J. Math. Kyoto Univ. (JMKYAZ)
37-1 (1997) 129-150

High order Itô-Taylor approximations to heat kernels*

By

Arturo KOHATSU-HIGA

1. Introduction

Let $(\Omega, \mathcal{F}, P)$ be the canonical space of the standard m dimensional Wiener process $W=(W^1, \cdots, W^m)$. On this space, define X as the solution of the following SDE (stochastic differential equation):

$$X_t = x + \sum_{i=1}^{m} \int_0^t B_i(X_s)\, dW_s^i + \int_0^t B_0(X_s)\, ds, \quad 0 \le t \le T, \tag{1.1}$$

where, $x \in \mathbf{R}^d$ and B_i: $\mathbf{R}^d \to \mathbf{R}^d$, $i=0, \cdots, m$ are smooth vector fields with bounded derivatives. It is known that under the condition that

$$(\det \sigma_{X_t})^{-1} \in \bigcap_{p \ge 1} L^p(\Omega), \tag{1.2}$$

X_t has a smooth density, say $q(t; x, y)$. Here σ_F denotes the Malliavin covariance matrix of the random variable F.

The purpose of this article is to find ways of approximating $q(t; x, y)$. Recently, Hu-Watanabe [4] and Bally-Talay [1], have obtained results on this problem. Let's introduce these results. Define the following sets of vector fields

$$\Sigma_0 = \{B_j, j=1, \cdots, m\}$$
$$\Sigma_j = \{[B_k, V]; V \in \Sigma_{j-1}, k=0, \cdots, m\}, j \ge 1,$$

where $[\cdot,\cdot]$ denotes the Lie bracket. Now define for $A \ge 1$, the quadratic forms

$$V_A(x, \eta) := \sum_{V \in \Sigma_{A-1}} \langle V(x), \eta \rangle^2$$

and set

$$V_A(x) = 1 \wedge \inf_{|\eta|=1} V_A(x, \eta). \tag{1.3}$$

Communicated by Prof. S. Watanabe, May 20, 1996

* This article was written while the author was visiting the Department of Mathematics at Kyoto University with a JSPS fellowship

It is known that $V_A(x) > 0$ for some $A \in \mathbf{N}$ implies (1.2) and therefore the existence and smoothness of $q(t; x, y)$.

Theorem 1.1 *(Bally-Talay). Let $A \in \{1, 2, \cdots\}$ be such that $U_A = \{x; V_A(x) > 0\}$ is non void and let x and y be in U_A, so that*

$$V_A(x) \wedge V_A(y) > 0.$$

Then there exists a nondecreasing function $K(\cdot)$, there exists some strictly positive constants c, r, r', r'' and a function $\pi_t(x, y)$ and for each $n > \frac{2}{|x-y|}$, there exists a function $R_t^n(x, y)$ such that the density of the perturbed Euler-Maruyama scheme $\tilde{q}^n$ with uniform step-size n^{-1} satisfies

$$q(t; x, y) - \tilde{q}^n(t; x, y) = -\frac{1}{n}\pi_t(x, y) + \frac{1}{n^2}R_t^n(x, y) \tag{1.4}$$

with

$$|\pi_t(x, y)| + |R_t^n(x, y)| \leq \frac{K(t)}{t^r V_A(x)^{r'} V_A(y)^{r''}} \exp\left(-c\frac{|x-y|^2}{t}\right). \tag{1.5}$$

The function $K(\cdot)$ depends on the $L^m(\mathbf{R}^d)$ norms (for some integer m) of a finite number of partial derivatives of the function $\rho_0(\cdot)$.

Here $\tilde{q}_t^n(x, y)$ is the density of the sum of the Euler-Maruyama scheme at time t and an independent random variable with a density defined through the function ρ_0 (for details see [1]).

In this article we intend to find an expansion of the type (1.4) for high order Itô-Taylor approximations, therefore including the Euler-Maruyama scheme. Furthermore we will get rid of the conditions $V_A(y) > 0$ and $n > \frac{2}{|x-y|}$. Nevertheless we lose the explicit expressions available for the coefficient functions in the result of Theorem 1.1. But we will also prove that the coefficient functions satisfy inequalities similar to (1.5).

Another result on this topic has been obtained using Donsker's delta functions by Hu-Watanabe [4]. Let F_n denote the strong Itô-Taylor approximation scheme of order β at time t associated to a partition of size n^{-1} (for a definition see [4] or [6]). The following is a simplified version of Theorem 3.2 in [4].

Theorem 1.2 *(Hu-Watanabe). Assume (1.2). Then for $\delta > 0$,*

$$\sup_y |\partial_y^\alpha (q(t; x, y) - E(\phi_{n^{-\delta}}(F_n - y)))| \leq \frac{C}{n^{\beta \wedge \delta}} \tag{1.6}$$

where $\phi_r(x)$ denotes the density of a d-dimensional normal randon variable with mean zero and covariance matrix $r^2 I_{d \times d}$ and for α a multi-index, ∂_y^α denotes the high order derivative with respect to the coordinates indicated in α. C is a constant that depends on α but it is independent of n and δ.

This theorem is obtained by applying some general results about approximation of densities of random variables on Wiener space, and therefore is inspired in strong approximation techniques. We applied a slight modification of weak approximation techniques to the problem of approximating $q(t; x, y)$ obtaining the natural improvement of rates of convergence. For example, in the case $\beta=0.5$, $\delta\geq 0.5$ (the Euler-Maruyama scheme), we have improved the rate in (1.6) to Cn^{-1}. The modification of the weak approximation technique that we will apply in this article can be explained as follows.

Consider the weak approximation problem

$$|E(f_m(\tilde{X}_t)-f_m(X_t))|\leq\frac{C(m)}{n^{\beta}}, \tag{1.7}$$

where $\tilde{X}_t$ is a high order Itô-Taylor weak approximation of order β, stepsize $\frac{1}{n}$, $\{f_m;\ m\in\mathbf{N}\}$ is a sequence of smooth functions with polynomial growth at infinity converging to the delta function. Therefore the idea to obtain our results is centered in proving that $\sup_m C(m)<\infty$. To prove this, one has to obtain very detailed expressions of the difference in (1.7). This will invariably take us to consider derivatives of f_m which are undesirable if one is to prove boundedness of $C(m)$. This problem is solved by using the integration by parts formula of Malliavin calculus. Then to finish one only takes limits with respect to m in (1.7). During this procedure we find the problem of the existence of densities for $\tilde{X}$. In order to obtain such a property we perturb $\tilde{X}$ slightly with an independent normally distributed random variable. The idea of perturbing a random variable to obtain existence and smoothness of densities has been successfully used before by Bally-Talay and Hu-Watanabe. This method gives a bound to the speed of the weak convergence of the approximation to $q(t; x, y)$. To obtain an expansion of the error in terms of the powers of the step size one uses the above idea combined with the methodology developed by Kloeden, Platen and Hoffmann in [7].

Our main results are Theorem 3.1 and Theorem 5.2. Theorem 3.1 expands the result of Theorem 1.2 by proving that the rates of convergence for this problem can be considered as weak approximation rates. Furthermore a generalization of this result (Theorem 4.1) is one of the ingredients in the proof of Theorem 5.2. Theorem 5.2 proves the existence of an expansion of the error in terms of the stepsize of the approximation, under the hypothesis that $V_A(x)>0$, therefore improving on the results of Theorem 1.1. Theorem 5.3 is a slight extension of Theorem 5.2. This extension deals with the approximation of a heat kernel with a potential.

These results could have a variety of applications when coupled with Monte Carlo methods in order to simulate approximations of heat kernels.

Furthermore the development of the error of approximation in terms of powers of the step size provides a way for constructing extrapolation methods.

2. Preliminaries

We start by introducing some notation mostly taken from Kloeden and Platen [6] . On the canonical m-dimensional Wiener space let $\varphi_t(s,x)$, for $0 \leq s \leq t \leq T$, be the flow defined by the following stochastic differential equation (sde):

$$\varphi_t(s, x) = x + \sum_{i=0}^{m} \int_s^t B_i(\varphi_u(s, x)) dW_u^i.$$

where B_j: $\mathbf{R}^d \to \mathbf{R}^d$ are smooth functions with bounded derivatives for $j = 0$, $1 \cdots, m$. We use the notation $dW_u^0 = du$. We then have that $X_t = \varphi_t(0, x)$. Let $\mathcal{M}_m = \{(j_1, \cdots, j_l); j_i \in \{0, \cdots, m\}, i \in \{1, \cdots, l\}, \text{ for } l = 1, \cdots\} \cup \{v\}$ where v denotes the multi-index of lenght 0. For a multi-index $\alpha = (j_1, \cdots, j_l)$ define the length of α as $l(\alpha) = l$, also define $n(\alpha)$ as the number of zeros in α, $-\alpha = (j_2, \cdots, j_l)$ and $\alpha - = (j_1, \cdots, j_{l-1})$. Then for f: $[0, T] \times \mathbf{R}^d \to \mathbf{R}$ define the following operators:

$$L^j f(s, x) = \sum_{k=1}^{d} B_j^k(x) \frac{\partial f}{\partial x_k}(s, x), j = 1, \cdots, m$$

$$L^0 f(s, x) = \frac{\partial f}{\partial s}(s, x) + \sum_{k=1}^{d} B_0^k(x) \frac{\partial f}{\partial x_k}(s, x) + \frac{1}{2} \sum_{k,l=1}^{d} \sum_{j=1}^{m} B_j^k(x) B_j^l(x) \frac{\partial^2 f}{\partial x_k \partial x_l}(s, x)$$

where v^j denotes the j-th component of the vector $v, j = 1, \cdots, d$.

