ESAIM: M2AN 55 (2021) S323–S367
https://doi.org/10.1051/m2an/2020043

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

OPERATOR SPLITTING AROUND EULER–MARUYAMA SCHEME AND HIGH ORDER DISCRETIZATION OF HEAT KERNELS

Yuga Iguchi[1] and Toshihiro Yamada[2,*]

Abstract. This paper proposes a general higher order operator splitting scheme for diffusion semigroups using the Baker–Campbell–Hausdorff type commutator expansion of non-commutative algebra and the Malliavin calculus. An accurate discretization method for the fundamental solution of heat equations or the heat kernel is introduced with a new computational algorithm which will be useful for the inference for diffusion processes. The approximation is regarded as the splitting around the Euler–Maruyama scheme for the density. Numerical examples for diffusion processes are shown to validate the proposed scheme.

Mathematics Subject Classification. 58J35, 60H07, 60H30, 60H35, 60J60, 65C05.

Received July 1, 2019. Accepted June 25, 2020.

1. Introduction

Approximating or estimating heat kernels is an important theme in mathematics since it naturally appears in many problems related to the topics on diffusions and partial differential equations. Because heat kernel is given as the density of the solution of stochastic differential equation, many kernel estimations rely on probabilistic methods. While there has been considerable studies on kernel estimation, high order discretization for heat kernels and its validity are not obtained at present.

This paper shows a general high order discretization algorithm for heat kernels with its theoretical foundation. Before illustrating the sketch of the discretization method, we briefly review the standard discretization scheme. Maruyama [17] proposed a discretization method for Itô's stochastic differential equations (SDEs) which is nowadays called the *Euler–Maruyama* scheme. The method is widely used in many fields due to its versatility and applicability. Let $\{X(t,x)\}_{t\geq 0}$ be the solution of an Itô SDE

$$dX(t,x) = b(X(t,x))\,\mathrm{d}t + \sum_{i=1}^{d} \sigma_i(X(t,x))dB_t^i, \quad X(0,x) = x \in \mathbb{R}^N, \tag{1.1}$$

Keywords and phrases. Heat kernel, high order discretization, operator splitting, Baker–Campbell–Hausdorff formula, Malliavin calculus.

[1] MUFG Bank, Ltd., 2-7-1 Marunouchi, Chiyoda, Tokyo, Japan.

[2] Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo, Japan.

*Corresponding author: toshihiro.yamada@r.hit-u.ac.jp

with functions b, σ_i, $i=1,\dots,d$, and consider the problem of computing $P_Tf(x)=E[f(X(T,x))]$ for a function f where $\{P_t\}_t$ is the semigroup of linear operators given by $(P_tf)(x)=E[f(X(t,x))]$. It is known that one has

$$\left|E[f(X(T,x))]-E\left[f\left(\bar{X}^{\mathrm{EM},(n)}(T,x)\right)\right]\right|=O(1/n), \tag{1.2}$$

where $\bar{X}^{\mathrm{EM},(n)}(T,x)$ is the n-step Euler–Maruyama scheme with the time interval T/n given by

$$\begin{aligned}\bar{X}^{\mathrm{EM},(n)}((k+1)T/n,x)&=\bar{X}^{\mathrm{EM},(n)}(kT/n,x)+b\left(\bar{X}^{\mathrm{EM},(n)}(kT/n,x)\right)T/n\\&\quad+\sum_{i=1}^d\sigma_i\left(\bar{X}^{\mathrm{EM},(n)}(kT/n,x)\right)\left(B^i_{(k+1)T/n}-B^i_{kT/n}\right).\end{aligned} \tag{1.3}$$

In the operator splitting perspective, (1.2) is written in the form

$$\left|P_Tf(x)-\left(Q^{\mathrm{EM}}_{T/n}\right)^nf(x)\right|=O(1/n), \tag{1.4}$$

where $Q^{\mathrm{EM}}_tf(x)=E[f(\bar{X}^{\mathrm{EM}}(t,x))]$ with $\bar{X}^{\mathrm{EM}}(t,x)=x+b(x)t+\sum_{i=1}^d\sigma_i(x)B^i_t$. Even if the test function f is non-smooth, (1.2) or (1.4) still gives first order discretization under the sufficient smoothness condition for the coefficients b, σ_i, $i=1,\dots,d$ with an appropriate ellipticity. Furthermore, the density $y\mapsto p_T(x,y)$ of $X(T,x)$ is also approximated by the density $y\mapsto p_T^{\bar{X}^{\mathrm{EM},(n)}}(x,y)$ of the Euler–Maruyama scheme $\bar{X}^{\mathrm{EM},(n)}(T,x)$ as

$$\left|p_T^X(x,y)-p_T^{\bar{X}^{\mathrm{EM},(n)}}(x,y)\right|\le\frac{1}{n}\frac{c_1}{T^\alpha}e^{-c_2|x-y|^2/T} \tag{1.5}$$

for some $c_1,c_2>0$ and $\alpha\ge N$ under the sufficient smoothness condition for the coefficients b, σ_i, $i=1,\dots,d$ and the uniformly elliptic condition for σ_i, $i=1,\dots,d$; see Bally and Talay [4]. See also [9, 10, 13] for the Euler–Maruyama scheme for the density for instance. The efficient computation scheme is obtained in Pedersen [19] as follows:

$$p_T^X(x,y)=E\left[p_{T/n}^{\bar{X}^{\mathrm{EM},(n)}}\left(\bar{X}^{\mathrm{EM},(n)}((n-1)T/n,x),y\right)\right]+O(1/n), \tag{1.6}$$

where $y\mapsto p_{T/n}^{\bar{X}^{\mathrm{EM},(n)}}(x,y)$ is the Gaussian density of the one-step Euler–Maruyama scheme. Here, the $(n-1)$-step Euler–Maruyama scheme $\bar{X}^{\mathrm{EM},(n)}((n-1)T/n,x)$ is used in the algorithm (1.6), and then the scheme (1.6) is implemented by the Monte-Carlo simulation, which enables us to treat statistical inference of diffusion processes.

The paper shows a new discretization method of the heat kernel as an extension of the Euler–Maruyama scheme in Bally and Talay [4] and the algorithm of Pedersen [19]. We firstly show a general operator splitting method for diffusion semigroups as expectations of Itô SDEs. The approximation is obtained through a generator expansion method around the Euler–Maruyama semigroup and the Baker–Campbell–Hausdorff type commutator calculation of non-commutative algebra. In particular, we introduce an optimal truncation of the semigroup expansion in order to give a low cost numerical computation. For the global error estimate of the discretization, the Kusuoka–Stroock theory on Malliavin calculus plays an important role. The high order splitting method for weak approximation of Itô SDEs of the order $O(1/n^m)$ is given by

$$|P_Tf(x)-(Q^{(m)}_{T/n})^nf(x)|=O(1/n^m) \tag{1.7}$$

with operators $Q^{(m)}_t$, $t>0$ constructed through the Baker–Campbell–Hausdorff expansion, which has the form $Q^{(m)}_tf(x)=E[f(\bar{X}^{\mathrm{EM}}(t,x))\{1+\pi^{(m),x}_t(B_t)\}]$ with the one-step Euler–Maruyama scheme $\bar{X}^{\mathrm{EM}}(t,x)$ and a functional of polynomials of Brownian motions $\pi^{(m),x}_t(B_t)$. Here, the optimal truncation is introduced for the Baker–Campbell–Hausdorff expansion using Malliavin calculus in order to attain a high order discretization with minimum computational effort. Consequently, for instance, we get a simple second order discretization for Itô SDEs

$$P_T f(x) = \left(Q^{(2)}_{T/n}\right)^n f(x) + O\left(1/n^2\right) \tag{1.8}$$

with a simple local operator $Q^{(2)}_t$ given by

$$Q^{(2)}_t f(x) = \left\{1 + \frac{t^2}{2}\sum_{i=1}^{2}[\mathscr{L}^z_0, \mathscr{L}^z_i]\right\} P^{0,z}_t f(x)|_{z=x} = E\left[f(\bar{X}^{\mathrm{EM}}(t,x))\left\{1 + \pi^{(2),x}_t(B_t)\right\}\right], \tag{1.9}$$

where $P^{0,z}_t f(x) = E[f(x + b(z)t + \sum_{i=1}^d \sigma_i(z)B^i_t)]$ and $[\mathscr{L}^z_0, \mathscr{L}^z_i]$ $(i = 1, 2)$ is the commutator of two differentiation operators $\mathscr{L}^z_0$ and $\mathscr{L}^z_i$ appearing in the Taylor expansion $\mathscr{L} \approx \mathscr{L}^z_0 + \sum_{i=1}^2 \mathscr{L}^z_i$. Then, in the second order discretization, the terms $[\mathscr{L}^z_0, \mathscr{L}^z_i]P^{0,z}_t f(x)|_{z=x}$, $i = 1, 2$, work as the correction to the Euler–Maruyama term $P^{0,z}_t f(x)|_{z=x} = E[f(\bar{X}^{\mathrm{EM}}(t,x))]$, and these can be obtained through an easy commutator computation of the Baker–Campbell–Hausdorff formula. The general operator splitting (1.7) can be represented as

$$\Bigg|E[f(X(T,x))] - E\Bigg[f(\bar{X}^{\mathrm{EM},(n)}(T,x)) \times \prod_{i=1}^{n}\left\{1 + \pi^{(m),\bar{X}^{\mathrm{EM},(n)}((i-1)T/n,x)}_{T/n}(B_{iT/n} - B_{(i-1)T/n})\right\}\Bigg]\Bigg| = O(1/n^m). \tag{1.10}$$

using the n-step Euler–Maruyama scheme $\bar{X}^{\mathrm{EM},(n)}(T,x)$. Since the discretization scheme is constructed by

$$\left(Q^{(m)}_{T/n}\right)^n f(x) = E\left[f\left(\bar{X}^{\mathrm{EM},(n)}(T,x)\right)\prod_{i=1}^{n}\left\{1 + \pi^{(m),\bar{X}^{\mathrm{EM},(n)}((i-1)T/n,x)}_{T/n}\left(B_{iT/n} - B_{(i-1)T/n}\right)\right\}\right]$$

and the term $E[f(\bar{X}^{\mathrm{EM},(n)}(T,x))]$ in (1.10) is the Euler–Maruyama scheme, the approximation is regarded as the operator splitting around the Euler–Maruyama scheme. The result (1.10) holds even if the function f is only bounded and measurable.

