}\bigg(\sum_{i_1,i_2=j}^{k} \mathbb{I}_s^{i_1\wedge i_2} u_{i_1} u_{i_2}\bigg)^m \mathbb{G}_s < \infty$$

for every $m \in \mathbb{N}$ and $j = 1, ..., k$.

If $j = k$, then this estimate is sufficient for our use. When $j < k$, we also use the non-degeneracy of the matrix

$$\mathcal{M}(T_{j+1}, T_k) \quad = \quad \left(\frac{1}{2}\int_{T_j}^{T_{i_1 \wedge i_2}} K(t)^2 dt\right)_{i_1, i_2 = j+1, ..., k}$$

and the estimate

$$(C_1 - C_s)[u^{\otimes 2}] \quad \geq \quad \mathcal{M}(T_{j+1}, T_k)\big[(u_{j+1}, ..., u_k)^{\otimes 2}\big] \tag{6.7}$$

in order to obtain the decay $|(u_{j+1}, ..., u_k)|^{-(\check{\mathsf{d}}+1-\varepsilon)}$. For (6.7), we note that

$$\begin{aligned}
\sum_{i_1, i_2 = 1}^{k} \int_s^{s \vee T_{i_1 \wedge i_2}} K(t)^2 dt \, u_{i_1} u_{i_2} \quad &= \quad \int_s^1 \left(\sum_{i=1}^k 1_{[0, T_i]}(t) u_i\right)^2 K(t)^2 dt \\
&\geq \quad \int_{T_j}^1 \left(\sum_{i=1}^k 1_{[0, T_i]}(t) u_i\right)^2 K(t)^2 dt \\
&= \quad \sum_{i_1, i_2 = j+1}^{k} \int_{T_j}^{T_{i_1 \wedge i_2}} K(t)^2 dt \, u_{i_1} u_{i_2}.
\end{aligned}$$

By (2.7) we have

$$\det \mathcal{M}(T_{j+1}, T_k)^{-1} \quad \in \quad L_{\infty-}, \tag{6.8}$$

and hence (6.7) and (6.8) imply that

$$|(u_{j+1}, ..., u_k)|^m \exp\left(-\frac{1}{2}(C_1 - C_s)[u^{\otimes 2}]\right) \quad \leq \quad \mathsf{C}_m \left|\mathcal{M}(T_{j+1}, T_k)^{-1}\right|^m \tag{6.9}$$

is $\mathbb{L}_{\infty-}$-bounded uniformly in $(u_{j+1}, ..., u_k)$ for every $m \in \mathbb{N}$. Finally, we may use one of the above estimates of the decay, depending on $|(u_1, ..., u_j, v_1, ..., v_q)| \geq |(u_{j+1}, ..., u_k)|$ or not.

Following the proof of [22, Theorem 4], i.e. the procedure (a)-(g) therein with the additional truncation

$$\psi_{n,s}^j \quad = \quad \psi\Big(2\big[1 + 4 \det \sigma_{(M_{T_1}^{n,1}, ..., M_{T_{j-1}}^{n,j-1}, M_s^{n,j}, F)} (s_n^j)^{-1}\big]^{-1}\Big),$$

we obtain the desired decay of

$$n \sum_{i=1}^{m_n} \int_{t_{i-1}}^{t_i} 1_{s \in (\tau_n^j, T_j)} \int_r^{t_i} \mathbb{E}\Big[\partial^{\mathsf{a}_0} K^n(t_{i-1})[\mathtt{i}u] \partial^{\mathsf{a}_1} D_r \Big\{ e_s^n(u) D_s\big(\Psi(u, v)\psi_n\big)\Big\}\Big] ds dr$$

for $\alpha = 0$. A similar estimate can be shown for a general α.