For $\alpha = (j_1, \cdots, j_l)$ define by induction

$$f_\alpha = \begin{cases} f & ; \text{if } l = 0 \\ L^{j_1} f_{-\alpha} & ; \text{if } l \geq 1. \end{cases} \tag{2.1}$$

In the case that the function f is not explicitly stated we shall always take it to be the identity function $f(t, x) = x$. Also define the following Wiener functionals for v: $[0, T] \times \Omega \to \mathbf{R}^d$ an adapted càdlàg, $L^2([0, T])$ integrable w.p. 1 stochastic process:

$$I_\alpha[v(\cdot)]_{s,t} = \begin{cases} v(t) & ; \text{if } l = 0 \\ \int_s^t I_{\alpha-}[v(\cdot)]_{s,u} dW_u^{j_l} & ; \text{if } l \geq 1. \end{cases}$$

$I_{\alpha,s,t} = I_\alpha[1]_{s,t}$. Also let $\Gamma_\beta = \{\alpha \in \mathcal{M}_m; l(\alpha) \leq \beta\}$ and $B(\Gamma_\beta) = \{\alpha \in \mathcal{M}_m - \Gamma_\beta; -\alpha \in \Gamma_\beta\}$. We also define multi-indexes for the derivatives in $\mathbf{R}^d$. For this, let $P_l \equiv P_l^d = \{1, \cdots, d\}^l$, for $p = (p_1, \cdots, p_l) \in P_l$, define

$$\partial_y^p = \frac{\partial^l}{\partial y_{p_1} \cdots \partial y_{p_l}}, \quad F_p(y) = \prod_{h=1}^{l} y^{p_h}.$$

Now we introduce some basic tools from Malliavin calculus that will be used throughout the text. For further reference see [9]. Let $C_b^\infty(\mathbf{R}^{mn})$ be the set of C^∞ functions f: $\mathbf{R}^{nm} \to \mathbf{R}$ which are bounded and have bounded derivatives of all orders. The class of real random variables of the form $f(W_{t_1}, \cdots, W_{t_n})$, $f \in C_b^\infty(\mathbf{R}^{nm})$, is denoted by $\mathcal{S}$. $\mathbf{D}^{1,p}$ designates the Banach space which is the completion of $\mathcal{S}$ with respect to the norm:

$$\|F\|_{1,p} = \{E|F|^p\}^{1/p} + \left(\sum_{j=1}^m E\left[\left\{\int_0^1 |D_s^j F|^2 ds\right\}^{p/2}\right]\right)^{1/p},$$

where

$$D_s^j F = \sum_{i=1}^n \frac{\partial f}{\partial x_{ji}}(W_{t_1}, \cdots, W_{t_m}) \mathbf{I}_{[0,t_i]}(s).$$

$\mathbf{D}^{\alpha,p}$ is defined analogously and its associated norm is denoted by $\|\cdot\|_{\alpha,p}$. Also, let $\mathbf{D}^\infty = \bigcap_{p\geq 1} \bigcap_{\alpha \geq 1} \mathbf{D}^{\alpha,p}$. As with ∂_y^p we will also use the notation D_u^p for $p \in P_l^m$ and $u \in [0, T]^l$. By extension we denote $\|\cdot\|_{0,p} \equiv \|\cdot\|_p$ the norm in $L^p(\Omega)$.

Now we will introduce the conditions associated with the existence and smoothness of the law of the solution to (1.1). Let's define the Malliavin covariance matrix of F as $\sigma_F^{ij} = \langle DF^i, DF^j \rangle_{L^2[0,T]}$.

Let V_A defined as in (1.3). It is known that $V_A(x) > 0$ for some $A \in \mathbf{N}$ implies (1.2) and therefore one obtains the existence and smoothness of the density of the law of the solution to (1.1). We will assume that $V_A(x) > 0$ for some $A \in \mathbf{N}$ throughout the rest of the article. Furthermore Kusuoka and Stroock [5] (Corollary 3.25) obtained the following estimation for $(\det \sigma_{X_t})^{-1}$.

$$\|(\det \sigma_{X_t})^{-1}\|_p \leq \frac{M_p(1+|x|)^\mu}{V_A(x)^c t^\kappa} \tag{2.2}$$

where M_p, μ, c and κ depend only on A. We will use this estimation throughout the text without further mentioning. Also, from now on we will use different notations for constants that may change from one line to the next although we use the same symbols. The dependence of these constants on the different data of the problem is explicitly stated at each equation. We will consider from now on that d, m, A, $\beta \in \mathbf{N}$, $T \in \mathbf{R}_+$, B_i, $i=0, 1, \cdots, m$ are fixed throughout the article.

3. Bound for the error of approximation

First we introduce the Itô-Taylor weak approximation to (1.1) of order β. We denote it by $\{\tilde{X}_t; t \in [0, T]\}$ and it is defined as

$$\tilde{X}_t = \sum_{\alpha \in \Gamma_\beta} f_\alpha(\tilde{X}_{\tau_n}) I_{\alpha,\tau_n,t} \quad \text{for } \tau_n \leq t \leq \tau_{n+1},$$
$$\tilde{X}_0 = x.$$

where $\pi: 0=\tau_0<\tau_1<\cdots<\tau_N=T$ is a partition of $[0, T]$ such that $\sup_{i=0,\dots,N-1}(\tau_{i+1}-\tau_i)\le\delta$. Define $\eta(t)=\sup\{\tau_i;\ \tau_i\le t\}$ and $n(t)\equiv n_t$ as the integer such that $\eta(t)=t_{n(t)}$. To simplify notation we will use $\mathcal{F}_n=\mathcal{F}_{\tau_n}$ where $\{\mathcal{F}_t;\ t\in[0, T]\}$ is the natural filtration generated by W. In the same spirit we will do the same for processes, that is, $X_n\equiv X_{\tau_n}$, etc.

Note that although strong and weak approximations are different in nature one can still obtain the following properties:

$$\begin{aligned} &X_t\in\mathbf{D}^\infty \text{ and } \|\tilde{X}_t-X_t\|_{b,e}\le C\delta^{\beta'}, \\ &\sup_{u\in[0,T]^b} E\sup_{t\in[0,T]}|D_u^\alpha\tilde{X}_t-D_u^\alpha X_t|^e\le C\delta^{\beta' e}, \end{aligned} \tag{3.1}$$

where $\beta'=\frac{\beta}{2}$, C is a positive constant that depends only on $b\in\{0, 1, \cdots\}$, $e\ge 2$ and $\alpha\in P_b^m$. The method of proof of these assertions is the same as in the proof of Theorem 3.1 in [4].

In order to find an expression for the error in powers of δ we first have to prove that the rate of convergence of the scheme $\tilde{X}$, in the weak sense, is of order β. We will prove this result after a series of Lemmas. Define $e(\delta)=C_1\delta^{\frac{\beta}{2}}$.

Theorem 3.1. *Assume that $V_A(x)>0$, $C_1>0$ and $t>0$. Then for $p\in\mathcal{M}_d$, $n(p)=0$, we have that there exists positive constants M, μ, κ and c that depend only on the multi-index p and therefore all the constants are independent of x, y, t, δ and the partition π such that*

$$\sup_y|\partial_y^p E[\phi_{e(\delta)}(\tilde{X}_t-y)]-\partial_y^p q(t;x,y)|\le\frac{M(1+|x|)^\mu}{t^\kappa V_A(x)^c}\delta^\beta,$$

where $\phi_r(x)$ denotes the density of an d-dimensional normal random variable with mean zero and covariance matrix $r^2 I_{d\times d}$. $q(t;x,y)$ denotes the density of X_t at $y\in\mathbf{R}^d$.

From now on we assume that $p\in\mathcal{M}_d$, $n(p)=0$. Also to make the future notation easier let $g(\delta)=C\delta^{\frac{\beta}{2}}$, $h(\delta)=K\delta^{\frac{\beta}{2}}$ such that $e(\delta)=\sqrt{K^2+C^2}\,\delta^{\frac{\beta}{2}}$ with $C>0$ and $K\ge 0$. Note that due to their different nature there should be no confusion between $e(\delta)$ and e used in (3.1).

Lemma 3.1. *$e^2(\delta)=h^2(\delta)+g^2(\delta)$ implies that*

$$\partial_y^p E[\phi_{e(\delta)}(\tilde{X}_t-y)]=\partial_y^p\bar{E}[\phi_{g(\delta)}(\tilde{X}_t+h(\delta)\bar{W}_T-y)] \tag{3.2}$$

where $\{\bar{W}_t;\ t\in[0, T]\}$ is a d-dimensional Wiener process independent of W. $\bar{E}$ denotes the expectation on the extended sample space supporting $(W, \bar{W})$.

The proof of this Lemma is obtained by a direct integration of $\bar{W}$ in the right side of (3.2). To avoid excessive notation we will stop using the $\bar{E}$

notation in the future assuming that is understood that we are considering the extended sample space supporting $(W, \bar{W})$. We will keep using the notation $\{\mathcal{F}_t; t\in[0, T]\}$ to denote the filtration generated by $\{W_s; s\in[0, T]\}$.

The next lemma is an application of the integration by parts formula in the version that we will frequently use in our calculations.

Lemma 3.2. *Let $s<t$. Let U and Z be $\mathbf{R}^d$ and $\mathbf{R}$ valued random variables respectively, that belong to $\mathbf{D}^\infty$. Then there exists positive integer constants α_1, α_2, C, p', e, b, that depend only on p such that*

$$\sup_y |E[\partial_y^p \phi_{g(\delta)}(\varphi_t(s, U)+h(\delta)\bar{W}_T-y)Z]|$$
$$\leq C\|(\det \sigma_{\varphi_t(s, U)+h(\delta)\bar{W}_T})^{-1}\|_{p'}^{\alpha_1}\|\varphi_t(s, U)+h(\delta)\bar{W}_T\|_{b,e}^{\alpha_2}\|Z\|_{b,e}.$$

The proof of this lemma is simple and involves applying integration by parts $l(p)+d$-times. Then one applies standard inequalities.

The next lemma provides a bound for $\|(\det \sigma_{\varphi_t(s, U)+h(\delta)\bar{W}_T})^{-1}\|_{p'}$ for some particular values of U. The basis of this Lemma can be traced back to [8].

Lemma 3.3. *The following inequality is satisfied for any $p'\in\mathbf{N}$, $K>0$*

$$\sup_{\{(s,\tau)\in[0,t]^2; s\geq\tau\geq s-2\delta\}} \|(\det \sigma_{\varphi_t(s, U_\tau)+h(\delta)\bar{W}_T})^{-1}\|_{p'}\leq\frac{M(1+|x|)^\mu}{V_A(x)^c t^\kappa},$$

where U_τ is equal to either of the following processes: $\tilde{X}_\tau$, $\tilde{X}_{n-1}+\theta(\tilde{X}_n-\tilde{X}_{n-1})$, $\tilde{X}_{n-1}+\theta(\varphi_n(\tau_{n-1}, \tilde{X}_{n-1})-\tilde{X}_{n-1})$ or $\varphi_n(\tau_{n-1}, \tilde{X}_{n-1})$, for $\tau_n\leq\tau<\tau_{n+1}$ where $\theta\in[0, 1]$ is a parameter in $[0, 1]$. Here M, μ, c and κ are positive constants that depend only on p', and in particular do not depend on x, t, θ, π or δ.