As the second main result, we next introduce a higher order discretization scheme of the heat kernel $p^X_T(x,y)$ as follows:

$$\Bigg|p^X_T(x,y) - E\Bigg[p^{\bar{X},(m)}_{T/n}\left(\bar{X}^{\mathrm{EM},(n)}((n-1)T/n,x),y\right) \times \prod_{i=1}^{n-1}\left\{1 + \pi^{(m),\bar{X}^{\mathrm{EM},(n)}((i-1)T/n,x)}_{T/n}(B_{iT/n} - B_{(i-1)T/n})\right\}\Bigg]\Bigg| \le \frac{1}{n^m}\frac{c_1}{T^\alpha}e^{-c_2|x-y|^2/T}, \tag{1.11}$$

where $c_1, c_2 > 0$ and $\alpha \ge N$ are some constants. Here, the function $y \mapsto p^{\bar{X},(m)}_t(x,y)$ is a small time approximation of $p^X_t(x,\cdot)$, and we note that the $(n-1)$-th (not n-th) product of polynomials of Brownian motions is used. The kernel discretization is a natural extension of (1.6) in [19] since the function $p^{\bar{X},(m)}_t(x,\cdot)$ has the form $p^{\bar{X},(m)}_t(x,\cdot) = p^{\bar{X}^{\mathrm{EM},(n)}}_t(x,\cdot)\{1 + \vartheta_t(x,\cdot)\}$ with a polynomial function $\vartheta_t(x,\cdot)$ and the method in (1.11) can be written as

$$p^X_T(x,y) = E\left[p^{\bar{X},(m)}_{T/n}\left(\bar{X}^{\mathrm{EM},(n)}((n-1)T/n,x),y\right)\left(1 + M^{n-1}_{(m)}(T,x)\right)\right] + O\left(\frac{1}{n^m}\right),$$

with $1 + M^{n-1}_{(m)}(T,x) = \prod_{i=1}^{n-1}\{1 + \pi^{(m),\bar{X}^{\mathrm{EM},(n)}((i-1)T/n,x)}_{T/n}(B_{iT/n} - B_{(i-1)T/n})\}$. Also the approximation can be a generalization of [23]. The scheme is simply implemented by the Quasi-Monte-Carlo method. We provide numerical examples for the scheme in order to confirm the validity.

The paper is organized as follows. Section 2 introduces the higher order splitting method for diffusion semigroups around the Euler–Maruyama scheme. Then, in Section 3, we show a new discretization algorithm for the heat kernel. Numerical examples of the proposed method are given in Section 4 with the comparison with the Euler–Maruyama scheme. Section 5 concludes the method. Appendix is devoted to the proofs of mathematical results.

2. Splitting method

For $N, n \in \mathbb{N}$, let $C_b^\infty(\mathbb{R}^N, \mathbb{R}^n)$ be the space of all infinitely continuously differentiable functions $f : \mathbb{R}^N \to \mathbb{R}^n$ such that f and all of its partial derivatives at any order are bounded. We write $C_b^\infty(\mathbb{R}^N)$ for $C_b^\infty(\mathbb{R}^N, \mathbb{R})$. For any $f \in C_b^\infty(\mathbb{R}^N)$, we define

$$\|f\|_\infty := \sup_{x\in\mathbb{R}^N} |f(x)|, \qquad \|\nabla^k f\|_\infty := \max_{j_1,\dots,j_k\in\{1,\dots,N\}} \left\| \frac{\partial^k f}{\partial x_{j_1}\dots\partial x_{j_k}} \right\|_\infty . \tag{2.1}$$

Let $\Omega = C_0([0,\infty), \mathbb{R}^d) = \{w : [0,\infty) \to \mathbb{R}^d;\ w(0) = 0,\ w \text{ is continuous}\}$, $\mathcal{F}$ be the Borel field over Ω and $\mathbb{P}$ be the Wiener measure. For $p \in [1,\infty)$, the L^p-space of real valued Wiener functionals is denoted by $L^p(\Omega)$; that is, $L^p(\Omega)$ is a real Banach space of all $\mathbb{P}$-measurable functionals $F : \Omega \to \mathbb{R}$ such that $\|F\|_p = E[|F|^p]^{1/p} < \infty$ with the identification $F = G$ if and only if $F(w) = G(w)$, a.s.

We will use the language of Malliavin calculus. See [12, 18] for details. Let $H = L^2([0,\infty), \mathbb{R}^d)$ with the inner product $\langle\cdot,\cdot\rangle_H$ and $B(h)$ be the Wiener integral $B(h) = \sum_{j=1}^d \int_0^\infty h^j(s)\,\mathrm{d}B_s^j$ for $h \in H$. Let $\mathcal{S}(\Omega)$ denote the class of smooth random variables of the form $F = f(B(h_1),\dots,B(h_n))$ where $f \in C_b^\infty(\mathbb{R}^n)$, $h_1,\dots,h_n \in H$, $n \geq 1$. For $F \in \mathcal{S}(\Omega)$, we define the derivative DF as the H-valued random variable $DF = \sum_{i=1}^n \partial_i f(B(h_1),\dots,B(h_n))h_i$ and $D_{j,s}F = \sum_{i=1}^n \partial_i f(B(h_1),\dots,B(h_n))h_i^j(s)$, $j = 1,\dots,d$, $s \geq 0$. For $F \in \mathcal{S}(\Omega)$, we set D^jF, $j \in \mathbb{N}$, as the $H^{\otimes j}$-valued random variable obtained by iterating j-times the operator D. Then D^j is a closable operator from $L^p(\Omega)$ into $L^p(\Omega, H^{\otimes j})$ for any $p \in [1,\infty)$. For $k \in \mathbb{N}$, $p \in [1,\infty)$, we define $\|F\|_{k,p}^p = E[|F|^p] + \sum_{j=1}^k E[\|D^jF\|_{H^{\otimes j}}^p]$, $F \in \mathcal{S}(\Omega)$. Then the space $\mathbb{D}^{k,p}$ is defined as the completion of $\mathcal{S}(\Omega)$ with respect to the norm $\|\cdot\|_{k,p}$. Moreover, let $\mathbb{D}^\infty$ be the space of smooth Wiener functionals in the sense of Malliavin $\mathbb{D}^\infty = \cap_{p\geq1}\cap_{k\in\mathbb{N}}\mathbb{D}^{k,p}$. Let δ be an unbounded operator from $L^2(\Omega, H)$ into $L^2(\Omega)$ such that the domain of δ, denoted by $\mathrm{Dom}(\delta)$, is the set of H-valued square integrable random variables u such that $|E[\langle DF, u\rangle_H]| \leq C\|F\|_2$, for all $F \in \mathbb{D}^{1,2}$ where C is some constant depending on u, and if $u \in \mathrm{Dom}(\delta)$, $\delta(u)$ is characterized by

$$E[\langle DF, u\rangle_H] = E[F\delta(u)] \tag{2.2}$$

for all $F \in \mathbb{D}^{1,2}$. $\delta(u)$ is called the Skorohod integral of the process u. When $u \in \mathrm{Dom}(\delta)$ has the form $u = Gh$ with $G \in \mathbb{D}^{1,2}$ and $h \in H$, the Skorohod integral is given by

$$\delta(Gh) = G\delta(h) - \langle DG, h\rangle_H. \tag{2.3}$$

Let $F = (F_1,\dots,F_N)$ be a Wiener functional such that $F_i \in \mathbb{D}^\infty$, $i = 1,\dots,N$, and the Malliavin covariance matrix $\sigma^F = (\langle DF_i, DF_j\rangle_H)_{1\leq i,j\leq N}$ is invertible a.s. and satisfies $\|(\det\sigma^F)^{-1}\|_p < \infty$, $p \geq 1$. For such a nondegenerate Wiener functional F, we have the integration by parts, that is, for all $g \in C_b^\infty(\mathbb{R}^N)$, $G \in \mathbb{D}^\infty$ and multi-index $\alpha = (\alpha_1,\dots,\alpha_k) \in \{1,\dots,N\}^k$, $k \in \mathbb{N}$, there exists $H_\alpha(F,G) \in \mathbb{D}^\infty$ such that

$$E[(\partial^\alpha g)(F)G] = E[g(F)H_\alpha(F,G)]. \tag{2.4}$$

In particular, $H_\alpha(F,G)$ is given by $H_\alpha(F,G) = H_{(\alpha_k)}(F, H_{(\alpha_1,\dots,\alpha_{k-1})}(F,G))$ with

$$H_{(i)}(F,G) = \delta\left(\sum_{j=1}^N G(\sigma^F)^{-1}_{i,j}DF_j\right), \quad i = 1,\dots,N. \tag{2.5}$$