For $s \in (T_{j-1}, \tau_n^j)$, we apply the integration-by-parts formula for

$$\left(M_{T_1}^{n,1}, ..., M_{T_{j-1}}^{n,j-1}, F\right)$$

to obtain the decay $|(u_1, ..., u_{j-1}, v_1, ..., v_q)|^{-(\check{\mathsf{d}}+1-\varepsilon)}$. In order to obtain the decay $|(u_j, ..., u_k)|^{-(\check{\mathsf{d}}+1-\varepsilon)}$, we use the nondegeneracy of

$$\mathcal{M}(\tau_n^j, T_k) \quad = \quad \left(\frac{1}{2}\int_{\tau_n^j}^{T_{i_1\wedge i_2}} K(t)^2 dt\right)_{i_1,i_2=j,\dots,k}.$$

Then we repeat a similar procedure as in the previous case to obtain the desired decay. We deduce (6.6) by combining the above estimates.

6.2 Proof of Theorem 3.2

We state the decompositions in the differential form for the ease of exposition. Applying Taylor expansion we conclude that

$$\begin{aligned} dV_t^n &= \sqrt{n}\left(a(X_{\varphi_n(t)}^n) - a(X_t)\right)dt + \sqrt{n}\left(b(X_{\varphi_n(t)}^n) - b(X_t)\right)dW_t \qquad (6.10) \\ &= \sqrt{n}\left(a(X_t^n) - a(X_t)\right)dt + \sqrt{n}\left(a(X_{\varphi_n(t)}^n) - a(X_t^n)\right)dt \\ &+ \sqrt{n}\left(b(X_t^n) - b(X_t)\right)dW_t + \sqrt{n}\left(b(X_{\varphi_n(t)}^n) - b(X_t^n)\right)dW_t \\ &= \left(\left(a'(X_t) + \tilde{a}_t^n\right)V_t^n + \tilde{a}_t'^n\right)dt \\ &+ \left(\left(b'(X_t) + \tilde{b}_t^n\right)V_t^n - \sqrt{n}bb'(X_{\varphi_n(t)}^n)(W_t - W_{\varphi_n(t)}) + \tilde{b}_t'^n\right)dW_t, \end{aligned}$$

where the processes $\tilde{a}^n, \tilde{a}'^n, \tilde{b}^n, \tilde{b}'^n$ are defined as

$$\begin{aligned} \tilde{a}_t^n &= a'(\widetilde{X}_t^n) - a'(X_t), \qquad \tilde{a}_t'^n = \sqrt{n}\left(a(X_{\varphi_n(t)}^n) - a(X_t^n)\right), \\ \tilde{b}_t^n &= b'(\widetilde{X}_t'^n) - b'(X_t), \qquad \tilde{b}_t'^n = \sqrt{n}\left(b(X_{\varphi_n(t)}^n) - b(X_t^n) + bb'(X_{\varphi_n(t)}^n)(W_t - W_{\varphi_n(t)})\right), \end{aligned}$$

and $\widetilde{X}_t^n, \widetilde{X}_t'^n$ are certain random variables with $|\widetilde{X}_t^n - X_t| \le |X_t^n - X_t|$, $|\widetilde{X}_t'^n - X_t| \le |X_t^n - X_t|$. In particular, it holds that $\widetilde{X}^n \xrightarrow{u.c.p.} X$ and $\widetilde{X}'^n \xrightarrow{u.c.p.} X$. Using Lemma 3.1 we thus can write

$$\begin{aligned} V_t^n = \Sigma_t^n \Bigg(\int_0^t (\Sigma_s^n)^{-1} \left(\tilde{a}_s'^n - \left(b'(X_t) + \tilde{b}_t^n\right)\left(\tilde{b}_s'^n - \sqrt{n}bb'(X_{\varphi_n(s)})(W_s - W_{\varphi_n(s)})\right)\right) ds \\ + \int_0^t (\Sigma_s^n)^{-1}\left(\tilde{b}_s'^n - \sqrt{n}bb'(X_{\varphi_n(s)})(W_s - W_{\varphi_n(s)})\right) dW_s \Bigg), \end{aligned}$$

where the process Σ^n is defined by

$$\Sigma_t^n = \exp\left(\int_0^t (b'(X_s) + \tilde{b}_s^n) dW_s + \int_0^t \left(a'(X_s) + \tilde{a}_s^n - \frac{1}{2}\left(b'(X_s) + \tilde{b}_s^n\right)^2\right) ds\right) \tag{6.11}$$