Proof. As in Lemma 2.1 in [4], let's denote $H_\delta=\det \sigma_{\varphi_t(s, U_\tau)+h(\delta)\bar{W}_T}$, $H=\det \sigma_{X_t}$.
Then one obtains that

$$EH_\delta^{-p'}\leq 2^{-p'}EH^{-p'}+(EH^{-4k})^{\frac{1}{2}}2^{2k}(EH_\delta^{-2p'})^{\frac{1}{2}}(E(H_\delta-H)^{4k})^{\frac{1}{2}}, \tag{3.3}$$

for any $k\in\mathbf{N}$. In any of the cases considered we have the following estimates

$$EH_\delta^{-2p'}\leq C(p')h(\delta)^{-4p'(d+m)}, \quad E(H_\delta-H)^{4k}\leq C(k)\delta^{\beta(k)}, \tag{3.4}$$

where $\beta(k)$ is a positive increasing function of k, such that $\beta(k)\uparrow\infty$ as $k\longrightarrow\infty$. For example, to prove (3.4) in the case $U=\tilde{X}$, one has that for any vector $v\in\mathbf{R}^d$,

$$v^T\sigma_{\varphi_t(s,\tilde{X}_\tau)+h(\delta)\bar{W}_T}v\geq h(\delta)^2T\|v\|^2.$$

Note that as $\det C\geq(\inf_{|v|=1}v^TCv)^{d+m}$ for any $(d+m)\times(d+m)$ matrix C, the proof of (3.4) follows. The proof of the Lemma is obtained by taking k big enough in (3.3). □

The next lemma deals with one aspect of the problem we are interested in solving. That is, the particular case when the approximation is the process itself.

Lemma 3.4. *Let $s \in [0, T]$ and Z be $\mathbf{R}^{\varepsilon}$ valued random variable that is $\mathcal{F}_T$-measurable ($\varepsilon \in \mathbf{N}$) such that $Z \in \mathbf{D}^{\infty}$. Furthermore, let G: $\mathbf{R}^d \times \mathbf{R}^{\varepsilon} \to \mathbf{R}$ be a smooth function with polynomial growth at infinity. Then for $\gamma \in \mathbf{N}$*

$$\partial_y^p E[\phi_{e(\delta)}(X_t - y) G(X_s, Z)] - \partial_y^p (q(t; x, y) E(G(X_s, Z)/X_t = y)) = \sum_{1 \leq j \leq \gamma} \{e(\delta)^{2j} A_j(s, t; x, y)\} + B_\gamma(s, t; x, y, \delta). \tag{3.5}$$

such that for $1 \leq j \leq \gamma$, there exists function A_j: $[0, T]^2 \times (\mathbf{R}^d)^2 \to \mathbf{R}$, B_γ: $[0, T]^2 \times (\mathbf{R}^d)^2 \times \mathbf{R}_+ \to \mathbf{R}$ positive constants M, μ, c,κ that depend only on p with

$$\sup_y |A_j(s, t; x, y)| \leq \frac{M(1+|x|)^{\mu}}{V_A(x)^c t^{\kappa}} \tag{3.6}$$

$$\sup_y \left| \frac{B_\gamma(s, t; x, y, \delta)}{\delta^{\beta(\gamma+1/2)}} \right| \leq \frac{M(1+|x|)^{\mu}}{V_A(x)^c t^{\kappa}} \tag{3.7}$$

Furthermore the following estimate is satisfied

$$\sup_y |\partial_y^p (q(t; x, y) E(G(X_s, Z)/X_t = y))| \leq \frac{M(1+|x|)^{\mu}}{V_A(x)^c t^{\kappa}} \tag{3.8}$$

In particular, note that the constants above do not depend on x, y, t, δ or s.

Sketch of the Proof. We will do the proof for $\gamma = 2$, $d = 1$. Let $a > 0$, $n \in \mathbf{N}$. Define $F(x_1) = \partial_y^p \phi_{n^{-a}}(\varphi_t(x) + x_1 - y) G(\varphi_s(x), Z)$. Then using Lemma 3.1 and Taylor's formula for F one obtains:

$$E[F(e(\delta)\bar{W}_T) - F(0)] = E\Big[\sum_{1 \leq j \leq 2} \frac{1}{j!} \Big(\frac{d^j}{dx_1^j} F(0) (e(\delta)\bar{W}_T)^j \Big) + \int_0^1 \int_0^v \frac{d^3}{dx_1^3} F(ue(\delta)\bar{W}_T)(ue(\delta)\bar{W}_T)^3 du dv \Big]$$

The residue B_1^n is defined as the last term in the above equation. Note that $F(0)$ and its derivatives are $\mathcal{F}_t$-measurable. Therefore by the independence between W and $\bar{W}$, one finds that all the odd order terms in the first term on the right of the above equation have expectation 0. From the even order terms one defines A_j^n. For example,

$$A_1^n(s, t; x, y) = -\frac{T}{2} E\Big[\frac{d^2}{dy^2} \partial_y^p \phi_{n^{-a}}(X_t - y) G(X_s, Z) \Big].$$

Using (2.2) and Lemma 3.2 with n^{-a} playing the role of $g(\delta)$ and $s = 0$, $K = 0$, one obtains that (3.6) is satisfied with A^n instead of A. Note that the constants do not depend on n.

Then by taking limits with respect to n one obtains (using, e.g., Theorem 2.1 in [4]) that the above converges uniformly in y to a function of A_1 defined as

$$A_1(s, t; x, y) = -\frac{T}{2}\frac{d^2}{dy^2}\partial_y^p(q(t; x, y)E(G(X_s, Z)/X_t=y)).$$

Similarly, $E[F(e(\delta)\bar{W}_T) - F(0)]$ converges to the term on the left side of (3.5).

Properties (3.6) and (3.7) are obtained by taking limits. Repeating the above argument one obtains (3.8). That is, note that the first term on the left side of (3.5) is also bounded by the expression on the right of (3.8) where the constants do not depend on n. Now using the triangular inequality and (3.5) one obtains (3.8). □

The steps toward the proof of Theorem 3.1 are similar to those in Theorem 14.5.2 in [6]. Therefore we will also need a result similar to Lemma 5.11.7 in [6].

Lemma 3.5. *With the definitions above we have:*

(i) $E((\tilde{X}_n - \varphi_n(\tau_{n-1}, \tilde{X}_{n-1}))^k/\mathscr{F}_{n-1})$

$$= -\int_{\tau_{n-1}}^{t}\cdots\int_{\tau_{n-1}}^{s_2} E[f_{\alpha*}^k(\varphi_{s_1}(\tau_{n-1}, \tilde{X}_{n-1}))/\mathscr{F}_{n-1}]ds_1\cdots ds_{\beta+1}, \tag{3.9}$$

with $\alpha^* = (0, \cdots 0,)$ *with* $l(\alpha^*) = \beta+1$.

(ii) $E(F_p(\tilde{X}_n - \tilde{X}_{n-1}) - F_p(\varphi_n(\tau_{n-1}, \tilde{X}_{n-1}) - \tilde{X}_{n-1})/\mathscr{F}_{n-1})$

$$= -\sum_{r=1}^{d}\frac{1}{r!}\sum_{k_1,\dots,k_r=1}^{d} E(\hat{F}_{k_1,\dots,k_r}(\tau_{n-1}, \tau_n)/\mathscr{F}_{n-1}) \tag{3.10}$$

where $$\hat{F}_{k_1,\dots,k_r}(\tau_{n-1}, \tau_n) = qF_{p'}(\tilde{X}_n - \tilde{X}_{n-1})\prod_{j=1}^{r}(\varphi_n(\tau_{n-1}, \tilde{X}_{n-1}) - \tilde{X}_n)^{k_j}. \tag{3.11}$$

For some $q\in\{1, \cdots, l(p)\}$ *and* $p'\in P_{l(p)-r}$ *which are functions of* p *and* r.

(iii) *In* (3.11) *for* $r=1$ *we have that for some* q *and* p' *as above,*

$$\hat{F}_k = \left\{\sum_{\alpha\in B(\Gamma_\beta)} I_\alpha[f_\alpha^k(\varphi_{s_1}(\tau_{n-1}, \tilde{X}_{n-1}))]_{n-1,n}\right\}\left\{q\sum_{j=0}^{m}\int_{\tau_{n-1}}^{\tau_n}\hat{b}_{p'}^j(s, \tilde{X}_{n-1})dW_s^j\right\}. \tag{3.12}$$

For the definition of $\hat{b}_{p'}^j$ *see* (5.11.12) *and* (5.11.13) *in* [6].
Furthermore there exists constants M, μ *that depend on* $b\in\{0, 1, \cdots\}$, $e\geq 2$, p *and* r *such that for* $r\geq 1$ *or* $l(p')\geq 2$, *we have*

$$\|E(\prod_{j=1}^{r}(\varphi_n(\tau_{n-1}, \tilde{X}_{n-1}) - \tilde{X}_n)^{k_j}F_{p'}(Z - \tilde{X}_{n-1})/\mathscr{F}_{n-1})\|_{b,e} \leq M(1+|x|)^\mu\delta^{\beta+1} \tag{3.13}$$

for $Z = \tilde{X}_n$, $\varphi_n(\tau_{n-1}, \tilde{X}_{n-1})$.

Note that $\widehat{b}^j_{p'}(s, \widetilde{X}_{n-1})$ always satisfies that $\tau_{n-1} \leq s < \tau_n$.

Proof. (3.9)-(3.12) have been proven in Lemma 5.11.7 of [6]. Therefore, we only prove (3.13). The proof is done by cases. We will do the proof in the case $b=1$, $r=1$, $l(p) \geq 2$, $Z=\widetilde{X}_n$. The proof in the other cases is similar. In this case by (3.12) we only need to consider for $t \leq \tau_{n-1}$, $l(\alpha)=\beta+1$, $\gamma=1, \cdots, d+m$:

$$D^\gamma_t E\left(I_\alpha[f^k_\alpha(\varphi_{s_1}(\tau_{n-1}, \widetilde{X}_n))]_{n-1,n} \int_{\tau_{n-1}}^{\tau_n} \widehat{b}^j_{p'}(s, \widetilde{X}_{n-1})\, dW^j_s / \mathcal{F}_{n-1}\right)$$
$$= E\Big(I_\alpha[D^\gamma_t\{f^k_\alpha(\varphi_{s_1}(\tau_{n-1}, \widetilde{X}_n))\}]_{n-1,n}$$
$$\int_{\tau_{n-1}}^{\tau_n} \widehat{b}^j_{p'}(s, \widetilde{X}_{n-1})\, dW^j_s + I_\alpha[f^k_\alpha(\varphi_{s_1}(\tau_{n-1}, \widetilde{X}_n))]_{n-1,n}$$
$$\int_{\tau_{n-1}}^{\tau_n} D^\gamma_t\{\widehat{b}^j_{p'}(s, \widetilde{X}_{n-1})\}\, dW^j_s / \mathcal{F}_{n-1}\Big).$$

Here we have used (3.1) and that $\widehat{b}^j_{p'}(s, \widetilde{X}_{n-1}) \in \mathbf{D}^\infty$ for all j, p, uniformly in s and n. Furthermore there exists M and r depending on p', e and b such that

$$\sup_s \|\widehat{b}^j_{p'}(s, \widetilde{X}_{n-1})\|_{b,e} \leq M(1+|x|)^r.$$

With this property and Lemma 5.7.2 in [6] the result follows.