Let us consider the solution $X = \{X(t,x)\}_{t\geq 0}$, $x \in \mathbb{R}^N$ of the following Itô's stochastic differential equation driven by a d-dimensional Brownian motion $B = \{B_t\}_{t\geq 0}$:

$$X(t,x) = x + \int_0^t b(X(s,x))\,\mathrm{d}s + \sum_{i=1}^d \int_0^t \sigma_i(X(s,x))\,\mathrm{d}B_s^i, \tag{2.6}$$

where $b \in C_b^\infty(\mathbb{R}^N,\mathbb{R}^N)$, $\sigma_i \in C_b^\infty(\mathbb{R}^N,\mathbb{R}^N)$, $i = 1,\ldots,d$. We assume the uniformly elliptic condition for the matrix $\sigma\sigma'(\cdot)$, where σ' is the transposition of σ . Let $\mathscr{L}$ be the generator given by

$$\mathscr{L}\varphi(x) = \sum_{i=1}^N b^i(x)\frac{\partial}{\partial x_i}\varphi(x) + \frac{1}{2}\sum_{k=1}^d\sum_{i,j=1}^N \sigma_k^i(x)\sigma_k^j(x)\frac{\partial^2}{\partial x_i \partial x_j}\varphi(x), \quad \varphi \in C_b^\infty(\mathbb{R}^N), \tag{2.7}$$

and $\{P_t\}_{t\geq 0}$ be the semigroup of linear operators

$$(P_t f)(x) := (e^{t\mathscr{L}} f)(x) := E[f(X(t,x))], \quad t \geq 0, \quad x \in \mathbb{R}^N, \tag{2.8}$$

where $f : \mathbb{R}^N \to \mathbb{R}$ is appropriately chosen. For a fixed $z \in \mathbb{R}^N$, we define a generator $\mathscr{L}_0^z$ whose coefficients are frozen at a point $z \in \mathbb{R}^N$ as follows:

$$\mathscr{L}_0^z\varphi(x) = \sum_{i=1}^N b^i(z)\frac{\partial}{\partial x_i}\varphi(x) + \frac{1}{2}\sum_{k=1}^d\sum_{i,j=1}^N \sigma_k^i(z)\sigma_k^j(z)\frac{\partial^2}{\partial x_i \partial x_j}\varphi(x), \quad \varphi \in C_b^\infty(\mathbb{R}^N), \tag{2.9}$$

and let $\{P_t^{0,z}\}_{t>0}$ be the semigroup of linear operators corresponding to $\mathscr{L}_0^z$ given by

$$(P_t^{0,z} f)(x) := (e^{t\mathscr{L}_0^z} f)(x) := E[f(\bar{X}^z(t,x))], \quad t \geq 0, \quad x \in \mathbb{R}^N, \tag{2.10}$$

with an appropriate function $f : \mathbb{R}^N \to \mathbb{R}$, where

$$\bar{X}^z(t,x) = x + b(z)t + \sum_{i=1}^d \sigma_i(z)B_t^i. \tag{2.11}$$

Then, $(P_t^{0,z} f)(x)$ can be explicitly given using the Gaussian density $y \mapsto p^{\bar{X}^z}(t,x,y)$ of $\bar{X}^z(t,x)$ as

$$\left(P_t^{0,z} f\right)(x) = \int_{\mathbb{R}^N} f(y)p^{\bar{X}^z}(t,x,y)\,\mathrm{d}y, \qquad t > 0, \quad x \in \mathbb{R}^N. \tag{2.12}$$

For a multi-index $\alpha = (\alpha_1,\ldots,\alpha_k) \in \{1,2,\ldots,N\}^k, k \in \mathbb{N}$, we write $|\alpha| := k$ and $\partial^\alpha := \frac{\partial^{|\alpha|}}{\partial x_{\alpha_1}\cdots\partial x_{\alpha_k}}$. Let $\bar{X}^{\mathrm{EM}}(t,x)$ be the one-step Euler–Maruyama scheme:

$$\bar{X}^{\mathrm{EM}}(t,x) := \bar{X}^z(t,x)\Big|_{z=x} = x + b(x)t + \sum_{i=1}^d \sigma_i(x)B_t^i$$

and $\{\bar{X}^{\mathrm{EM},(n)}(t,x)\}_t$ be the continuous Euler–Maruyama scheme with the uniform time grids given by

$$\begin{aligned}\bar{X}^{\mathrm{EM},(n)}(t,x) = \bar{X}^{\mathrm{EM},(n)}(kT/n,x) &+ b\left(\bar{X}^{\mathrm{EM},(n)}(kT/n,x)\right)(t - kT/n)\\ &+ \sum_{i=1}^d \sigma_i\left(\bar{X}^{(\mathrm{EM},n)}(kT/n,x)\right)\left(B_t^i - B_{kT/n}^i\right),\end{aligned}$$

for $t \in [kT/n, (k+1)T/n]$. The aim of this paper is to provide a general discretization scheme for $P_T f(x)$ around the Euler–Maruyama scheme $E[f(\bar{X}^{\mathrm{EM},(n)}(T,x))]$.

In order to construct a discretization scheme, we expand $\mathscr{L}$ around $\mathscr{L}_0^z$ in the sense that; for all $m \in \mathbb{N}$, there is $C_{b,\sigma}(m) > 0$ depending on the derivatives of b, σ and m such that for all $\varphi \in C_b^\infty(\mathbb{R}^N)$ and $x, z \in \mathbb{R}^N$,

$$\left| \mathscr{L}\varphi(x) - \mathscr{L}_0^z \varphi(x) - \sum_{i=1}^m \mathscr{L}_i^z \varphi(x) \right| \leq C_{b,\sigma}(m) \sum_{i=1}^2 \|\nabla^i \varphi\|_\infty |x-z|^{m+1}, \tag{2.13}$$

where the differential operators $\mathscr{L}_i^z$, $i \in \mathbb{N}$ are defined by

$$\begin{aligned} \mathscr{L}_i^z := & \sum_{j=1}^N \sum_{\alpha \in \{1,2,\ldots,N\}^i} \frac{1}{|\alpha|!} \partial^\alpha b^j(\cdot)|_{\cdot = z} \prod_{j=1}^{|\alpha|} (x_{\alpha_j} - z_{\alpha_j}) \frac{\partial}{\partial x_j} \\ & + \frac{1}{2} \sum_{k=1}^d \sum_{j_1,j_2=1}^N \sum_{\alpha \in \{1,2,\ldots,N\}^i} \frac{1}{|\alpha|!} \partial^\alpha (\sigma_k^{j_1} \sigma_k^{j_2})(\cdot)|_{\cdot = z} \prod_{j=1}^{|\alpha|} (x_{\alpha_j} - z_{\alpha_j}) \frac{\partial^2}{\partial x_{j_1} \partial x_{j_2}}. \end{aligned} \tag{2.14}$$

For example, we have

$$\mathscr{L}_1^z \varphi(x) = \sum_{i,l=1}^N (x_l - z_l) \partial_l b^i(z) \frac{\partial}{\partial x_i} \varphi(x) + \sum_{k=1}^d \sum_{i,j,l=1}^N (x_l - z_l) \partial_l \sigma_k^i(z) \sigma_k^j(z) \frac{\partial^2}{\partial x_i \partial x_j} \varphi(x), \tag{2.15}$$

$$\begin{aligned} \mathscr{L}_2^z \varphi(x) = & \frac{1}{2} \sum_{i,l_1,l_2=1}^N (x_{l_1} - z_{l_1})(x_{l_2} - z_{l_2}) \partial_{l_1} \partial_{l_2} b^i(z) \frac{\partial}{\partial x_i} \varphi(x) \\ & + \sum_{k=1}^d \sum_{i,j,l_1,l_2=1}^N (x_{l_1} - z_{l_1})(x_{l_2} - z_{l_2}) \{ \partial_{l_1} \partial_{l_2} \sigma_k^i(z) \sigma_k^j(z) + \partial_{l_1} \sigma_k^i(z) \partial_{l_2} \sigma_k^j(z) \} \frac{\partial^2}{\partial x_i \partial x_j} \varphi(x), \end{aligned} \tag{2.16}$$

for $\varphi \in C_b^\infty(\mathbb{R}^N)$ and $x \in \mathbb{R}^N$. We use non-commutative relationship of $\mathscr{L}_0^z$ and $\mathscr{L}_i^z$, $i \in \mathbb{N}$ in the approximation scheme. Let $\mathcal{DO}$ be the space of smooth differential operators over $\mathbb{R}^N$. Then $\mathcal{DO}$ is a non-commutative algebra over $\mathbb{R}$. For $\mathscr{D}_1, \mathscr{D}_2 \in \mathcal{DO}$, we define the commutator $[\mathscr{D}_1, \mathscr{D}_2] = \mathscr{D}_1 \mathscr{D}_2 - \mathscr{D}_2 \mathscr{D}_1$. In addition, we define for $\mathscr{D}_1, \mathscr{D}_2, \ldots, \mathscr{D}_n \in \mathcal{DO}$, $n \in \mathbb{N}$, $\prod_{k=1}^n \mathscr{D}_k := \mathscr{D}_1, \mathscr{D}_2, \ldots, \mathscr{D}_n$.

2.1. Splitting around Euler–Maruyama

We shall use abbreviate notation

$$\bar{X}(t,x) \equiv \bar{X}^{\mathrm{EM}}(t,x) \quad \text{and} \quad \bar{X}^{(n)}(t,x) \equiv \bar{X}^{\mathrm{EM},(n)}(t,x), \tag{2.17}$$

for simplicity.

We now show a splitting method using a new operator constructed through the Baker–Campbell–Hausdorff expansion around the Euler–Maruyama scheme combined with Malliavin calculus.