Comparing the representation of V_t^n with (3.4), we just need to show that

$$\Sigma^n \xrightarrow{u.c.p.} \Sigma, \tag{6.12}$$

$$\int_0^t \Sigma_s^{-1} \tilde{a}_s'^n ds \xrightarrow{u.c.p.} 0, \tag{6.13}$$

$$\int_0^t \Sigma_s^{-1} \tilde{b}_s'^n dW_s \xrightarrow{u.c.p.} 0, \tag{6.14}$$

$$\int_0^t \Sigma_s^{-1} \left(b'(X_t) + \tilde{b}_t^n\right) \left(\tilde{b}_s'^n - \sqrt{n} b b'(X_{\varphi_n(s)})(W_s - W_{\varphi_n(s)})\right) ds \xrightarrow{u.c.p.} 0, \tag{6.15}$$

where the process Σ has been defined in (3.3). Since both $\tilde{a}_s^n$ and $\tilde{b}_s^n$ are bounded as assumed in the beginning of Section 6, and $\tilde{a}^n \xrightarrow{u.c.p.} 0$, $\tilde{b}^n \xrightarrow{u.c.p.} 0$ (because $\widetilde{X}^n \xrightarrow{u.c.p.} X$, $\widetilde{X}'^n \xrightarrow{u.c.p.} X$ and $a, b \in C^2(\mathbb{R})$), we readily deduce the convergence at (6.12). To show the convergence at (6.13) we use the decomposition $\tilde{a}_s'^n = \tilde{a}_s'^{n,1} + \tilde{a}_s'^{n,2}$ with

$$\begin{aligned}
\tilde{a}_s'^{n,1} &= -\sqrt{n} a'(X_{\varphi_n(s)}^n) \int_{\varphi_n(s)}^s b(X_{\varphi_n(u)}^n) dW_u, \\
\tilde{a}_s'^{n,2} &= \sqrt{n}\left(a'(X_s''^n) - a'(X_{\varphi_n(s)}^n)\right)\left(X_{\varphi_n(s)}^n - X_s^n\right) \\
&\quad - \sqrt{n} a'(X_{\varphi_n(s)}^n) \int_{\varphi_n(s)}^s a(X_{\varphi_n(u)}^n) du,
\end{aligned}$$

where $X_s''^n$ is a certain random variable with $|X_s''^n - X_s^n| \le |X_s^n - X_{\varphi_n(s)}^n|$. Since $X_s''^n \xrightarrow{u.c.p.} X_s$ and all involved objects are assumed to be bounded, we conclude by (6.1) that

$$\mathbb{E}[|\tilde{a}_s'^{n,2}|] \le C\epsilon_n$$

with $\epsilon_n \to 0$ as $n \to \infty$. Thus, we obtain

$$\int_0^t \Sigma_s^{-1} \tilde{a}_s'^{n,2} ds \xrightarrow{u.c.p.} 0$$

by an application of (6.2). Now, we notice that $\mathbb{E}[\tilde{a}_s'^{n,1} | \mathcal{F}_{(i-1)/n}] = 0$ and $\mathbb{E}[|\tilde{a}_s'^{n,1}|^2] \le C$. Thus, we deduce that

$$\int_0^t \Sigma_s^{-1} \tilde{a}_s'^{n,1} ds = \int_0^t \Sigma_{\varphi_n(s)}^{-1} \tilde{a}_s'^{n,1} ds + \int_0^t \left(\Sigma_s^{-1} - \Sigma_{\varphi_n(s)}^{-1}\right) \tilde{a}_s'^{n,1} ds \xrightarrow{u.c.p.} 0,$$