The following Lemma will be the base for proof of Theorem 3.1 and Theorem 5.2. Its proof resembles the proof of 14.5.2 in [6]. From now on we will assume without loss of generality that $t \in \pi$.

Lemma 3.6. *There exists measurable functions $G(t, p, l, \omega, n, x, \delta)$, $\widetilde{G}(t, p, \omega, n, z, x, \delta)$, $\omega \in \Omega$, x, $z \in \mathbf{R}^d$, such that*

$$E(\phi_{e(\delta)}(\widetilde{X}_t - y) - \phi_{e(\delta)}(X_t - y)) = \sum_{n=1}^{n_t} \Bigg\{ \Bigg\{ \sum_{l=1}^{2\beta+1} \sum_{p \in \widetilde{P}_l} E\Big[\partial^p_y \phi_{g(\delta)}(\varphi_t(\tau_n, \widetilde{X}_{n-1})$$
$$+ h(\delta)\bar{W}_T - y)\, G(t, p, l, \omega, n, x, \delta)\Big]\Bigg\} + E(R_{n,\beta}(U_1) + R_{n,\beta}(U_2))\Bigg\} \tag{3.14}$$

where

$$ER_{n,\beta}(U) = \sum_{p \in \widetilde{P}_{2(\beta+1)}} E\Big(\int_0^1 \cdots \int_0^{s_2} \partial^p_y \phi_{g(\delta)}(\varphi_t(\tau_n, Z(\theta, U)) + h(\delta)\bar{W}_T - y)$$
$$\widetilde{G}(t, p, \omega, n, Z(\theta, U), x, \delta)\, d\theta\, ds_2 \cdots ds_{2(\beta+1)}\Big)$$

and $\widetilde{P}_l = \bigcup_{i=1}^{l} P_i$, $Z(\theta, U) = \widetilde{X}_{n-1} + \theta(U - \widetilde{X}_{n-1})$. Furthermore G and $\widetilde{G}$ satisfy for $U_1 = \widetilde{X}_n$, or $U_2 = \varphi_n(\tau_{n-1}, \widetilde{X}_{n-1})$ and e, $b \in \mathbf{N}$:

$$\|G(t, p, l, \cdot, n, x, \delta)\|_{b,e} + \sum_{i=1}^{2} \|\widetilde{G}(t, p, \cdot, n, Z(\theta, U_i), x, \delta)\|_{b,e} \leq M(1+|x|)^\mu \delta^{\beta+1}, \tag{3.15}$$

where M, and μ depend only on p, l, e and b.

Proof. Let $u(s, z) = E(\phi_{e(\delta)}(\varphi_t(s, z) - y))$. Then $u(t, z) = \phi_{e(\delta)}(z-y)$ and $L^0 u(s, z) = 0$ for $s<t$, $z \in \mathbf{R}^d$. Now we calculate, using Itô's formula and Taylor's expansion:

$$\begin{aligned}
E\phi_{e(\delta)}(\tilde{X}_t - y) - \phi_{e(\delta)}(X_t - y)) &= E\Big(\sum_{n=1}^{n_t}\{(u(\tau_n, \tilde{X}_n) - u(\tau_{n-1}, \tilde{X}_{n-1})) \\
&\quad - (u(\tau_n, \varphi_n(\tau_{n-1}, \tilde{X}_{n-1})) - u(\tau_{n-1}, \tilde{X}_{n-1}))\}\Big) \\
&= E\Big(\sum_{n=1}^{n_t}\{(u(\tau_n, \tilde{X}_n) - u(\tau_n, \tilde{X}_{n-1})) \\
&\quad - (u(\tau_n, \varphi_n(\tau_{n-1}, \tilde{X}_{n-1})) - u(\tau_n, \tilde{X}_{n-1}))\}\Big) \\
&= E\Big[\sum_{n=1}^{n_t}\Big(\sum_{l=1}^{2\beta+1}\frac{1}{l!}\Big(\sum_{p\in P_l}\Big\{\partial_z^p u(\tau_n, \tilde{X}_{n-1})(F_p(\tilde{X}_n - \tilde{X}_{n-1}) \\
&\quad - F_p(\varphi_n(\tau_{n-1}, \tilde{X}_{n-1}) - \tilde{X}_{n-1}))\Big\}\Big) + R_{n,\beta}(\tilde{X}_n) - R_{n,\beta}(\varphi_n(\tau_{n-1}, \tilde{X}_{n-1}))\Big)\Big],
\end{aligned} \tag{3.16}$$

where

$$R_{n,\beta}(Z) = \sum_{p\in P_{2(\beta+1)}} \int_0^1 \cdots \int_0^{s_2} \partial_z^p u(\tau_n, \tilde{X}_{n-1} + \theta(Z - \tilde{X}_{n-1}))\, d\theta ds_2 \cdots ds_{2(\beta+1)} F_p(Z - \tilde{X}_{n-1}). \tag{3.17}$$

The idea in the rest of the proof to follow is to consider each term in (3.16), apply the definition of u and define G and $\tilde{G}$. Then we use Lemma 3.5 to prove that the random variables G and $\tilde{G}$ obtained in this form satisfy (3.15). We divide the calculation in parts.

1. By the definition of u and Lemma 3.1, we have that there exists polynomials of degree less than $l(p)$, $P_{p'}$ such that

$$\begin{aligned}
\partial_z^p u(\tau_n, \tilde{X}_{n-1}) &= \partial_z^p E(\phi_{e(\delta)}(\varphi_t(\tau_n, \tilde{X}_{n-1}) - y)/\mathcal{F}_n) \\
&= \sum_{p' \hookrightarrow p} E((\partial_y^{p'}\phi_{g(\delta)})(\varphi_t(\tau_n, \tilde{X}_{n-1}) + h(\delta)\bar{W}_T - y) \\
&\qquad P_{p'}\Big(\frac{d^j\varphi_t(\tau_n, \tilde{X}_{n-1})}{dx^j}, j=1, \cdots, l(p)\Big)/\mathcal{F}_n),
\end{aligned} \tag{3.18}$$

where the set $p' \hookrightarrow p$, is composed by all possible indices p', $l(p') \in \{1, \cdots, l(p)\}$, selected from p in the same order as they appear in p. Obviously, $P_{p'}(\frac{d^j\varphi_t(\tau_n, \tilde{X}_{n-1})}{dx^j}, j=1, \cdots, l(p)) \in \mathbf{D}^\infty$ and furthermore for $\alpha \in P_b^{d+m}$

$$\sup_{u\in[0,T]^b} E\Big(\sup_{t\in[\tau_n, T]} |D_u^\alpha P_{p'}\Big(\frac{d^j\varphi_t(\tau_n, \tilde{X}_{n-1})}{dx^j}, j=1, \cdots, l(p)\Big)|^e\Big) \le M(1+|x|)^\mu, \tag{3.19}$$

where M and μ depend only on p, e and b. We will apply (3.19) to **2.** and **3.**

below. In the rest of this proof we will denote such a polynomial by $P_{p'}$ to simplify the notation.

Now, we can define G as appropriate sums of terms of the type

$$P_{p'}\Big(\frac{d^j\varphi_t(\tau_n, \widetilde{X}_{n-1})}{dx^j}, j=1, \cdots, l(p)\Big)E[(F_p(\widetilde{X}_n-\widetilde{X}_{n-1}) \\ -F_p(\varphi_n(\tau_{n-1}, \widetilde{X}_{n-1})-\widetilde{X}_{n-1}))/\mathscr{F}_{n-1}].$$

$\widetilde{G}$ is defined analogously. Now we will prove that they satisfy (3.15).

2. We now study the first term on the right of (3.16) in parts. Consider

$$A=E\{\partial_z^p u(\tau_n, \widetilde{X}_{n-1})E(F_p(\widetilde{X}_n-\widetilde{X}_{n-1})-F_p(\varphi_n(\tau_{n-1}, \widetilde{X}_{n-1})-\widetilde{X}_{n-1})/\mathscr{F}_{n-1})\},$$

for $n=1, \cdots, n_t$, $p\in P_l$, $l=1, \cdots, 2\beta+1$.