Theorem 2.1. *For $T \geq 1$ and $m \in \mathbb{N}$, there exists a constant $C = C(T,m) > 0$ such that*

$$\|P_T f - \left(Q_{T/n}^{(m)}\right)^n f\|_\infty \leq C \|f\|_\infty \frac{1}{n^m}, \tag{2.18}$$

for all bounded and measurable functions $f:\mathbb{R}^N\to\mathbb{R}$ *and* $n\geq 1$, *where* $\{Q_t^{(m)}\}_{t>0}$ *is a family of operators given by*

$$\begin{aligned}Q_t^{(m)}f(x) = P_t^{0,z}f(x)|_{z=x} + \sum_{i=1}^{m-1}\sum_{i\leq\sum_{j=1}^i k_j\leq 2m+1}\sum_{\substack{1\leq\alpha_1\leq k_1\\ 0\leq\alpha_2\leq k_2\\ \cdots\\ 0\leq\alpha_i\leq k_i\\ \sum_{l=1}^i\alpha_l+i\leq m}}\frac{1}{\alpha!}I(\alpha)\\ \times t^{\sum_{l=1}^i\alpha_l+i}\prod_{l=1}^i\Big(\underbrace{[\mathscr{L}_0^z,[\mathscr{L}_0^z,\ldots,[\mathscr{L}_0^z,\mathscr{L}_{k_l}^z]]\ldots]]}_{\alpha_l\text{-times}}\Big)P_t^{0,z}f(x)|_{z=x}\\ = E\left[f(\bar{X}(t,x))\left\{1+\pi_t^{(m),x}(B_t)\right\}\right]\end{aligned}\tag{2.19}$$

for some $\pi_t^{(m),x}(B_t)\in\mathbb{D}^\infty$, *where* $\alpha! := (\alpha_1)!(\alpha_2)!\ldots(\alpha_i)!$ *and* $I(\alpha)$, $\alpha=(\alpha_1,\ldots,\alpha_i)$, *is defined as*

$$I(\alpha)=\int_0^1\int_{t_i}^1\ldots\int_{t_2}^1(t_1)^{\alpha_i}(t_2)^{\alpha_{i-1}}\ldots(t_i)^{\alpha_1}\,\mathrm{d}t_1\ldots\mathrm{d}t_i.\tag{2.20}$$

Proof of Theorem 2.1. See Section 2.3. □

In the discretization, an optimal truncation of the Baker–Campbell–Hausdorff expansion is used in (2.19) which is justified by Malliavin calculus. See Proposition 2.3 in the proof of Theorem 2.1.

As a corollary, we show a useful representation for the proposed splitting method and give a property on the variance.

Corollary 2.1. *It holds that*

$$\left(Q_{T/n}^{(m)}\right)^n f(x)=E\left[f\left(\bar{X}^{(n)}(T,x)\right)\prod_{i=1}^n\left\{1+\pi_{T/n}^{(m),\bar{X}^{(n)}((i-1)T/n,x)}\left(B_{iT/n}-B_{(i-1)T/n}\right)\right\}\right],\quad x\in\mathbb{R}^N.\tag{2.21}$$

Furthermore, we have the following result for the variance:

$$E\left[\left|f\left(\bar{X}^{(n)}(T,x)\right)\prod_{i=1}^n\left\{1+\pi_{T/n}^{(m),\bar{X}^{(n)}((i-1)T/n,x)}\left(B_{iT/n}-B_{(i-1)T/n}\right)\right\}\right|^2\right]<\infty,\tag{2.22}$$

for any $x\in\mathbb{R}^N$.

Proof. It is obvious from the proof of Theorem 2.1 (especially Lem. 2.1). □

Remark 2.1 (On the construction of the splitting method)**.** We explain the strategy for the construction of the splitting method for $(P_Tf)(x)=E[f(X(T,x))]$, $T\geq 1$ for a non-smooth test function f. Basically, the discretization is derived through two steps, the local and global approximations.

We first consider the local approximation. According to the generator expansion around the frozen generator:

$$\mathscr{L}=\mathscr{L}_0^z+\mathscr{L}_1^z+\mathscr{L}_2^z+\ldots,\tag{2.23}$$

the semigroup $\{P_t\}_t$ is expanded as

$$P_t\varphi(x)=P_t^{0,z}\varphi(x)\Big|_{z=x}+P_t^{1,z}\varphi(x)\Big|_{z=x}+P_t^{2,z}\varphi(x)\Big|_{z=x}+\ldots\tag{2.24}$$

for a bounded and measurable function φ, where the family of $P_t^{0,z}$, $t>0$ is the semigroup frozen at z corresponding to $\mathscr{L}_0^z$. Note that the family of $P_t^{0,z}|_{z=x}$, $t>0$ is regarded as the one-step Euler–Maruyama semigroup. The expansion (2.24) is obtained based on the expansion of the "parametrix":

$$P_t\varphi(x) = P_t^{0,z}\varphi(x)\big|_{z=x} + \int_0^t P_{t-s}(\mathscr{L}-\mathscr{L}_0^z)P_s^{0,z}\varphi(x)\,\mathrm{d}s\big|_{z=x} \tag{2.25}$$
$$= P_t^{0,z}\varphi(x)\Big|_{z=x} + \int_0^t P_{t-s}^{0,z}(\mathscr{L}-\mathscr{L}_0^z)P_s^{0,z}\varphi(x)\,\mathrm{d}s\big|_{z=x} + \ldots = P_t^{0,z}\varphi(x)\big|_{z=x} + \sum_{i=1}^{2m+1} P_t^{i,z}\varphi(x)\big|_{z=x} + \ldots,$$

where (2.25) and (2.23) (or (2.13)) are recursively used, and $P_t^{i,z}\varphi(x)$ is given by

$$P_t^{i,z}\varphi(x) = \sum_{k_1+\ldots+k_i=i}\int_0^t\int_{t_i}^t\cdots\int_{t_2}^t P_{t_i}^{0,z}\mathscr{L}_{k_1}^z P_{t_{i-1}-t_i}^{0,z}\mathscr{L}_{k_2}^z\ldots\mathscr{L}_{k_i}^z P_{t-t_1}^{0,z}\varphi(x)\,\mathrm{d}t_1\ldots\mathrm{d}t_i. \tag{2.26}$$

When we truncate the expansion at the order $2m+1$ as $P_t\varphi(x)\approx P_t^{0,z}\varphi(x)|_{z=x}+\sum_{i=1}^{2m+1}P_t^{i,z}\varphi(x)|_{z=x}$, the error term can be written as

$$\mathscr{R}_t^{\varphi}(x) = \sum_{k=0}^{2m+1}\int_0^t P_{t-s}\left(\mathscr{L}-\sum_{i=0}^{2m+1-k}\mathscr{L}_i^z\right)P_s^{k,z}\varphi(x)\,\mathrm{d}s|_{z=x}. \tag{2.27}$$

We will discuss details in the proof of Proposition 2.2.

While we obtain an approximation $P_t\varphi(x)\approx P_t^{0,z}\varphi(x)|_{z=x}+\sum_{i=1}^{2m+1}P_t^{i,z}\varphi(x)|_{z=x}$, we do not use this form directly in numerical computation and apply the Baker–Campbell–Hausdorff formula to $P_t^{i,z}\varphi(x)$, $i\geq 1$. Furthermore, we do not use all terms of $\sum_{i=1}^{2m+1}P_t^{i,z}\varphi(x)|_{z=x}$ in order to obtain more smart local approximation for splitting method.

After obtaining the approximation $P_t\varphi(x)\approx P_t^{0,z}\varphi(x)|_{z=x}+\sum_{i=1}^{2m+1}P_t^{i,z}\varphi(x)|_{z=x}$, we give an optimal truncation using the Baker–Campbell–Hausdorff formula and the Malliavin calculus, which may be regarded as a "truncated parametrix". The following local approximation is introduced:

$$P_t\varphi(x)\approx Q_t^{(m)}\varphi(x) \tag{2.28}$$

with

$$Q_t^{(m)}\varphi(x) = P_t^{0,z}\varphi(x)|_{z=x} + \sum_{i=1}^{m-1}\sum_{i\leq\sum_{j=1}^i k_j\leq 2m+1}\sum_{\substack{1\leq\alpha_1\leq k_1\\0\leq\alpha_2\leq k_2\\\cdots\\0\leq\alpha_i\leq k_i\\\sum_{l=1}^i\alpha_l+i\leq m}}\int_0^t\int_{t_i}^t\cdots\int_{t_2}^t\frac{(t_1)^{\alpha_i}(t_2)^{\alpha_{i-1}}\ldots(t_i)^{\alpha_1}}{\alpha_1!\alpha_2!\ldots\alpha_i!}\,\mathrm{d}t_1\ldots\mathrm{d}t_i$$
$$\times\prod_{l=1}^{i}\Big(\underbrace{[\mathscr{L}_0^z,[\mathscr{L}_0^z,\ldots,[\mathscr{L}_0^z,\mathscr{L}_{k_l}^z]]\ldots]}_{\alpha_l\text{-times}}\Big)P_t^{0,z}\varphi(x)|_{z=x} \tag{2.29}$$

The above truncation in (2.28) is "optimal" in the sense that $(Q_{T/n}^{(m)})^n f(x)$ will give $O(1/n^m)$-order approximation for $P_Tf(x)$ with the minimum computational effort. The local approximation error is explicitly given as follows $P_t\varphi(x)-Q_t^{(m)}\varphi(x)=\mathscr{R}_t^{\varphi}(x)+t^{m+1}\Psi_t^{\varphi}(x)$ where $\mathscr{R}_t^{\varphi}$ is given in (2.27) which can be estimated by $\|\mathscr{R}_t^{\varphi}\|_\infty\leq t^{m+1}C\|\varphi\|_\infty$ using Malliavin calculus, and Ψ_t^{φ} is a function of the form $\Psi_t^{\varphi}(x)=\sum_{l\leq\nu}h_l(t)g_l(\cdot)\partial^{\beta^l}P_t^{0,z}\varphi(\cdot)|_{z=\cdot}$ with $\nu\in\mathbb{N}$, multi-indices $\beta^l\in\{1,\ldots,N\}^l$, $l\leq\nu$, non-decreasing h_l and smooth and bounded g_l.