which follows by a combination of (6.2) and (6.3). Indeed, it holds that

$$\int_0^t \Sigma^{-1}_{\varphi_n(s)} \tilde{a}'^{n,1}_s ds = -\sqrt{n} \sum_{i=1}^{[nt]} \Sigma^{-1}_{\frac{i-1}{n}} a'(X^n_{\frac{i-1}{n}}) b'(X^n_{\frac{i-1}{n}}) \int_{\frac{i-1}{n}}^{\frac{i}{n}} (W_s - W_{\frac{i-1}{n}}) ds + o_{\mathbb{P}}(1),$$

and (6.3) can be applied to the last line. Consequently, we have (6.13). Finally, we show the convergence at (6.14). Observe the decomposition

$$\begin{aligned}\tilde{b}'^n_s &= \sqrt{n} \left(b'(X''^n_s) - b'(X^n_{\varphi_n(s)}) \right) \left(X^n_{\varphi_n(s)} - X^n_s \right) \\ &\quad - \sqrt{n} b'(X^n_{\varphi_n(s)}) \left(\int_{\varphi_n(s)}^s a(X^n_u) du + \int_{\varphi_n(s)}^s b(X^n_u) - b(X^n_{\varphi_n(s)}) dW_u \right).\end{aligned}$$

As for the term $\tilde{a}'^{n,2}_s$ we deduce that $\mathbb{E}[|\tilde{b}'^n_s|] \leq C\epsilon_n$ with $\epsilon_n \to 0$ as $n \to \infty$. Hence, we obtain (6.14). The proof of (6.15) combines the proof methods of (6.13) and (6.14). Consequently,

$$\sup_{t \in [0,1]} |V^n_t - \overline{V}^n_t| \xrightarrow{\mathbb{P}} 0,$$

which completes the proof of Theorem 3.2. □

6.3 Proof of Theorem 3.3

The derivation of the second order stochastic expansion is more involved than the expansion of Theorem 3.2, but the underlying methodology is similar. For simplicity of exposition we sometimes use the same notations as in the previous section although they might have a different meaning. Instead of the first order approximation in the last line of (6.10), we may further develop

$$\begin{aligned} dV^n_t &= a'(X_t) V^n_t dt + (b'(X_t) V^n_t - \sqrt{n} b b'(X^n_{\varphi_n(t)})(W_t - W_{\varphi_n(t)})) dW_t \qquad (6.16) \\ &+ \Big(\frac{1}{2\sqrt{n}} a''(X_t)(V^n_t)^2 - \sqrt{n} b a'(X^n_{\varphi_n(t)})(W_t - W_{\varphi_n(t)}) \\ &- \sqrt{n} a a'(X^n_{\varphi_n(t)})(t - \varphi_n(t)) - \frac{\sqrt{n}}{2} b^2 a''(X^n_{\varphi_n(t)})(W_t - W_{\varphi_n(t)})^2 \Big) dt \\ &+ \Big(\frac{1}{2\sqrt{n}} b''(X_t)(V^n_t)^2 - \frac{\sqrt{n}}{2} b^2 b''(X^n_{\varphi_n(t)})(W_t - W_{\varphi_n(t)})^2 \\ &- \sqrt{n} a b'(X^n_{\varphi_n(t)})(t - \varphi_n(t)) \Big) dW_t + \tilde{a}^n_t dt + \tilde{b}^n_t dW_t, \end{aligned}$$

where $\tilde{a}^n_t$ and $\tilde{b}^n_t$ are stochastic processes, whose negligibility in the involved asymptotic expansions is shown in exactly the same manner as in (6.12)-(6.15) (although these terms have a different meaning in this subsection).