2.a In the case $l=1$, we have $p=(k)$, therefore by (3.9)

$$A=-\int_{\tau_{n-1}}^{\tau_n}\cdots\int_{\tau_{n-1}}^{s_2}E[\partial_z^k u(\tau_n, \widetilde{X}_{n-1})f_{\alpha*}^k(\varphi_{s_1}(\tau_{n-1}, \widetilde{X}_{n-1}))]ds_1\cdots ds_{\beta+1},$$

with $\alpha^*=(0, \cdots, 0)$ with $l(\alpha^*)=\beta+1$. Applying step **1.** to A, we have that the above expression can be written as finite sums of terms of the type

$$\int_{\tau_{n-1}}^{\tau_n}\cdots\int_{\tau_{n-1}}^{s_2}E[\partial_y^k\phi_{g(\delta)}(\varphi_t(\tau_n, \widetilde{X}_{n-1})+h(\delta)\bar{W}_T-y)P_{p'}f_{\alpha*}^k(\varphi_{s_1}(\tau_{n-1}, \widetilde{X}_{n-1}))] \\ ds_1\cdots ds_{\beta+1}.$$

Furthermore, using (3.1) and properties of the flow φ_t we have that there exists M and μ depending only on p, k, e and b such that

$$\sup_s\|P_{p'}f_{\alpha*}^k(\varphi_s(\tau_{n-1}, \widetilde{X}_{n-1}))\|_{b,e}\leq M(1+|x|)^{\mu}.$$

2.b In the case $l\geq 2$, $r=1$, we use (3.10) and (3.12) to obtain that the following is one of the summands of A (the other summands are considered in **2.c**)

$$E(\partial_z^p u(\tau_n, \widetilde{X}_{n-1})E(I_\alpha[f_\alpha^k(\varphi_{s_1}(\tau_{n-1}, \widetilde{X}_{n-1}))]_{n-1,n}\int_{\tau_{n-1}}^{\tau_n}\widehat{b}_{p''}^j(s, \widetilde{X}_{n-1})dW_s^j/\mathscr{F}_{n-1})), \tag{3.20}$$

for $k=1, \cdots, d$, $\alpha\in B(\Gamma_\beta)$, $p''\in P_{l-r}$ and $j=0, \cdots, m$. As in **2.a**, here we also find from **1.** and (3.13) that (3.20) can be written in terms of

$$E(\partial_y^{p'}\phi_{g(\delta)}(\varphi_t(\tau_n, \widetilde{X}_{n-1})+h(\delta)\bar{W}_T-y)P_{p'}E(I_\alpha[f_\alpha^k(\varphi_{s_1}(\tau_{n-1}, \widetilde{X}_{n-1}))]_{n-1,n} \\ \int_{\tau_{n-1}}^{\tau_n}\widehat{b}_{p''}^j(s, \widetilde{X}_{n-1})dW_s^j/\mathscr{F}_{n-1})),$$

and that there exists constants M and μ depending only on p, l, e and b such that that for $p''\in P_{l-r}$,

$$\|P_{p'}E(I_\alpha[f_\alpha^k(\varphi_{s_1}(\tau_{n-1},\tilde{X}_{n-1}))]_{n-1,n}\int_{\tau_{n-1}}^{\tau_n}\hat{b}^j_{p''}(s,\tilde{X}_{n-1})dW^j_s/\mathcal{F}_{n-1})\|_{b,e}$$
$$\leq M(1+|x|)^\mu\delta^{\beta+1}$$

2.c Now we consider the case, $r\geq 2$, that involves

$$E(\partial_z^p u(\tau_n,\tilde{X}_{n-1})E(\hat{F}_{k_1,\dots,k_r}(\tau_{n-1},\tau_n)/\mathcal{F}_{n-1})).$$

Considering (3.11), one obtains that for $p''\in P_{l-r}$

$$\|P_{p'}E(\prod_{j=1}^r(\varphi_n(\tau_{n-1},\tilde{X}_{n-1})-\tilde{X}_n)^{k_j}qF_{p''}(\tilde{X}_n-\tilde{X}_{n-1})/\mathcal{F}_{n-1})\|_{b,e}\leq M(1+|x|)^\mu\delta^{\beta+1}$$

where M and μ depend only on p, l, e and b. The above inequality follows from (3.13) and **1.**

By considering the calculations in **1.** and **2.** we have finish the proof of the first part of the Lemma. Now it only remains to consider the residues $R_{n,\beta}(\tilde{X}_n)$ and $R_{n,\beta}(\varphi_{\tau_n}(\tau_{n-1},\tilde{X}_{n-1}))$.

3. By (3.17) it is enough to prove that for $Z=\tilde{X}_n$, $\varphi_{\tau_n}(\tau_{n-1},\tilde{X}_{n-1})$, $p\in P_{2(\beta+1)}$

$$\|P_{p'}F_p(Z-\tilde{X}_{n-1})\|_{b,e}\leq M(1+|x|)^\mu\delta^{\beta+1},$$

for some constants M and μ depending only on p, l, b and e. As before, this inequality also follows from Lemma 3.5 and **1.**. Note that in this case $P_{p'}$ denotes a polynomial like in (3.18) with $\tilde{X}_{n-1}+\theta(Z-\tilde{X}_{n-1})$ instead of $\tilde{X}_{n-1}$.

Putting together all the steps from **1.** through **3.** the result follows.

Now, we are ready to give the proof of Theorem 3.1.

Proof of Theorem 3.1. First by Lemma 3.4 for $G=1$, we reduce our consideration to

$$|E[\partial_y^p\phi_{e(\delta)}(\tilde{X}_t-y)-\partial_y^p\phi_{e(\delta)}(X_t-y)]|.$$

Now we apply Lemma 3.6 to obtain an expression like (3.14). We analyze each term as we did in the proof of Lemma 3.6.

First we apply Lemma 3.2 to the remainders. We then find that these terms are bounded by an expression of the form

$$\sum_{n=1}^{n_t}\sum_{i=1}^{2}C(i)\sup_{\theta\in[0,1]}(\|(\det\sigma_{\varphi_t(\tau_n,Z(\theta,U_i))+h(\delta)\bar{V}_t})^{-1}\|_q^{\alpha_1}\|\tilde{G}(Z(\theta,U_i))\|_{b,e}\|\varphi_t(\tau_n,Z(\theta,U_i))\|_{\tilde{b},\tilde{e}}^{\alpha_2}),$$

for some positive constants α_1, α_2, q, b, e, $C(i)$. Then using Lemma 3.3, the estimate in (3.15) and classical methods to find stochastic derivatives of sde's we obtain that this term is also bounded by a term of the type $\frac{M(1+|x|)^\mu}{t^\kappa V_A(x)^c}\delta^\beta$.

The only term left in (3.14) can be handled similarly. In fact, this term is bounded by

$$\sum_{n=1}^{n_t}\sum_{l=1}^{2\beta+1}\sum_{p\in P_l} C(l)\|(\det \sigma_{\varphi_t(\tau_n,\tilde{X}_{n-1})+h(\delta)\tilde{W}_T})^{-1}\|_q^{\alpha_1}\|G\|_{b,e}\|\varphi_t(\tau_n,\tilde{X}_{n-1})\|_{b,e}^{\alpha_2},$$

for some positive constants α_1, α_2, q, b, e, $C(l)$ and G is the function that appears in Lemma 3.6 with $\|G\|_{b,e}\leq M(1+|x|)^{\mu}\delta^{\beta+1}$, where M and μ depend only on l, p, b, and e. Then by Lemma 3.3 the result follows. □

4. An extension of Theorem 3.1.

Before we start to develop the error of approximation in terms of the stepsize δ we will work out an extension of Theorem 3.1 that will be useful to understand a finite inductive argument to be used in the proof of Theorem 5.2.

For this, we expand the notation of F_p to the following case $F_p(f_{\alpha_i}(y))=\prod_{i=1}^{l}(f_{\alpha_i}(y))^{j_i}$, for $\alpha_1,\cdots,\alpha_l\in\mathcal{M}_m$ and $p=(j_1,\cdots,j_l)$. We also need to introduce some new notation. Let $a\in\{1,\cdots,N\}$ and define

$$P_a^{\pi_t}=\{(\tau_{j_a},\cdots,\tau_{j_0});j_a,\cdots,j_0\in\{0,\cdots,N\},\ \tau_{j_a},\cdots,\tau_{j_0}\in\pi,\ 0\leq\tau_{j_a}<\cdots<\tau_{j_0}\leq t\}.$$

For $\tau=(\tau_{j_a},\cdots,\tau_{j_0})\in P_a^{\pi_t}$ define the projection π_b for $b\leq a$ as $\pi_b(\tau)=(\tau_{j_a},\cdots,\tau_{j_{a-b}})$. Also let

$$\tau_t=\begin{cases}(\tau_{j_a},\cdots,\tau_{j_0},t) & ;\text{ if }\tau_{j_0}<t\\ \tau & ;\text{ if }\tau_{j_0}=t.\end{cases}$$

In the case $\tau_{j_0}<t$, we understand that $\tau_{j_{-1}}=t$.

For $s\leq\tau_{j_a-1}$ and $j\in\mathbf{N}$ define

$$\varphi_\tau(s,x)=\varphi_{j_0-1}(\tau_{j_1},\varphi_{j_1-1}(\tau_{j_2},\cdots(\tau_{j_a},\varphi_{j_a-1}(s,x))\cdots),$$
$$\frac{\partial^j}{\partial x^j}\varphi_\tau(s,x)=\Big(\frac{\partial^j}{\partial x^j}\varphi_{j_0-1}(\tau_{j_1},\cdot)\Big)(\varphi_{j_1-1}(\tau_{j_2},\cdots(\tau_{j_a},\varphi_{j_a-1}(s,x))\cdots).$$

In the particular case that we consider φ_{τ_t} or φ_{τ_v}, for $v\notin\pi$ we will then replace τ_{j_0-1} by t or v in the above definition respectively. That is, for example:

$$\varphi_{\tau_t}(s,x)=\varphi_t(\tau_{j_0-1},\varphi_{j_0-1}(\tau_{j_1},\cdots(\tau_{j_a},\varphi_{j_a-1}(s,x))\cdots).$$

Let $u_\delta^0(s,z)=E[\partial_y^p\phi_{e(\delta)}(\varphi_t(s,z)-y)]$. Then define inductively for sequences $\tau\in P_a^{\pi_t}$, $(p_a,\cdots,p_0)\in(\tilde{P}_\infty)^{a+1}$ and $\alpha_i^j\in\mathcal{M}$ for $i=1,\cdots,l(p_j);j=0,\cdots,a$.

$$u_\delta^{a+1}(s,z)=E[\partial_z^{p_a}u_\delta^a(\tau_{j_a},\varphi_{j_a-1}(s,z))F_{p_a}(f_{\alpha_i^a}(\varphi_{j_a-1}(s,z)))]$$

It is clear that u_δ^a as defined above also depends on $y\in\mathbf{R}^d$. To simplify the notation we have not written this dependence explicitly in the notation.

The purpose of this section is to give the analogous of Theorem 3.1 for u_δ^a for fixed a. Here we will only give the analogous Lemmas as in the previous sections. Their proofs are similar to the ones in the previous section. Therefore we will not give them here. We start with a representation for u^a. From now on we consider that $a\in\mathbf{N}$, $\tau\in P_a^{\pi_t}$, $(p_a,\cdots,p_0)\in(\tilde{P}_\infty)^{a+1}$, $\alpha_i^j\in\mathcal{M}$ for $i=1,\cdots,l(p_j)$; $j=0,\cdots,a$ are fixed unless explicitly stated otherwise.