Next, we explain the global approximation for $P_Tf(x)$. When the test function f is smooth, the result $\|P_Tf-(Q_{T/n}^{(m)})^n\varphi\|_\infty=O(n^{-m})$ immediately follows. Actually, if φ is a function of C_b^∞-class, we easily see that

$\|P_t\varphi - Q_t^{(m)}\varphi\|_\infty = O(t^{m+1})$ holds by the estimate $\|\Psi_t^\varphi\|_\infty \leq C$ with the result $\sup_{t>0}\|\partial^{\beta^l}P_t^{0,\cdot}\varphi(\cdot)\|_\infty < \infty$, and then one has $\|P_Tf - (Q_{T/n}^{(m)})^n f\|_\infty = n \times O((T/n)^{m+1}) = O(n^{-m})$.

However, when f is non-smooth, we cannot employ this argument. The important point on the construction of the splitting method in the paper is to use the explicit local error functions $\mathscr{R}_t^\varphi$ and Ψ_t^φ in the estimate of the global approximation for P_Tf with bounded and measurable test functions f, where the Malliavin calculus plays a crucial role. In particular, by applying Kusuoka–Stroock's integration by parts [14] for elliptic Itô processes, we are able to provide the global approximation $P_Tf(x) = (Q_{T/n}^{(m)})^n f(x) + O(1/n^m)$ for bounded and measurable test functions f, in other words, the weak approximation for Itô SDE is obtained.

Remark 2.2 (Comparison with the probabilistic parametrix methods)**.** We mention the features of the proposed scheme (Thm. 2.1 (and Cor. 2.1)) by comparing with the probabilistic parametrix methods of Bally and Kohatsu-Higa [2] and Labordere *et al.* [15].

Bally and Kohatsu-Higa [2] obtained an exact formula with an estimator for $P_Tf(x)$ through the parametrix method, where there is no discretization error (weak approximation error), in other words, the estimator provided by [2] is "unbiased". However, the cost of the scheme may be the divergent of the variance. In general, the estimator gives infinite variance except for some special cases, which is partially improved by Labordere *et al.* [15].

Our scheme will be regarded as a biased-simulation method because it involves weak approximation error. However the bias is quite small since it gives higher order discretization. Also the variance of the estimator is finite (see Cor. 2.1), which is consequence of the use of the Baker–Campbell–Hausdorff formula for the "truncated parametrix" (Rem. 2.1).

2.2. Examples of the splitting method

As examples for the splitting method in Theorem 2.1 (and Cor. 2.1), we show the following simple second and third order methods in Corollaries 2.2 and 2.3.

Corollary 2.2 (Second order weak approximation)**.** *For $T \geq 1$, there exists a constant $C > 0$ such that*

$$\|P_Tf - (Q_{T/n}^{(2)})^n f\|_\infty \leq C\|f\|_\infty \frac{1}{n^2}, \tag{2.30}$$

for all bounded and measurable functions $f : \mathbb{R}^N \to \mathbb{R}$ and $n \geq 1$, where $\{Q_t^{(2)}\}_{t>0}$ is a family of operators given by

$$Q_t^{(2)} f(x) = P_t^{0,z} f(x)|_{z=x} + \frac{t^2}{2}\sum_{i=1}^{2}\left[\mathscr{L}_0^z, \mathscr{L}_i^z\right] P_t^{0,z} f(x)|_{z=x} \tag{2.31}$$

$$= E\left[f\left(\bar{X}(t,x)\right)\left\{1 + \pi_t^{(2),x}(B_t)\right\}\right]. \tag{2.32}$$

Here, the Malliavin weight $\pi_t^{(2),x}(B_t)$ is given by

$$\pi_t^{(2),x}(B_t) = \frac{1}{2}\sum_{i_1,i_2=1}^{d}\sum_{k_1=1}^{N}\mathcal{L}_{i_1}\sigma_{i_2}^{k_1}(x)H_{(k_1)}\left(\bar{X}(t,x), B_t^{i_1}B_t^{i_2} - t1_{i_1=i_2}\right) \tag{2.33}$$

$$+ \frac{1}{2}\sum_{i_1=1}^{d}\sum_{k_1=1}^{N}\left\{\mathcal{L}_0\sigma_{i_1}^{k_1}(x) + \mathcal{L}_{i_1}b^{k_1}(x)\right\}H_{(k_1)}\left(\bar{X}(t,x), B_t^{i_1}t\right) \tag{2.34}$$

$$+ \frac{1}{4}\sum_{i_1,i_2=1}^{d}\sum_{k_1,k_2=1}^{N}\mathcal{L}_{i_1}\sigma_{i_2}^{k_1}(x)\mathcal{L}_{i_1}\sigma_{i_2}^{k_2}(x)H_{(k_1,k_2)}\left(\bar{X}(t,x), t^2\right) \tag{2.35}$$

$$+ \frac{1}{2}\sum_{k_1=1}^{N}\mathcal{L}_0 b^{k_1}(x)H_{(k_1)}\left(\bar{X}(t,x), t^2\right), \tag{2.36}$$

with the differential operators appearing in the Itô Taylor expansion:

$$\mathcal{L}_i = \sum_{j=1}^N \sigma_i^j(x)\frac{\partial}{\partial x_j}, \quad i=1,\dots,d,$$

$$\mathcal{L}_0 = \sum_{j=1}^N b^j(x)\frac{\partial}{\partial x_j} + \frac{1}{2}\sum_{j_1,j_2=1}^N \sum_{i=1}^d \sigma_i^{j_1}(x)\sigma_i^{j_2}(x)\frac{\partial^2}{\partial x_{j_1}\partial x_{j_2}}.$$

Remark 2.3 (Explicit formula for the weight in Cor. 2.2)**.** We provide below the explicit forms of

$$H_{(k_1)}(\bar{X}(t,x), B_t^{i_1}B_t^{i_2} - t1_{i_1=i_2}),\ H_{(k_1)}(\bar{X}(t,x), B_t^{i_1}t),\ H_{(k_1,k_2)}(\bar{X}(t,x),1),$$

$i_1,i_2=1,\dots,d$, $k_1,k_2=1,\dots,N$, in the Malliavin weight $\pi_t^{(2),x}(B_t)$ in Corollary 2.2. Let $A(x)=(A_i^j(x))_{1\le i,j\le N}$ be the inverse matrix of $\sigma\sigma'(x)$. By the computation of Skorohod integral (2.3) with (2.4) and (2.5), we have

$$\begin{aligned} H_{(k_1)}(\bar{X}(t,x), B_t^{i_1}B_t^{i_2} - t1_{i_1=i_2}) &= \sum_{j_1=1}^N \sum_{i_3=1}^d A_{j_1}^{k_1}(x)V_{i_3}^{j_1}(x)t^{-1}\{B_t^{i_1}B_t^{i_2}B_t^{i_3} \\ &\quad - B_t^{i_1}t1_{i_2=i_3} - B_t^{i_2}t1_{i_1=i_3} - B_t^{i_3}t1_{i_1=i_2}\}, \end{aligned} \tag{2.37}$$

$$H_{(k_1)}(\bar{X}(t,x), B_t^{i_1}t) = \sum_{j_1=1}^N \sum_{i_3=1}^d A_{j_1}^{k_1}(x)V_{i_3}^{j_1}(x)\{B_t^{i_1}B_t^{i_3} - t1_{i_1=i_3}\}, \tag{2.38}$$

$$H_{(k_1,k_2)}(\bar{X}(t,x), t^2) = \sum_{j_1,j_2=1}^N \sum_{i_3,i_4=1}^d A_{j_1}^{k_1}(x)V_{i_3}^{j_1}(x)A_{j_2}^{k_2}(x)V_{i_4}^{j_2}(x)\{B_t^{i_3}B_t^{i_4} - t1_{i_3=i_4}\}, \tag{2.39}$$

for $k_1,k_2=1,\dots,N$ and $i_1,i_2=1,\dots,d$.

Corollary 2.3 (Third order weak approximation)**.** *For $T\ge 1$, there exists a constant $C>0$ such that*

$$\|P_T f - (Q_{T/n}^{(3)})^n f\|_\infty \le C\|f\|_\infty \frac{1}{n^3}, \tag{2.40}$$

for all bounded and measurable functions $f:\mathbb{R}^N\to\mathbb{R}$ and $n\ge 1$, where $\{Q_t^{(3)}\}_{t>0}$ is a family of operators given by

$$\begin{aligned} Q_t^{(3)}f(x) &= P_t^{0,z}f(x)|_{z=x} + \frac{t^2}{2}\sum_{i=1}^2 [\mathscr{L}_0^z, \mathscr{L}_i^z]P_t^{0,z}f(x)|_{z=x} \\ &\quad + \frac{t^3}{6}\sum_{i=2}^4 [\mathscr{L}_0^z,[\mathscr{L}_0^z,\mathscr{L}_i^z]]P_t^{0,z}f(x)|_{z=x} + \frac{t^3}{6}\sum_{i_1=1}^2\sum_{i_1+i_2\le 4}[\mathscr{L}_0^z,\mathscr{L}_{i_1}^z]\mathscr{L}_{i_2}^z P_t^{0,z}f(x)|_{z=x}. \end{aligned}$$