Now, recall the definition of the first order approximation $\overline{V}_t^n$ at (3.4). By Lemma 3.1 this process satisfies the stochastic differential equation

$$d\overline{V}_t^n = a'(X_t)\overline{V}_t^n dt + b'(X_t)\overline{V}_t^n dW_t - \sqrt{n}\Sigma_t\Sigma_{\varphi_n(t)}^{-1} bb'(X_{\varphi_n(t)}^n)(W_t - W_{\varphi_n(t)})dW_t$$
$$- \sqrt{n}\Sigma_t\Sigma_{\varphi_n(t)}^{-1} b'(X_t)bb'(X_{\varphi_n(t)}^n)(W_t - W_{\varphi_n(t)}))dt.$$

Observing the definition of the stochastic process $dR_t^n = R_t^n(1)dt + R_t^n(2)dW_t$ at (3.5), we deduce by Lemma 3.1 and the negligibility of the terms $\tilde{a}_t^n$, $\tilde{b}_t^n$ the decomposition

$$V_t^n - \overline{V}_t^n = \Sigma_t \int_0^t \Sigma_s^{-1} \left(\left(R_s^n(1) - b'(X_s)R_s^n(2)\right)\right) ds + R_s^n(2)dW_s\right) + o_{\mathbb{P}}(n^{-1/2}),$$

where Σ has been defined in (3.3). This finishes the proof of Theorem 3.3. □

6.4 Proof of Theorem 4.3

We will verify conditions $(B1)$, $(B2)$ and $(B3)$ for Theorem 2.1 under (A), $(C1)$, $(C2)$ and $(C3)$. Recall that

$$K(s) = -\Sigma_s^{-1} bb'(X_s),$$

$\ell = 2k + \overline{q} + 8$, $\ell_* = 2[\overline{q}/2] + 4$ and we are assuming that a, b are in $C^\infty(\mathbb{R})$ and all their derivatives of positive order are bounded. As mentioned just before assumption (A), the functionals $c_\alpha(z)$ in the representation (2.10) of the full random symbol σ and also in (4.9) are associated with $\underline{\sigma}$ of (4.8) and $\overline{\sigma}$ of (2.9).

Conditions $(B1)$(i), (ii) are obvious. Condition $(B1)$(iii) is assumed by $(C1)$. In the present situation, $\widehat{F}_n = 0$ since $F_n = F$. Condition $(B2)$(i) follows from (A) and $(C2)$(i). $(B2)$(ii) will be checked later after constructing s_n. Condition $(B2)$(iii) is already obtained in (4.7). The property $(B2)$(iv) has been observed to derive the expression (4.8).

We shall consider non-degeneracy of the Malliavin covariance matrix $\sigma_{(\mathbb{X}_1,\mathbb{X}_2)}$ of $(\mathbb{X}_1, \mathbb{X}_2)$, where

$$\mathbb{X}_1 = \left(M_{S_1}^{n,1}, ..., M_{S_{j-1}}^{n,j-1}, M_{S_j}^{n,j}\right) \quad \text{and} \quad \mathbb{X}_2 = \left(\Sigma_{T_1}, ..., \Sigma_{T_k}\right)$$

for $S_1 = T_1, ..., S_{j-1} = T_{j-1}$ and S_j is either $s \in [(T_{j-1} + T_j)/2, T_j]$. We will estimate the Malliavin covariance matrix $\sigma_{(\mathbb{X}_1,\mathbb{X}_2)}$. Let $\theta_i = i/n$. Let

$$\eta_i(t) = \sqrt{n}\big(W(\theta_i \wedge t) - W(\theta_{i-1} \wedge t)\big)$$

and

$$\xi_i(t) = n\left(\big(W(\theta_i \wedge t) - W(\theta_{i-1} \wedge t)\big)^2 - \big(\theta_i \wedge t - \theta_{i-1} \wedge t\big)\right).$$