Lemma 4.1. *There exists finite sets of multi-indices Q_k, $k=0,\cdots,a$ and polynomials $P_q(x)$ for $q\in Q_k$ such that for $s\leq\tau_{j_a-1}$ and $a\geq 0$,*

$$u_\delta^{a+1}(s,z)=\sum_{q_0,\dots,q_a,q\in Q_a}E\Big[\partial_y^q\phi_{g(\delta)}(\varphi_{\tau_t}(s,z)+h(\delta)\bar{W}_T-y)P_q\Big(\frac{\partial^j}{\partial x^j}\varphi_{\pi_k(\tau_t)}(s,z),\ k=1,\cdots,a+1;$$

$$j=0,\cdots,l(Q_a)\Big)\prod_{k=0}^{a}\partial_x^{q_k}\{F_{p_k}(f_{\alpha_i^k}(\varphi_{j_k-1}(\tau_{j_k+1},\cdot)))\}(\varphi_{\pi_{a-k-1}(\tau)}(s,z))\Big]\quad(4.1)$$

Here $l(Q_a)=\max\{l(q);\,q\in Q_a\}$.

Lemma 4.1 is the equivalent to the definition for u in Lemma 3.6 and it is proven by inductin.

Lemma 4.2. *The following inequality is satisfied for $\tau\in P_a^{\pi_t}$, $K>0$*

$$\sup_{\{(s,\lambda)\in[0,\tau_{j_a-1}]^2,\ s\geq\lambda\geq s-2\delta\}}\|(\det\sigma_{\tau_t(s,U_\lambda)+h(\delta)\bar{W}_T})^{-1}\|_{p'}\leq\frac{M(1+|x|)^\mu}{V_A(x)^c t^\kappa},$$

where U_λ is equal to either of the following processes: $\tilde{X}_\lambda$, $\tilde{X}_{j_{a+1}-1}+\theta(\tilde{X}_{j_{a+1}}-\tilde{X}_{j_{a+1}-1})$, $\tilde{X}_{j_{a+1}-1}+\theta(\varphi_{j_{a+1}}(\tau_{j_{a+1}-1},\tilde{X}_{j_{a+1}-1})-\tilde{X}_{j_{a+1}-1})$ or $\varphi_{j_{a+1}}(\tau_{j_{a+1}-1},\tilde{X}_{j_{a+1}-1})$, for $\tau_{j_{a+1}}\leq\lambda<\tau_{j_{a+1}+1}$ where $\theta\in[0,1]$ is a fix parameter and $j_{a+1}=0,\cdots,j_a-1$. Here M, μ, c and κ are positive constants that depend only on p' and a, and furthermore do not depend on θ,π or δ.

Lemma 4.3. *For $\gamma\in\mathbf{N}$ and $j_{a+1}\in\{0,\cdots,j_a-1\}$, we have*

$$E[\partial_z^{p_a}u_\delta^q(\tau_{j_a},\varphi_{j_a-1}(\tau_{j_{a+1}},X_{j_{a+1}-1}))F_{p_a}(f_{\alpha_i^a}(\varphi_{j_a-1}(\tau_{j_{a+1}},X_{j_{a+1}-1})))]-\sum_{q_0,\dots,q_a,q\in Q_a}\Big(\partial_y^q\Big\{q(t;x,y)$$

$$E\Big(P_q\Big(\frac{\partial^j}{\partial x^j}\varphi_{j_{a-k}}(\tau_{j_{a-k+1}},\cdot)\,(X_{j_{a-k+1}}),\ k=1,\cdots,a+1;\ j=0,\cdots,l(Q_a)\Big)$$

$$\prod_{k=0}^{a}\partial_x^{q_k}\{F_{p_k}(f_{\alpha_i^k}(\varphi_k(\cdot,\cdot)))\}(\tau_k,X_k)/X_t=y\Big)\Big\}\Big)=\sum_{1\leq j\leq\gamma}\{e(\delta)^{2j}A_j^q(\tau_t;x,y)\}+B_\gamma^q(\tau_t;x,y,\delta).$$

such that for $1\leq j\leq\gamma$, there exists functions A_j^q: $[0,T]^{a+1}\times(\mathbf{R}^d)^2\to\mathbf{R}$, B_γ^q: $[0,T]^{a+1}\times(\mathbf{R}^d)^2\times\mathbf{R}_+\to\mathbf{R}$ and positive constants M, μ, c, κ such that they depend only on $(p_0,\cdots,p_a)\in(\tilde{P}_\infty)^{a+1}$, $\alpha_i^j\in\mathcal{M}$ for $i=1,\cdots,l(p_j)$; $j=0,\cdots,a$ and $a\in\mathbf{N}$ with

$$\sup_y|A_j^q(\tau_t;x,y)|\leq\frac{M(1+|x|)^\mu}{V_A(x)^c t^\kappa},$$

$$\sup_y\left|\frac{B_\gamma^q(\tau_t;x,y,\delta)}{\delta^{\beta(\gamma+\frac{1}{2})}}\right|\leq\frac{M(1+|x|)^\mu}{V_A(x)^c t^\kappa}.$$

Furthermore the following estimate is satisfied

$$\sup_{y}\sum_{q_0,\dots,q_a,q\in Q_a}\Big|\partial_y^q(q(t;x,y)E\Big(P_q\Big(\frac{\partial^j}{\partial x^j}\varphi_{j_{a-k}}(\tau_{j_{a-k+1}},\cdot)(X_{j_{a-k+1}}),k=1,\cdots,a+1;j=0,\cdots,l(Q_a)\Big)$$
$$\prod_{k=0}^{a}\partial_x^{q_k}\{F_{p_k}(f_{\alpha_i^k}(\varphi_k(\cdot,\cdot)))\}(\tau_k,X_k)/X_t=y)\Big|\le\frac{M(1+|x|)^{\mu}}{V_A(x)^c t^{\kappa}} \tag{4.2}$$

Here the set Q_a is the set obtained from Lemma 4.1. Also note that in particular the constants M, μ, c and κ are independent of δ, τ_t and the partition π.

Lemma 4.4. *There exists measurable functions $G_a(\tau_{j_a},p_a,l,\omega,j_{a+1},x,\delta)$, $\tilde{G}_a(t,p_a,\omega,j_{a+1},z,x,\delta)$, $\omega\in\Omega$, $x,z\in\mathbf{R}^d$, such that*

$$E(u_\delta^q(\tau_{j_a},\tilde{X}_{j_a-1})-u_\delta^q(\tau_{j_a},X_{j_a-1}))=\sum_{j_{a+1}=1}^{j_a-1}\{\{\sum_{l=1}^{2\beta+1}\sum_{p_a\in\tilde{P}_l}E[\partial_z^{p_a}u_\delta^q(\tau_{j_a},\varphi_{j_a}(\tau_{j_{a+1}},\tilde{X}_{j_{a+1}-1}))$$
$$G_a(\tau_{j_a},p_a,l,j_{a+1},x,\delta)]\}+E(R_{j_{a+1},\beta}(U_1)+R_{j_{a+1},\beta}(U_2))\}$$

where

$$ER_{j_{a+1},\beta}(U)=\sum_{p_a\in\tilde{P}_{2(\beta+1)}}E\Big(\int_0^1\cdots\int_0^{s_2}\partial_z^{p_a}u_\delta^q(\tau_{j_a},Z(\theta,U)))$$
$$\tilde{G}_a(\tau_{j_a},p_a,w,j_{a+1},Z(\theta,U),x,\delta)d\theta ds_2\cdots ds_{2(\beta+1)}\Big)$$

and $\tilde{P}_l=\bigcup_{i=1}^{l}P_i$, $Z(\theta,U)=\tilde{X}_{j_{a+1}-1}+\theta(U-\tilde{X}_{j_{a+1}-1})$. Furthermore G and $\tilde{G}$ satisfy for $U_1=\tilde{X}_{j_{a+1}}$, or $U_2=\varphi_{j_{a+1}}(\tau_{j_{a+1}-1},\tilde{X}_{j_{a+1}-1})$ and fixed e, $b\in\mathbf{N}$:

$$\|G_a(\tau_{j_a},p_a,l,\cdot,j_{a+1},x,\delta)\|_{b,e}+\sum_{i=1}^{2}\|\tilde{G}_a(t,p_a,\cdot,j_{a+1},Z(\theta,U_i),x,\delta)\|_{b,e}\le M(1+|x|)^{\mu}\delta^{\beta+1},$$

where M, and μ depend only on p_a, l, e and b.

Theorem 4.1. *For $a\ge 0$ and $C_1>0$, there exists positive constants M, μ, c, κ such that they depend on $(p_0,\cdots,p_a)\in(\tilde{P}_\infty)^{a+1}$, $\alpha_i^j\in\mathcal{M}$ for $i=1,\cdots,l(p_j)$; $j=0,\cdots,a$ and $a\in\mathbf{N}$ with*

$$|E[\partial_z^{p_a}u_\delta^q(\tau_{j_a},\tilde{X}_{j_a-1})F_{p_a}(f_{\alpha_i^a}(\tilde{X}_{j_a-1}))-\partial_z^{p_a}u_\delta^q(\tau_{j_a},X_{j_a-1})F_{p_a}(f_{\alpha_i^a}(X_{j_a-1}))]|\le\frac{M(1+|x|)^{\mu}}{t^{\kappa}V_A(x)^c}\delta^{\beta}.$$

To prove the above theorem one has to go through a similar calculations as in Lemma 3.6 and Theorem 3.1. The dependence of the constants upon a will be unimportant for future developments as a will be always smaller than a fix positive integer.

5. Expansion of the error in powers of the step-size

In this section we will develop the approximation error in terms of δ^{γ} for

$\gamma=\beta, \cdots, \rho-1$, $\rho\geq\beta+1$, $\rho\in\mathbf{N}$, fixed. We will state the theorem first for $\rho=2\beta$ and after a preparatory Lemma we will give its proof. Then we will state the main theorem for the general case. This section resembles Section 14.6 in [6]. We assume that the partition π, is uniform, that is $\tau_n-\tau_{n-1}=\delta$ for all $n=1, \cdots, N$.