2.3. Proof of Theorem 2.1

For a bounded and measurable function $\varphi:\mathbb{R}^N\to\mathbb{R}$, we define functions $P_t^{i,z}\varphi$, $t>0$, $i=1,2,\dots,2m+1$ given by

$$P_t^{i,z}\varphi(x) = \sum_{k=0}^{i-1}\int_0^t P_{t-s}^{0,z}\mathscr{L}_{i-k}^z P_s^{k,z}\varphi(x)\,\mathrm{d}s, \quad x,z\in\mathbb{R}^N, \tag{2.41}$$

which play a role in the construction of the approximation for $P_t\varphi$. We note that for any $i\in\mathbb{N}$, $k=0,\dots,i-1$, $x\in\mathbb{R}^N$, one has $s\mapsto P_{t-s}^{0,z}\mathscr{L}_{i-k}^z P_s^{0,z}\varphi(x)\in L^1([0,t])$ by the integration by parts argument. Let us see this in the case $i = 1$ ($k = 0$). From the the integration by parts and the definition of $\mathscr{L}_1^z$ in (2.15), we have the two representations for $P_{t-s}^{0,z}\mathscr{L}_1^z P_s^{0,z}\varphi(x)$:

$$P_{t-s}^{0,z}\mathscr{L}_1^z P_s^{0,z}\varphi(x) = \sum_{\substack{\alpha\in\{1,2,\ldots,N\}^k,\\ k\le 2}} \sum_{l=1}^{N} E\left[\partial^\alpha P_s^{0,z}\varphi\left(\bar{X}^z(t-s,x)\right) g_{\alpha,l}(z)\left(\bar{X}^{z,l}(t-s,x)-z_l\right)\right] \tag{2.42}$$

$$= \sum_{\substack{\alpha\in\{1,2,\ldots,N\}^k,\\ k\le 2}} \sum_{l=1}^{N} E\left[P_s^{0,z}\varphi\left(\bar{X}^z(t-s,x)\right) H_\alpha\left(\bar{X}^z(t-s,x), g_{\alpha,l}(z)\left(\bar{X}^{z,l}(t-s,x)-z_l\right)\right)\right] \tag{2.43}$$

for some $g_{\alpha,l}\in C_b^\infty(\mathbb{R}^N)$, $\alpha\in\{1,2,\ldots,N\}^k$, $k=1,2$, $l=1,\ldots,N$. Hence, we can choose positive constants C_1 and C_2 such that

$$\begin{aligned}\int_0^t \left|P_{t-s}^{0,z}\mathscr{L}_1^z P_s^{0,z}\varphi(x)\right| \mathrm{d}s &= \int_0^{t/2} \left|P_{t-s}^{0,z}\mathscr{L}_1^z P_s^{0,z}\varphi(x)\right| \mathrm{d}s + \int_{t/2}^t \left|P_{t-s}^{0,z}\mathscr{L}_1^z P_s^{0,z}\varphi(x)\right| \mathrm{d}s\\ &\le C_1\int_0^{t/2}\frac{\|\varphi\|_\infty}{t-s}\,\mathrm{d}s + C_2\int_{t/2}^t \frac{\|\varphi\|_\infty}{s}\,\mathrm{d}s = (C_1+C_2)\|\varphi\|_\infty 2\log 2\end{aligned} \tag{2.44}$$

where the estimate of Kusuoka and Stroock [14] (or the basic Gaussian calculus, in this case) is applied to (2.42) and (2.43). Then $s\mapsto P_{t-s}^{0,z}\mathscr{L}_1^z P_s^{0,z}\varphi(x)\in L^1([0,t])$ holds for bounded and measurable test functions. The integrability of the integrand of $P_t^{i,z}\varphi(x)$ for $i\ge 2$ follows by induction.

Before showing the proof of Theorem 2.1, we prepare three useful results (Props. 2.1–2.3).

First, we give the Baker–Campbell–Hausdorff type formula in order to compute $P_t^{i,z}\varphi(x)|_{z=x}$, $i\ge 1$. See the book of [5] for the topics on the Baker–Campbell–Hausdorff formula initialized in [1,6,11]. In our application, we need the following specific version of Baker–Campbell–Hausdorff formula for expectations with bounded and measurable test functions.

Proposition 2.1 (Baker–Campbell–Hausdorff formula). *Let $0<s<t\le 1$, $i\in\mathbb{N}$ and $\widehat{\mathscr{L}_i}\in\mathcal{DO}$ be a differential operator of the form $\widehat{\mathscr{L}_i}=c\psi_i(\cdot)\partial^\beta$ where c is a constant, $\psi_i(\cdot)$ is a polynomial of the degree at most i and ∂^β is a partial derivative with a multi-index $\beta\in\{1,\ldots,N\}^\ell$, $\ell\in\mathbb{N}$. Then we have the explicit formula:*

$$e^{s\mathscr{L}_0^z}\widehat{\mathscr{L}_i}e^{(t-s)\mathscr{L}_0^z}\varphi(\cdot) = \sum_{k=0}^{i}\frac{s^k}{k!}\underbrace{\left[\mathscr{L}_0^z,\left[\ldots\left[\mathscr{L}_0^z,\left[\mathscr{L}_0^z,\widehat{\mathscr{L}_i}\right]\right]\ldots\right]\right]}_{k\text{-times}} e^{t\mathscr{L}_0^z}\varphi(\cdot), \tag{2.45}$$

for any bounded and measurable function $\varphi:\mathbb{R}^N\to\mathbb{R}$, where $\underbrace{[\mathscr{L}_0^z,[\ldots[\mathscr{L}_0^z,[\mathscr{L}_0^z,\widehat{\mathscr{L}_i}]]\ldots]]}_{0\text{-times}}\equiv\widehat{\mathscr{L}_i}$.

Proof. We provide the proof in Appendix A. □

Using the explicit Baker–Campbell–Hausdorff formula (2.45) and the Malliavin calculus, we have the following estimate.

Proposition 2.2. *For $m\in\mathbb{N}$, there exists a constant $C=C(m)>0$ such that*

$$\sup_{x\in\mathbb{R}^N}\left|P_t\varphi(x)-\sum_{i=0}^{2m+1}P_t^{i,z}\varphi(x)|_{z=x}\right|\le Ct^{m+1}\|\varphi\|_\infty \tag{2.46}$$

for any $t\in(0,1]$ and bounded and measurable function $\varphi:\mathbb{R}^N\to\mathbb{R}$.

Proof. We provide the proof in Appendix B. ☐

Furthermore, applying the explicit Baker–Campbell–Hausdorff formula for $\sum_{i=0}^{2m+1} P_t^{i,z}\varphi(x)|_{z=x}$, we give the optimal truncation in order to give $O(t^{m+1})$-order local approximation for $P_t\varphi(x)$.

Proposition 2.3. *For $m \in \mathbb{N}$ and a bounded and measurable function $\varphi : \mathbb{R}^N \to \mathbb{R}$, it holds that*

$$\sum_{i=0}^{2m+1} P_t^{i,z}\varphi(x)|_{z=x} = Q_t^{(m)}\varphi(x) \tag{2.47}$$

$$+ \sum_{i=1}^{m-1} \sum_{i\leq\sum_{j=1}^i k_j\leq 2m+1} \sum_{\substack{1\leq\alpha_1\leq k_1\\ 0\leq\alpha_2\leq k_2\\ \cdots\\ 0\leq\alpha_i\leq k_i\\ \sum_{l=1}^i\alpha_l+i>m}} \frac{t^{\sum_{l=1}^i\alpha_l+i}}{\alpha!} I(\alpha) \prod_{l=1}^i \Big(\underbrace{[\mathscr{L}_0^z,[\mathscr{L}_0^z,\ldots,[\mathscr{L}_0^z,\mathscr{L}_{k_l}^z]]\ldots]]}_{\alpha_l\text{-times}}\Big) P_t^{0,z}\varphi(x)|_{z=x} \tag{2.48}$$

$$+ \sum_{i=m}^{2m+1} \sum_{i\leq\sum_{j=1}^i k_j\leq 2m+1} \sum_{\substack{1\leq\alpha_1\leq k_1\\ 0\leq\alpha_2\leq k_2\\ \cdots\\ 0\leq\alpha_i\leq k_i}} \frac{t^{\sum_{l=1}^i\alpha_l+i}}{\alpha!} I(\alpha) \prod_{l=1}^i \Big(\underbrace{[\mathscr{L}_0^z,[\mathscr{L}_0^z,\ldots,[\mathscr{L}_0^z,\mathscr{L}_{k_l}^z]]\ldots]]}_{\alpha_l\text{-times}}\Big) P_t^{0,z}\varphi(x)|_{z=x}, \tag{2.49}$$

where $\{Q_t^{(m)}\}_{t>0}$ is a family of operators given in Theorem 2.1. In particular, the terms (2.48) and (2.49) are simply written in the following form:

$$t^{m+1}\sum_{l\leq\nu} h_l(t)g_l(x)\partial^{\beta^{(l)}} P_t^{0,z}\varphi(x)|_{z=x}, \tag{2.50}$$

for some $\nu\in\mathbb{N}$, functions h_l, $l\leq\nu$ at most polynomial growth, $g_l\in C_b^\infty(\mathbb{R}^N)$, $l\leq\nu$ and multi-indices $\beta^{(l)}$, $l\leq\nu$.

Proof. We will show this in Appendix C. ☐

Using these results, we proceed to the proof of Theorem 2.1.

Proof of Theorem 2.1. For any bounded and measurable function $\varphi : \mathbb{R}^N \to \mathbb{R}$ and $t\in(0,1]$, we define a function $\mathscr{R}_t^\varphi(\cdot)$ on $\mathbb{R}^N$ as $\mathscr{R}_t^\varphi(x) := P_t\varphi(x) - \sum_{i=0}^{2m+1} P_t^{i,z}\varphi(x)|_{z=x}$, $x\in\mathbb{R}^N$. Then, by Proposition 2.3, the local error $P_t\varphi(x) - Q_t^{(m)}\varphi(x)$ is given by

$$P_t\varphi(x) - Q_t^{(m)}\varphi(x) = \mathscr{R}_t^\varphi(x) + t^{m+1}\Psi_t^\varphi(x), \tag{2.51}$$

where $\Psi_t^\varphi : \mathbb{R}^N \to \mathbb{R}$ is defined as $\Psi_t^\varphi(\cdot) := \sum_{l\leq\nu} h_l(t)g_l(\cdot)\partial^{\beta^{(l)}} P_t^{0,z}\varphi(\cdot)|_{z=\cdot}$ with those functions appearing in (2.50). When the bounded and measurable function φ is sufficiently smooth, we immediately see that the function $\partial^{\beta^{(l)}} P_t^{0,z}\varphi(\cdot)|_{z=\cdot}$ in $\Psi_t^\varphi(\cdot)$ has the form $\partial^{\beta^{(l)}} P_t^{0,z}\varphi(x)|_{z=x} = E[\partial^{\beta^{(l)}}\varphi(\bar{X}(t,x))]$. However, when φ is only bounded and measurable we cannot use the derivatives of φ, of which case is discussed in Lemma 2.2.