Then, as in [21], we have

$$D_r M_{S_\mu}^{n,\mu} = \sum_{i=1}^n 2K(\theta_{i-1})\eta_i(S_\mu) 1_{(\theta_{i-1}\wedge S_\mu, \theta_i \wedge S_\mu]}(r)$$
$$+n^{-1/2} \sum_{i=1}^{n-1} \left(\sum_{i'=i+1}^n D_r K(\theta_{i-1})\xi_{i'}(S_\mu)\right) 1_{(\theta_{i-1}\wedge S_\mu, \theta_i \wedge S_\mu]}(r)$$

for $\mu = 1, ..., j$. Therefore,

$$\begin{aligned}\sigma(n,\mu_1,\mu_2) &:= \langle DM^n_{S_{\mu_1}}, DM^n_{S_{\mu_2}}\rangle_{\mathbb{H}}\\ &= \sum_{i=1}^n \int_{\theta_{i-1}}^{\theta_i}\left[2K(\theta_{i-1})\eta_i(S_{\mu_1}) + n^{-1/2}\sum_{i'=i+1}^n D_rK(\theta_{i-1})\xi_{i'}(S_{\mu_1})\right]1_{[0,S_{\mu_1}]}(r)\\ &\qquad\times\left[2K(\theta_{i-1})\eta_i(S_{\mu_2}) + n^{-1/2}\sum_{i'=i+1}^n D_rK(\theta_{i-1})\xi_{i'}(S_{\mu_2})\right]1_{[0,S_{\mu_2}]}(r)dr\\ &\quad+O_{\mathbb{L}^p}(n^{-1/2})\\ &= \tilde\sigma(n,\mu_1,\mu_2) + O_{\mathbb{L}^p}(n^{-1/2})\end{aligned}$$

for $\mu_1, \mu_2 = 1, ..., j$, where $\tilde\sigma(n,\mu_1,\mu_2) = \tilde\sigma_1(n,\mu_1,\mu_2) + \tilde\sigma_2(n,\mu_1,\mu_2)$ with

$$\tilde\sigma_1(n,\mu_1,\mu_2) = \sum_{i=1}^n \int_{\theta_{i-1}}^{\theta_i}\big(2K(\theta_{i-1})\eta_i(S_{\mu_1})\big)1_{[0,S_{\mu_1}]}(r)\big(2K(\theta_{i-1})\eta_i(S_{\mu_2})\big)1_{[0,S_{\mu_2}]}(r)dr$$

and

$$\begin{aligned}\tilde\sigma_2(n,\mu_1,\mu_2) &= \sum_{i=1}^n \int_{\theta_{i-1}}^{\theta_i}\left(n^{-1/2}\sum_{i'=i+1}^n D_rK(\theta_{i-1})\xi_{i'}(S_{\mu_1})\right)1_{[0,S_{\mu_1}]}(r)\\ &\qquad\times\left(n^{-1/2}\sum_{i'=i+1}^n D_rK(\theta_{i-1})\xi_{i'}(S_{\mu_2})\right)1_{[0,S_{\mu_2}]}(r)dr\end{aligned}$$

for $\mu_1, \mu_2 = 1, ..., j$. Moreover, for $G = (G^\nu)_{\nu=1,\dots,\overline{q}}$,

$$\begin{aligned}\sigma(n,\mu,\nu) &:= \langle DM^n_{S_\mu}, DG^\nu\rangle_{\mathbb{H}}\\ &= \sum_{i=1}^n \int_{\theta_{i-1}}^{\theta_i}\left[2K(\theta_{i-1})\eta_i(S_\mu) + n^{-1/2}\sum_{i'=i+1}^n D_rK(\theta_{i-1})\xi_{i'}(S_\mu)\right]1_{[0,S_{\mu_1}]}(r)\\ &\qquad\times D_rG^\nu dr + O_{\mathbb{L}^p}(n^{-1/2})\\ &= \tilde\sigma(n,\mu,\nu) + O_{\mathbb{L}^p}(n^{-1/2}),\end{aligned}$$

where

$$\tilde\sigma(n,\mu,\nu) = \sum_{i=1}^n \int_{\theta_{i-1}}^{\theta_i}\left(n^{-1/2}\sum_{i'=i+1}^n D_rK(\theta_{i-1})\xi_{i'}(S_\mu)\right)1_{[0,S_\mu]}(r)D_rG^\nu dr$$