Theorem 5.1. *Assume the same conditions as in Theorem* 3.1. *Then there exists functions* $\psi_\gamma(t, x, y)$, $\gamma=\beta, \cdots, 2\beta-1$ *and constants* M, μ, κ, c *that depend only on* p. *In particular, the constants are independent of* t, x, y, δ *and the partition* π. *The functions* ψ_γ *are independent of* δ *and the partition* π. *The constants and the functions* ψ_γ *satisfy*

$$\sup_y|\partial_y^p E[\phi_{e(\delta)}(\tilde{X}_t-y)]-\partial_y^p q(t; x, y)-\sum_{\gamma=\beta}^{2\beta-1}\psi_\gamma(t, x, y)\delta^\gamma|\leq\frac{M(1+|x|)^\mu}{t^\kappa V_A(x)^c}\delta^{2\beta} \quad (5.1)$$

where $\psi_\gamma(t, x, y)$ *satisfies*

$$\sup_y|\psi_\gamma(t, x, y)|\leq\frac{M(1+|x|)^\mu}{t^\kappa V_A(x)^c} \quad (5.2)$$

Lemma 5.1. *For* $u(s, z)=E[\phi_{e(\delta)}(\varphi_t(s, z)-y)]$ *we have*

$$|\delta E[\partial_z^p u(\tau_n, X_{n-1})F_p(f_{\alpha_i}(X_{n-1}))]-\sum_{r=0}^{2\beta-1}\delta^r\int_{\tau_{n-1}}^{\tau_n}E(\Psi_r(s, X_s))ds|\leq\frac{M(1+|x|)^\mu}{t^\kappa V_A(x)^c}\delta^{2\beta+1}$$

for some measurable functions Ψ_r *and some positive constants* M, μ, κ *and* c. Ψ_r, *that depend only on* p *and the* α_i*'s. In particular they do not depend on the partition* π. *Furthermore* $E(\Psi_r(s, X_s))$ *satisfies*

$$\sup_{r,s}|E(\Psi_r(s, X_s))|\leq\frac{M(1+|x|)^\mu}{t^\kappa V_A(x)^c} \quad (5.3)$$

Proof. Here, we will only sketch the proof as it is very similar to the proof of Lemma 14.6.2 in [6]. In general consider $w(s, x)$ a smooth function with polynomial growth at infinity. Applying the Itô-Taylor expansion we have

$$E[\int_{\tau_{n-1}}^{\tau_n}w(s, X_s)ds-(\tau_n-\tau_{n-1})w(\tau_{n-1}, X_{n-1})]$$
$$=\sum_{r=1}^{2\beta-1}E((L^0)^r w(\tau_{n-1}, X_{n-1}))\frac{(\tau_n-\tau_{n-1})^{r+1}}{(r+1)!}+E(\int_{\tau_{n-1}}^{\tau_n}I_{\alpha*}[w_{\alpha*}(\cdot, X.)]_{\tau_{n-1},s}ds),$$

with $l(\alpha^*)=2\beta$. Now we use the above expansion repeatedly for each $(L^0)^r w$ in place of w, which proves the existence of constants C_r, B_r such that

$$E\Big[\int_{\tau_{n-1}}^{\tau_n}w(s, X_s)ds-(\tau_n-\tau_{n-1})w(\tau_{n-1}, X_{n-1})-\sum_{r=1}^{2\beta-1}\int_{\tau_{n-1}}^{\tau_n}C_r(L^0)^r w(s, X_s)ds\frac{(\tau_n-\tau_{n-1})^r}{(r+1)!}\Big]$$
$$=E\Big(\sum_{i=1}^{2\beta-1}\int_{\tau_{n-1}}^{\tau_n}B_i I_{\alpha_i^*}[((L^0)^{i-1}w)(\cdot, X.)]_{\tau_{n-1},s}ds\frac{(\tau_n-\tau_{n-1})^{i-1}}{i!}\Big). \quad (5.4)$$

Here $l(\alpha_i^*)=2\beta-i+1$. Now we replace $w(s,z)=(\partial_z^p u(s,z))F_p(f_{\alpha_i}(z))$ and prove that the term on the right side above is of order $\delta^{2\beta+1}$. In fact, note that as in the previous section, $E[((L^0)^{i-1}w)(s,X_s)]$ can be written in terms of expressions like in (4.1), because $L^0u(s,z)=0$ for $s<t$, $z\in\mathbf{R}^d$.

From here one only needs to apply Lemma 3.2, (2.2) and classical flow estimates to obtain bounds of the type $\frac{M(1+|x|)^\mu}{t^\kappa V_A(x)^c}$.

Now one uses the same argument as in (14.6.8) in [6] and Lemma 3.2 to conclude that

$$|\delta E(\partial_z^p u(\tau_n,X_{n-1})F_p(f_{\alpha_i}(X_{n-1})))-\sum_{r=0}^{2\beta-1}\delta^r\int_{\tau_{n-1}}^{\tau_n}E((\Phi_r w)(s,X_s))ds|$$
$$\leq\frac{M(1+|x|)^\mu}{t^\kappa V_A(x)^c}\delta^{2\beta+1},$$

where Φ is a differential operator in z. That is one proves using Taylor's expansion that

$$|\delta Ew(\tau_n,X_{n-1})-\delta Ew(\tau_{n-1},X_{n-1})-\delta\sum_{l=1}^{2\beta-1}E\Big(\Big(\frac{\partial}{\partial s}\Big)^l w(\tau_{n-1},X_{n-1})\Big)\frac{\delta^l}{l!}|$$
$$\leq\frac{M(1+|x|)^\mu}{t^\kappa V_A(x)^c}\delta^{2\beta+1},$$

and uses the fact that $L^0u(s,z)=0$ for $s<t$, $z\in\mathbf{R}^d$.

Next use the definition of w and u to obtain an expression for $E(\Phi_r w(s,X_s))$. This calculation, as in (4.1), shows that this expectation can be written in terms of

$$\sum_{q,q_1\in Q_{-1}}E\Big[\partial_y^q\phi_{e(\delta)}(X_t-y)P_q\Big(\frac{\partial^j}{\partial x^j}\varphi_t(s,\cdot)(X_s),j=0,\cdots,l(Q_{-1})\Big)\partial_y^{q_1}F_p(f_{\alpha_i}(\cdot))(X_s)\Big].\quad(5.5)$$

Here Q_{-1} is a set of multi-indices. Applying Lemma 3.4 to the above expression gives a δ-free function that can be used to define Ψ. The estimate (5.3) for such a function is obtained via (3.6)-(3.8). □

Proof of Theorem 5.1. Following the same rationale of Theorem 3.1. and Lemma 3.6 (which we will use repeatedly), we have from (3.16):

$$E[\partial_y^p\phi_{e(\delta)}(\tilde{X}_t-y)]-E[\partial_y^p\phi_{e(\delta)}(X_t-y)]=E[\sum_{n=1}^{n_t}\{(\sum_{l=1}^{4\beta+1}\frac{1}{l!}\sum_{p\in P_l}\partial_y^p u(\tau_n,\tilde{X}_{n-1})(F_p(\tilde{X}_n-\tilde{X}_{n-1})$$
$$-F_p(\varphi_n(\tau_{n-1},\tilde{X}_{n-1})-\tilde{X}_{n-1})))+R_{n,2\beta}(\tilde{X}_n)-R_{n,2\beta}(\varphi_n(\tau_{n-1},\tilde{X}_{n-1}))\}].\quad(5.6)$$

Let's define for $\tau_{n-1}\leq s<\tau_n$

$$\eta_s(x)=\sum_{\alpha\in\Gamma_{2\beta}}f_\alpha(x)I_{\alpha,\tau_{n-1},s}.$$

That is, η is the weak approximation of order 2β starting from x at time τ_{n-1}. As in Lemma 3.6, we will divide the study of (5.6) in cases.

a. The residual terms $R_{n,2\beta}(\tilde{X}_n)$ and $R_{n,2\beta}(\varphi_n(\tau_{n-1}, \tilde{X}_{n-1}))$ are of order $\delta^{4\beta+1}$ as proved in Lemma 3.6.

b. We will divide the differences in (5.6) of the type $F_p(\tilde{X}_n - \tilde{X}_{n-1}) - F_p(\varphi_{\tau_n}(\tau_{n-1}, \tilde{X}_{n-1}) - \tilde{X}_{n-1})$ in two. The first is

$$E[\partial_z^p u(\tau_n, \tilde{X}_{n-1}) E(F_p(\eta_n(\tilde{X}_{n-1}) - \tilde{X}_{n-1}) - F_p(\varphi_{\tau_n}(\tau_{n-1}, \tilde{X}_{n-1}) - \tilde{X}_{n-1}) / \mathcal{F}_{n-1})].$$

This term is also of order $\delta^{2\beta+1}$ as proved in Lemma 3.6, **2.**, because η_n is a weak approximation of order 2β.

c. Let's consider the term left from **b.**: Assume that $p = (j_1, \cdots, j_l)$

$$E(F_p(\tilde{X}_n - \tilde{X}_{n-1}) - F_p(\eta_n(\tilde{X}_{n-1}) - \tilde{X}_{n-1}) / \mathcal{F}_{n-1})$$
$$= \sum_{m=1}^{l} (-1)^m \sum_{k_i, p'} \sum_{\alpha_i \in \Gamma_{2\beta} - \Gamma_\beta} E(\prod_{i=1}^{m} (f_{\alpha_i}^{k_i}(\tilde{X}_{n-1}) I_{\alpha_i, \tau_{n-1}, \tau_n}) F_{p'}(\eta_n(\tilde{X}_{n-1}) - \tilde{X}_{n-1}) / \mathcal{F}_{n-1}).$$

where the first sum runs over all $k_1, \cdots, k_m$ taken without replacement from the set $\{j_1, \cdots, j_l\}$ and p' is the same index as p with the indices $k_1, \cdots, k_m$ removed. Define

$$g(\alpha, \alpha_1, \cdots, \alpha_{l-1}, \tau_n - \tau_{n-1}) = E(I_{\alpha, \tau_{n-1}, \tau_n} \prod_{i=1}^{l-1} I_{\alpha_i, \tau_{n-1}, \tau_n} / \mathcal{F}_{n-1}),$$

for $\alpha \in \Gamma_{2\beta} - \Gamma_\beta$ and $\alpha_i \in \Gamma_{2\beta} - \{v\}$ for $i = 1, \cdots, l-1$ and $\{k, k_1, \cdots, k_{l-1}\} = \{j_1, \cdots, j_l\}$. It is known (for example apply the same method as in Theorem 4.1 [2]) that g is a polynomial function in $\tau_n - \tau_{n-1} = \delta$ of the type $a_1\delta^{\beta+1} + \cdots + a_{l\beta}\delta^{(l+1)\beta}$.