We show the bound for the global error $P_Tf(x) - (Q_{T/n}^{(m)})^n f(x)$ without employing the regularity on f. We note that by the integration by parts on Wiener space, for any bounded and measurable function $f:\mathbb{R}^N\to\mathbb{R}$, $Q_t^{(m)}f(x)$ is represented as

$$Q_t^{(m)}f(x) = E\left[f\big(\bar{X}(t,x)\big)\big\{1+\pi_t^{(m),x}(B_t)\big\}\right], \tag{2.52}$$

where $\pi_t^{(m),x}(B_t) \in \mathbb{D}^\infty$ is given as a sum of polynomials of Brownian motion $\{B_t\}_{t>0}$ with coefficients depending on $b, \sigma_1, \ldots, \sigma_d$ and theirs partial derivatives evaluated at $x \in \mathbb{R}^N$. Here we have

$$P_T f(x) - \left(Q_{T/n}^{(m)}\right)^n f(x) = \sum_{k=0}^{n-1} \left(Q_{T/n}^{(m)}\right)^k (P_{T/n} - Q_{T/n}) P_{T-(k+1)T/n} f(x) \tag{2.53}$$

$$= \sum_{k=0}^{n-1} \left(Q_{T/n}^{(m)}\right)^k \mathscr{R}_{T/n}^{P_{T-(k+1)T/n}f}(x) + (T/n)^{m+1} \sum_{k=0}^{n-1} \left(Q_{T/n}^{(m)}\right)^k \Psi_{T/n}^{P_{T-(k+1)T/n}f}(x)$$

$$= \sum_{k=0}^{n-1} E\left[\mathscr{R}_{T/n}^{P_{T-(k+1)T/n}f}\left(\bar{X}^{(n)}(kT/n, x)\right) \times \prod_{j=1}^{k} \left\{1 + \pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)T/n,x)}\left(B_{jT/n} - B_{(j-1)T/n}\right)\right\}\right]$$

$$+ (T/n)^{m+1} \sum_{k=0}^{n-1} E\left[\Psi_{T/n}^{P_{T-(k+1)T/n}f}\left(\bar{X}^{(n)}(kT/n, x)\right) \times \prod_{j=1}^{k} \left\{1 + \pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)T/n,x)}(B_{jT/n} - B_{(j-1)T/n})\right\}\right], \tag{2.54}$$

where the representation (2.52) is iteratively applied. We note that the following properties hold for the Euler–Maruyama scheme and the weights in (2.54).

Lemma 2.1. *We have* $\bar{X}^{(n)}((k-1)T/n, x) \in \mathbb{D}^\infty$ *and* $\prod_{j=1}^{k}\{1 + \pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)T/n,x)}(B_{jT/n} - B_{(j-1)T/n})\} \in \mathbb{D}^\infty$ *for all* $k = 1, \ldots, n$. *In other words, for* $K \in \mathbb{N}$, $p = 2e$, $e \in \mathbb{N}$, *there exists* $C(T) > 0$ *such that*

$$\sup_{k=1,\ldots,n} \left\|\bar{X}^{(n)}((k-1)T/n, x)\right\|_{K,p} \leq C(T), \tag{2.55}$$

$$\sup_{k=1,\ldots,n} \left\| \prod_{j=1}^{k} \left\{1 + \pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)T/n,x)}(B_{jT/n} - B_{(j-1)T/n})\right\}\right\|_{K,p} \leq C(T). \tag{2.56}$$

Proof. The result $\bar{X}^{(n)}((k-1)T/n, x) \in \mathbb{D}^\infty$, $k = 1, \ldots, n$ or (2.55) is given in Lemma 5.1 of Bally and Talay [3]. We will give the proof of (2.56) in Appendix D. □

First, we will estimate the terms $(Q_{T/n}^{(m)})^k \mathscr{R}_{T/n}^{P_{T-(k+1)T/n}f}(x)$, $k = 0, 1, 2, \ldots, n-1$. By Lemma 2.1, it holds that there exists $C(T) > 0$ which does not depend on k such that

$$\|(Q_{T/n}^{(m)})^k \mathscr{R}_{T/n}^{P_{T-(k+1)T/n}f}\|_\infty \leq \|\mathscr{R}_{T/n}^{P_{T-(k+1)T/n}f}\|_\infty C(T)$$

for any $k = 0, 1, \ldots, n-1$. Then, using Proposition 2.2 leads to

$$\left\|(Q_{T/n}^{(m)})^k \mathscr{R}_{T/n}^{P_{T-(k+1)T/n}f}\right\|_\infty \leq C(T)\|f\|_\infty \frac{1}{n^{m+1}} \tag{2.57}$$

with some constant $C(T) > 0$ which does not depend on $k = 0, 1, \ldots, n-1$.

In order to complete the proof, it suffices to show that there exists a constant $C > 0$ independent of n such that $\|(Q_{T/n}^{(m)})^k \Psi_{T/n}^{P_{T-(k+1)T/n}f}\|_\infty \leq C\|f\|_\infty$ for $k = 0, 1, 2, \ldots, n-1$. The term $(Q_{T/n}^{(m)})^k \Psi_{T/n}^{P_{T-(k+1)T/n}f}(x)$ is represented with the Euler–Maruyama scheme as follows: for $k = 0, 1, 2, \ldots, n-2$,

$$\left(Q_{T/n}^{(m)}\right)^k \Psi_{T/n}^{P_{T-(k+1)T/n}f}(x) = \sum_{l \leq \nu} h_l(T/n) E\left[\partial^{\beta^{(l)}} P_{T-(k+1)T/n} f\left(\bar{X}^{(n)}((k+1)T/n, x)\right) G_l^{(k)}\right], \tag{2.58}$$

where $\mathcal{F}_{kT/n}$-measurable random variable $G_l^{(k)} \in \mathbb{D}^\infty$, $l \leq \nu$ is given by

$$G_l^{(k)} = g_l\left(\bar{X}^{(n)}((k-1)T/n, x)\right) \prod_{j=1}^{k} \left\{1 + \pi_{T/n}^{(m), \bar{X}^{(n)}((j-1)T/n, x)}\left(B_{jT/n} - B_{(j-1)T/n}\right)\right\}. \tag{2.59}$$

Note that now we exclude the case where $k = n-1$ since we cannot have $\partial^{\beta^{(l)}} f$ for a bounded and measurable function $f : \mathbb{R}^N \to \mathbb{R}$ while we can use $\partial^{\beta^{(l)}} P_{T-(k+1)T/n} f$ for $k = 1, \ldots, n-2$.

For $s \in (0, T/2)$ and a multi-index α, since we can choose $C > 0$ and $q > 1$ such that $\|\partial^\alpha P_{T-s} f\|_\infty \leq \|f\|_\infty C/T^q$ by Friedman [8] or Kusuoka and Stroock [14], we immediately have

$$\|(Q_{T/n}^{(m)})^k \Psi_{T/n}^{P_{T-(k+1)T/n}f}\|_\infty \leq C(T)\|f\|_\infty, \quad k = 0, 1, \ldots, [n/2].$$

For $s \in [T/2, T)$, another application of Kusuoka and Stroock [14] enables us to give the upper bound of (2.58). We can remove the derivatives of $P_{T-(k+1)T/n} f$ through the integration by parts with respect to the elliptic Itô process $\{\bar{X}^{(n)}(s, x)\}_{s>0}$ and then have the estimate: for all multi-index α, there are $C > 0$ and $q > 1$ such that

$$\sup_{x \in \mathbb{R}^N} |E[\partial^\alpha P_{T-s} f(\bar{X}^{(n)}(s, x))G]| \leq C\frac{1}{T^q}\|P_{T-s} f\|_\infty \leq C\frac{1}{T^q}\|f\|_\infty,$$

for all $G \in \mathbb{D}^\infty$ such that for all $k \geq 1, p \in (1, \infty)$, $\|G\|_{k,p} \leq C$. Hence, one has

$$\|(Q_{T/n}^{(m)})^k \Psi_{T/n}^{P_{T-(k+1)T/n}f}\|_\infty \leq C(T)\|f\|_\infty, \quad k = [n/2] + 1, \ldots, n-2.$$

Even if $k = n-1$, we can show the upper bound of (2.58) as follows:

Lemma 2.2. *There exist $C > 0$ and $q > 0$ such that for any bounded and measurable function $f : \mathbb{R}^N \to \mathbb{R}$,*

$$\|(Q_{T/n}^{(m)})^{n-1} \Psi_{T/n}^f\|_\infty \leq C\frac{1}{T^q}\|f\|_\infty. \tag{2.60}$$

Proof. We will give the proof in Appendix E. ☐

Therefore, we have

$$\left\|(Q_{T/n}^{(m)})^k (P_{T/n} - Q_{T/n}) P_{T-(k+1)T/n} f\right\|_\infty \leq C(T)\frac{1}{n^{m+1}}\|f\|_\infty, \quad k = 0, 1, \ldots, n-1,$$

and in conclusion

$$\left\|P_T f - (Q_{T/n}^{(m)})^n f\right\|_\infty \leq C(T)\sum_{k=0}^{n-1} \frac{1}{n^{m+1}}\|f\|_\infty = C(T)\frac{1}{n^m}\|f\|_\infty.$$

The proof of Theorem 2.1 is finished. ☐

3. Discretization of heat kernel

Let $y \mapsto p_t^X(x,y)$ be the density of $X(t,x)$ or the fundamental solution (the heat kernel) associated with the backward heat operator $\partial_t + \mathscr{L}$. We aim at constructing a general discretization method for $p_T^X(x,y)$, $T \geq 1$, $x, y \in \mathbb{R}^N$.