Let

$$\tilde\sigma(n,\nu_1,\nu_2) = \int_0^1 D_rG^{\nu_1}D_rG^{\nu_2}dr.$$

Then it is easy to see that the matrix

$$\begin{bmatrix} \big(\tilde\sigma_2(n,\mu_1,\mu_2)\big) & \big(\tilde\sigma(n,\mu,\nu)\big)\\ \big(\tilde\sigma(n,\mu,\nu)\big)^\star & \big(\tilde\sigma(n,\nu_1,\nu_2)\big)\end{bmatrix}$$

is nonnegative definite. As we will see, the matrix $\big(\tilde{\sigma}(n,\nu_1,\nu_2)\big)$ is positive definite almost surely. Therefore,

$$\big(\tilde{\sigma}_2(n,\mu_1,\mu_2)\big)-\big(\tilde{\sigma}(n,\mu,\nu)\big)\big(\tilde{\sigma}(n,\nu_1,\nu_2)\big)^{-1}\big(\tilde{\sigma}(n,\mu,\nu)\big)^{\star}$$

is nonnegative definite, and hence

$$\begin{aligned} & \det\begin{bmatrix} \big(\tilde{\sigma}(n,\mu_1,\mu_2)\big) & \big(\tilde{\sigma}(n,\mu,\nu)\big) \\ \big(\tilde{\sigma}(n,\mu,\nu)\big)^{\star} & \big(\tilde{\sigma}(n,\nu_1,\nu_2)\big) \end{bmatrix} \\ = \ & \det\Big[\ \big(\tilde{\sigma}(n,\mu_1,\mu_2)\big)-\big(\tilde{\sigma}(n,\mu,\nu)\big)\big(\tilde{\sigma}(n,\nu_1,\nu_2)\big)^{-1}\big(\tilde{\sigma}(n,\mu,\nu)\big)^{\star}\ \Big] \\ & \times\det\big(\tilde{\sigma}(n,\nu_1,\nu_2)\big) \\ \geq \ & \det\big(\tilde{\sigma}_1(n,\mu_1,\mu_2)\big)\det\big(\tilde{\sigma}(n,\nu_1,\nu_2)\big) \ =: \ \mathcal{M}_n. \end{aligned}$$

Now $\mathcal{M}_n$ converges in $\mathbb{L}_{\infty-}$ to

$$\mathcal{M}_\infty \ := \ \det\left[\int_0^{S_{\mu_1}\wedge S_{\mu_2}} 4K(t)^2dt\right]_{\mu_1,\mu_2=1,\dots,j} \times\det\left[\int_0^1 D_rG^{\nu_1}D_rG^{\nu_2}dr\right]_{\nu_1,\nu_2=1,\dots,\overline{q}}$$

with rate $n^{-1/2}$. Define s_n^j by

$$s_n^j := \frac{1}{2}\mathcal{M}_\infty.$$

Then $\sup_{n\in\mathbb{N}}\big\|s_n^j\big\|_{\ell,p}<\infty$ for every $p>1$ and every j, so that $(B2)$(ii) holds additionally by (A). But $\mathcal{M}_\infty$ is non-degenerate, i.e.

$$\mathcal{M}_\infty^{-1} \ \in \ \mathbb{L}_{\infty-} \tag{6.17}$$

due to $(C1)$ and $(C3)$. This shows $(B3)$(ii). Moreover, the estimate $\mathcal{M}_n-\mathcal{M}_\infty = O_{\mathbb{L}_{\infty-}}(n^{-1/2})$ and (6.17) proves $(B3)$(i). Hence, the proof of Theorem 4.3 is completed. □

Acknowledgment

Mark Podolskij and Bezirgen Veliyev acknowledge financial support from the project "Ambit fields: probabilistic properties and statistical inference" funded by Villum Fonden (No. 11745) and from CREATES funded by the Danish National Research Foundation. The research of Nakahiro Yoshida is supported in part by Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research No. 17H01702 (Scientific Research), No. 26540011 (Challenging Exploratory Research); Japan Science and Technology Agency CREST JPMJCR14D7; and by a Cooperative Research Program of the Institute of Statistical Mathematics.