Therefore we only need to consider terms of the type

$$E[\partial_z^p u(\tau_n, \tilde{X}_{n-1}) F_p(f_{\alpha_i}(\tilde{X}_{n-1}))] g(\alpha, \alpha_1, \cdots, \alpha_{l-1}, \tau_n - \tau_{n-1}). \qquad (5.7)$$

Then we apply Theorem 4.1 for $a = 0$ to prove that is enough to consider instead of (5.7):

$$E[\partial_z^p u(\tau_n, X_{n-1}) F_p(f_{\alpha_i}(X_{n-1}))] g(\alpha, \alpha_1, \cdots, \alpha_{l-1}, \tau_n - \tau_{n-1}). \qquad (5.8)$$

The proof finishes by applying Lemma 5.1 to (5.8). □

Now we will give the analogous of Theorem 5.1. in the case where an expansion of higher powers of δ is desired.

Theorem 5.2. *Assume the same conditions as in Theorem 3.1. Then there exists functions $\psi_\gamma(t, x, y)$, $\gamma = \beta, \cdots, \rho - 1$ and positive constants M, μ, κ, c that depend only on p. In particular the constants are independent of t, x, y, δ and the partition π. The functions ψ_γ are independent of δ and the partition π and they satisfy*

$$\sup_y |\partial_y^p E[\phi_{e(\delta)}(\tilde{X}_t - y)] - \partial_y^p q(t;, x, y) - \sum_{\gamma=\beta}^{\rho-1} \psi_\gamma(t, x, y) \delta^\gamma| \leq \frac{M(1+|x|)^\mu}{t^\kappa V_A(x)^c} \delta^\rho$$

where $\psi_T(t, x, y)$ *satisfies*

$$\sup_y |\psi_T(t, x, y)| \leq \frac{M(1+|x|)^\mu}{t^\kappa V_A(x)^c}.$$

In order to prove the above one also needs a generalization of Lemma 5.1. From now on we restrict a introduced in Section 4 to $a \leq [\frac{\rho}{\beta}] - 1$, where $[\cdot]$ denotes the greatest integer function.

Lemma 5.2.

$$|\sum_{\tau \in P_a^\pi} \delta^a E[\partial_z^{p_a} u_\delta^a(\tau_{j_a}, X_{j_a-1}) F_p(f_{\alpha_i^a}(X_{j_a-1}))] - \sum_{r=0}^{\rho-1} \delta^r \int_0^t \cdots$$
$$\int_0^{s_2} E(\Psi_r^a(s_1, \cdots, s_a, X_{s_1}, \cdots, X_{s_a})) ds_1 \cdots ds_a| \leq \frac{M(1+|x|)^\mu}{t^\kappa V_A(x)^c} \delta^{\rho+1},$$

for some measurable functions Ψ_r^a *and positive constants* M, μ, κ *and* c *that depend only on* p, $(p_0, \cdots, p_a) \in (\tilde{P}_\infty)^{a+1}$, $a \in \mathbf{N}$ *and the* $\alpha_i^j \in \mathcal{M}$ *for* $i=1, \cdots, l(p_j)$; $j=0, \cdots, a$. *In particular the constants do not depend on the partition* π, δ, t, x *or* y *and the functions* Ψ_r^a *do not depend on the partition* π *or* δ. *Furthermore* $E(\Psi_r^a(s_1, \cdots, s_a, X_{s_1}, \cdots, X_{s_a}))$ *satisfies*

$$\sup_{r, s_1, \ldots, s_a} |E(\Psi_r^a(s_1, \cdots, s_a, X_{s_1}, \cdots, X_{s_a}))| \leq \frac{M(1+|x|)^\mu}{t^\kappa V_A(x)^c}.$$

Note that all the n's that appeared in the Lemma 5.1 and Theorem 5.1 and their proofs become j_0 in the present situation. Similary u becomes u^0. The general idea of the proofs of Lemma 5.2 and Theorem 5.2 is to repeat the same argument a times. The iteration of the argument in the proof of Theorem 5.2 gives the indices $j_1, \cdots, j_a$.

In fact, the proof of Lemma 5.2 is obtained by performing finite induction on a. That is, one repeats the steps in Lemma 5.1 a times in a conditional form (also replacing 2β by ρ).

Sketch of the proof of Theorem 5.2. The general case is obtained by iteration of the argument in the proof of Theorem 5.1, as in [7]. We have already proved the result for $\rho \leq 2\beta$.

Consider in general the difference (for definitions see Section 4)

$$E[u_\delta^{a-1}(\tau_{j_{a-1}}, \tilde{X}_{j_{a-1}-1}) - u_\delta^{a-1}(\tau_{j_{a-1}}, X_{j_{a-1}-1})] \tag{5.9}$$

The proof of Theorem 5.1 can be used up to (5.7). Then we get that instead of (5.7) and (5.8), we have to compare

$$E[\partial_z^{p_a} u_\delta^a(\tau_{j_a}, \tilde{X}_{j_a-1}) F_{p_a}(f_{\alpha_i^a}(\tilde{X}_{j_a-1}))] - E[\partial_z^{p_a} u_\delta^a(\tau_{j_a}, X_{j_a-1}) F_{p_a}(f_{\alpha_i^a}(X_{j_a-1}))].$$

Note that p_a and α_i^a determined above do not depend on $j_0, \cdots, j_a$. By Theorem 4.1, this difference converges to 0 at a rate $O(\delta^\beta)$. This gives an expansion of the difference (5.9) in terms of δ^j, $j=\beta, \cdots, 2\beta$.

Now consider $\rho \leq 3\beta$. The proof of Theorem 5.1 can be used up to (5.7) replacing 2β by ρ. By the previous argument for $a=0$, the difference between (5.7) and (5.8) has a expansion in terms of δ^j, $j=\beta, \cdots, 2\beta$. Therefore the result follows. The rest of the proof follows by finite induction.

Note that the function ψ_γ can be written explicitly following the above proof carefully. In fact for $\gamma=1$ the function ψ_1 is explicitly stated in [1].

With the same methodology used here one can actually achieve other generalizations. For example, in the case that the heat kernel to approximate is associated with the operator

$$L'f(s, x) = L^0 f(s, x) + V(x) f(s, x)$$

for $V: \mathbf{R}^d \to \mathbf{R}$ a smooth bounded function with bounded derivatives, an extension of Theorem 5.2 can be proved. Let's first introduce a middle step in this generalization. The following result extends Theorem 14.6.1 in [6].

Corollary 5.1. *Let f be a smooth function with polynomial growth at infinity. Then there exists a positive constant C and functions ψ_γ, $\gamma=\beta+1, \cdots, \rho-1$ independent of δ, such that for $\rho \geq \beta$*

$$|E[\exp(\int_0^t \tilde{V}_s(\tilde{X}_{\eta(s)})ds) f(\tilde{X}_t)] - \exp(\int_0^t V(X_s)ds) f(X_t) - \sum_{\gamma=\beta}^{\rho-1} \psi_\gamma(t, x)\delta^\gamma| \leq C\delta^\rho.$$

Here the approximation $\tilde{V}_s(x)$ is defined as

$$\tilde{V}_t(x) = \sum_{\alpha \in \Gamma_{\beta-1}} V_\alpha(x) I_{\alpha, \tau_n, t} \qquad \tau_n < t \leq \tau_{n+1}.$$

With a slight modification of the technique shown in this article one can also prove the following result.

Theorem 5.3. *Assume the same conditions as in Theorem 3.1. Then there exists functions $\psi_\gamma(t, x, y)$ $\gamma=\beta, \cdots, \rho-1$ and positive constants M, μ, κ, c depending only on p satisfying the equation below. In particular the constants are independent of t, x, y, δ and the partition π. The functions ψ_γ are independent of δ and the partition π.*

$$\sup_y |\partial_y^p E[\exp(\int_0^t \tilde{V}_s(\tilde{X}_{\eta(s)})ds)\, \phi_{e(\delta)}(\tilde{X}_t - y)] - \partial_y^p \bar{q}(t; x, y) - \sum_{\gamma=\beta}^{\rho-1} \psi_\gamma(t, x, y)\delta^\gamma|$$
$$\leq \frac{M(1+|x|)^\mu}{t^\kappa V_A(x)^c}\delta^\rho,$$

where $\psi_\gamma(t, x, y)$ satisfies

$$\sup_y |\psi_\gamma(t, x, y)| \leq \frac{M(1+|x|)^\mu}{t^\kappa V_A(x)^c}.$$

where $\bar{q}(t, x, y) = E(\exp(\int_0^t V(X_s)ds) / X_t = y) q(t; x, y)$.

The introduction of the term $\exp(\int_0^t V(X_s)ds)$ does not bring any major complication to the methodology applied here. The only point that one has to be careful about is that when one is working with the formula (3.16) and its subsequent analysis one finds terms of the following type

$$\partial_z^p u(\tau_n, \tilde{X}_n) \int_{\tau_n}^{\tau_{n+1}} (\tilde{V}_s(\tilde{X}_n) - V(\varphi_s(\tau_n, \tilde{X}_n)))ds.$$

where $u(s, z) = E(\exp(\int_s^t V(\varphi_v(s, z))dv) \phi_{e(\delta)}(\varphi_t(s, z) - y))$.

Acknowledgements. The author wishes to thank Prof. D. Nualart, Prof. S. Watanabe and Prof. I. Shigekawa for advice and encouragement.

UNIVERSITAT POMPEU FABRA
DEPARTMENT OF ECONOMICS
RAMON TRIAS FARGAS 25-27
08005 BARCELONA
SPAIN

References

[1] V. Bally and D. Talay, The law of the Euler scheme for stochastic differential equations: II. Convergence rate of the density, to appear in Monte Carlo Methods and Applications (1996).
[2] J. G. Gaines, The algebra of iterated stochastic integrals, Stochastics and Stochastics Reports, **49** (1994), 169-179.
[3] N. Ikeda and S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland, 1989.
[4] Y. Hu and S. Watanabe, Donsker delta functions and approximations of heat kernels by the time discretization method, preprint (1995).
[5] S. Kusuoka and D. Stroock, Applications of the Malliavin calculus II, J. Fac. Sci. Univ. Tokyo Sect IA Math., **32** (1985), 1-76.
[6] P. E. Kloeden and E. Platen, Numerical Solution of Stochastic Differential Equations, Applications of Mathematics **23**, Springer, New York, 1992.
[7] P. E. Kloeden, E. Platen and N. Hoffmann, Extrapolation methods for the weak approximation of Itô-difussions, SIAM J. Numer. Anal., **32** (1995), 1519-1534.
[8] R. Léandre, Minoration en temps petit de la densité d'une diffusion dégénérée, Journal of Functional Analysis, **74** (1987), 399-414.
[9] D. Nualart, The Malliavin Calculus and Related Topics, Springer Verlag, 1995.