References

[1] Aldous, D. J., and G.K. Eagleson (1978): On mixing and stability of limit theorems. *Ann. of Probab.* 6(2), 325–331.

[2] Arnold, S. (2006): Approximation schemes for SDEs with discontinuous coefficients. *PhD thesis*, ETH Zürich.

[3] Bally, V. and D. Talay (1996): The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function. *Probab. Theory Related Fields* 104 (1), 43–60.

[4] Bally, V. and D. Talay (1996): The law of the Euler scheme for stochastic differential equations. II. Convergence rate of the density. *Monte Carlo Methods Appl.* 2(2), 93–128.

[5] Barndorff-Nielsen, O.E., Graversen, S.E., Jacod, J., Podolskij, M. and N. Shephard (2006): A central limit theorem for realised power and bipower variations of continuous semimartingales. In: Yu. Kabanov, R. Liptser and J. Stoyanov (Eds.), From Stochastic Calculus to Mathematical Finance. Festschrift in Honour of A.N. Shiryaev, Heidelberg: Springer, 2006, 33–68.

[6] Chan, K. S. and O. Stramer (1998): Weak consistency of the Euler method for numerically solving stochastic differential equations with discontinuous coefficients. *Stochastic Process. Appl.* 76(1), 33–44.

[7] Dalalyan, A. and N. Yoshida (2011): Second-order asymptotic expansion for a non-synchronous covariation estimator. *Ann. Inst. H. Poincaré Probab. Statist.* 47, 3, 748-789.

[8] Halidias, N. and P.E. Kloeden (2006): A note on strong solution of stochastic differential equations with discontinuous a drift coefficient.*J. Appl. Math. Stoch. Anal.*

[9] Jacod, J. and P. Protter (1998): Asymptotic error distributions for the Euler method for stochastic differential equations. *Ann. Probab.*, 26, 267–307.

[10] Jacod, J. and A.N. Shiryaev (2003): *Limit Theorems for Stochastic Processes*, 2d ed., Springer-Verlag: Berlin.

[11] Kloeden, P. E. and E. Platen (1992): *Numerical solution of stochastic differential equations.* Applications of Mathematics (New York), vol. 23, Springer-Verlag Berlin.

[12] Kurtz, T. G. and P. Protter (1991): Wong–Zakai corrections, random evolutions and numerical schemes for SDEs. In *Stochastic Analysis* 331–346. Academic Press, New York.

[13] Mikulevičius, R. and E. Platen (1991): Rate of convergence of the Euler approximation for diffusion processes. *Math. Nachr.* 15, 233–239.

[14] Nualart, D. (2006): *The Malliavin calculus and related topics.* Springer-Verlag, Berlin, 2nd edition.

[15] Podolskij, M. and N. Yoshida (2016): Edgeworth expansion for functionals of continuous diffusion processes. *Ann. Appl. Probab.* 26(6), 3415–3455.

[16] Podolskij, M., Veliyev, B. and N. Yoshida (2017): Edgeworth expansion for the pre-averaging estimator. *Stochastic Process. Appl.* 127(11), 3558–3595.

[17] Renyi, A. (1963): On stable sequences of events. *Sankhya* A 25, 293-302.

[18] Ikeda, N. and S. Watanabe (1989): *Stochastic differential equations and diffusion processes*. North-Holland Publishing Co., Amsterdam, 2nd edition.

[19] L. Yan (2002): The Euler scheme with irregular coefficients. *Ann. Probab.* 30(3), 1172–1194.

[20] Yoshida, N. (1997): Malliavin calculus and asymptotic expansion for martingales. *Probab. Theory Related Fields*, 109, 301-342.

[21] Yoshida, N. (2012): Asymptotic expansion for the quadratic form of the diffusion process. Working paper. Available on arXiv:1212.5845.

[22] Yoshida, N. (2013): Martingale expansion in mixed normal limit. *Stochastic Process. Appl.* 123(3), 887–933.