Let us denote by $p_t^{\bar{X}^z}(x,\cdot)$ the density of $\bar{X}^z(t,x)$ given in (2.11), *i.e.* $\bar{X}^z(t,x) = x + b(z)t + \sum_{i=1}^d \sigma_i(z)B_t^i$. Since $\bar{X}^z(t,x)$ follows N-dimensional normal distribution with the mean $x+b(z)t$ and the variance-covariance $t\sigma(z)\sigma'(z)$, the density $p_t^{\bar{X}^z}(x,y)$, $(t,x,y) \in (0,\infty)\times\mathbb{R}^N\times\mathbb{R}^N$ is explicitly given by

$$p_t^{\bar{X}^z}(x,y) = \frac{1}{(2\pi t)^{N/2}\sqrt{\det \Sigma(z)}} \exp\left(-\frac{1}{2t}\Big(y - \{x+b(z)t\}\Big)' A(z)\Big(y-\{x+b(z)t\}\Big)\right), \tag{3.1}$$

where $\Sigma(z) = \sigma(z)\sigma'(z)$ and $A(z) = \Sigma(z)^{-1}$. Note that the uniform ellipticity of $\sigma\sigma'$ guarantees that the $N \times N$ matrix $\Sigma(z)$ is invertible for all $z \in \mathbb{R}^N$. When we substitute $z = x$ into $p_t^{\bar{X}^z}(x,y)$, we simply write it as $p_t^{\bar{X}}(x,y)$. Using the weight $\pi_t^{(m),x}(B_t)$ in Theorem 2.1 or (2.52), *i.e.* $Q_t^{(m)}f(x) = E[f(\bar{X}(t,x))\{1+\pi_t^{(m),x}(B_t)\}]$, we give a new discretization of the heat kernel $p_T^X(x,\cdot)$.

The main result is as follows.

Theorem 3.1. *Assume $T \geq 1$, $m \in \mathbb{N}$. There exist some constants $C(m), c > 0$, $q \geq \frac{N}{2}$ independent of T and non-decreasing function $K(\cdot)$ such that*

$$\begin{aligned}&\left| p_T^X(x,y) - E\left[p_{T/n}^{\bar{X},(m)}\left(\bar{X}^{(n)}((n-1)T/n,x),y\right) \prod_{j=1}^{n-1}\left\{1+\pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)T/n,x)}\left(B_{jT/n}-B_{(j-1)T/n}\right)\right\}\right]\right| \\ &\quad \leq \frac{K(T)}{n^m}\frac{C(m)}{T^q}\exp\left(-c\frac{|y-x|^2}{T}\right),\end{aligned} \tag{3.2}$$

for all $x, y \in \mathbb{R}^N$, where $p_t^{\bar{X},(m)}(x,y)$, $(t,x,y) \in (0,\infty)\times\mathbb{R}^N\times\mathbb{R}^N$ is given by

$$\begin{aligned} p_t^{\bar{X},(m)}(x,y) &= p_t^{\bar{X}}(x,y) \\ &\quad + \sum_{i=1}^{m-1} \sum_{i \leq \sum_{j=1}^i k_j \leq 2m+1} \sum_{\substack{1\leq\alpha_1\leq k_1 \\ 0\leq\alpha_2\leq k_2 \\ \cdots \\ 0\leq\alpha_i\leq k_i \\ \sum_{l=1}^i \alpha_l + i \leq m}} \frac{t^{\sum_{l=1}^i \alpha_l + i}}{\alpha!} I(\alpha) \prod_{l=1}^i \Big(\underbrace{[\mathscr{L}_0^z,[\mathscr{L}_0^z,\ldots,[\mathscr{L}_0^z,\mathscr{L}_{k_l}^z]]\ldots]]}_{\alpha_l\text{-times}}\Big) p_t^{\bar{X}^z}(x,y)|_{z=x}. \end{aligned} \tag{3.3}$$

Proof. See Appendix F. ☐

We show the representation of $p_t^{\bar{X},(m)}(x,y)$. First, we note that the differential operator appearing in (3.3) is simply written in the following form:

$$\prod_{l=1}^i \big(\underbrace{[\mathscr{L}_0^z,[\mathscr{L}_0^z,\ldots,[\mathscr{L}_0^z,\mathscr{L}_{k_l}^z]]\ldots]]}_{\alpha_l\text{-times}}\big)|_{z=x} = \sum_{|\beta|\leq\nu} g_\beta(x)\partial^\beta, \tag{3.4}$$

where $\nu \in \mathbb{N}$, $\beta \in \{1,2,\ldots,N\}^l$, $l \leq \nu$ and each $g_\beta \in C_b^\infty(\mathbb{R}^N)$ is a linear combination of b, σ and their partial derivatives. Then, we have

$$\prod_{l=1}^i \big(\underbrace{[\mathscr{L}_0^z,[\mathscr{L}_0^z,\ldots,[\mathscr{L}_0^z,\mathscr{L}_{k_l}^z]]\ldots]]}_{\alpha_l\text{-times}}\big) p_t^{\bar{X}^z}(x,y)|_{z=x} = \sum_{|\beta|\leq\nu} g_\beta(x)\partial^\beta p_t^{\bar{X}^z}(x,y)|_{z=x}. \tag{3.5}$$

For a multi-index $\beta = (\beta_1, \ldots, \beta_l) \in \{1, \ldots, N\}^l$, $l \in \mathbb{N}$, the derivative $\partial^\beta p_t^{\bar{X}^z}(x, y)$ is explained as

$$\partial^\beta p_t^{\bar{X}^z}(x, y) = (-1)^l \times H_{\Sigma(z)t}^{\beta_1,\ldots,\beta_l}(y - x - b(z)t) p_t^{\bar{X}^z}(x, y) \tag{3.6}$$

using the l-th order multivariate Hermite polynomial $H_a^{\beta_1,\ldots,\beta_{i-1},\beta_i}(\xi)$ given by

$$H_a^{\beta_1,\ldots,\beta_{i-1},\beta_i}(\xi) = -\left([a^{-1}\xi]_{\beta_i} - \partial_{\xi_{\beta_i}}\right) H_a^{\beta_1,\ldots,\beta_{i-1}}(\xi), \qquad i = 1, \ldots, l, \tag{3.7}$$

for an invertible matrix $a \in \mathbb{R}^{N\times N}$ and a vector $\xi \in \mathbb{R}^N$, with $H_a^{\beta_1}(\xi) = -[a^{-1}\xi]_{\beta_1}$. Let $\vartheta_\beta : (0, \infty) \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ be a function given by

$$\vartheta_\beta(t, x, y) := (-1)^l \times H_{\Sigma(z)t}^{\beta_1,\ldots,\beta_l}(y - x - b(z)t)\Big|_{z=x}, \tag{3.8}$$

then $\partial^\beta p_t^{\bar{X}^z}(x, y)|_{z=x}$ has the form

$$\partial^\beta p_t^{\bar{X}^z}(x, y)|_{z=x} = \vartheta_\beta(t, x, y)\ p_t^{\bar{X}}(x, y). \tag{3.9}$$

Therefore, $p_t^{\bar{X},(m)}(x, y)$ is transformed into the following formula:

$$p_t^{\bar{X},(m)}(x, y) = \widetilde{\mathcal{M}}^{(m)}(t, x, y)\ p_t^{\bar{X}}(x, y), \quad (t, x, y) \in (0, \infty) \times \mathbb{R}^N \times \mathbb{R}^N, \tag{3.10}$$

where the weight on finite dimension $\widetilde{\mathcal{M}}^{(m)} : (0, \infty) \times \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ is given by

$$\widetilde{\mathcal{M}}^{(m)}(t, x, y) = 1 + \sum_{i=1}^{m-1} \sum_{i \le \sum_{j=1}^i k_j \le 2m+1} \sum_{\substack{1\le\alpha_1\le k_1 \\ 0\le\alpha_2\le k_2 \\ \cdots \\ 0\le\alpha_i\le k_i \\ \sum_{l=1}^i \alpha_l + i \le m}} \sum_{|\beta|\le\nu} \frac{t^{\sum_{l=1}^i \alpha_l + i}}{\alpha!} I(\alpha) g_\beta(x) \vartheta_\beta(t, x, y). \tag{3.11}$$

Based on the above discussion and Theorem 3.1, we derive the second order discretization of heat kernel $p_T^X(x, y)$ as follows:

Corollary 3.1 (Second order discretization of heat kernel)**.** *Assume $T \ge 1$. There exist some constants $C, c > 0$, $q \ge \frac{N}{2}$ independent of T and non-decreasing function $K(\cdot)$ such that*

$$\left| p_T^X(x, y) - E\left[p_{T/n}^{\bar{X},(2)}\left(\bar{X}^{(n)}((n-1)T/n, x), y\right) \prod_{j=1}^{n-1} \left\{ 1 + \pi_{T/n}^{(2),\bar{X}^{(n)}((j-1)T/n,x)}\left(B_{jT/n} - B_{(j-1)T/n}\right)\right\}\right]\right|$$
$$\le \frac{K(T)}{n^2} \frac{C}{T^q} \exp\left(-c\frac{|y - x|^2}{T}\right),$$

for all $x, y \in \mathbb{R}^N$, where $\pi_t^{(2),x}(B_t), (t, x) \in (0, \infty) \times \mathbb{R}^N$ is given in Corollary 2.2 and $p_t^{\bar{X},(2)}(x, y)$, $(t, x, y) \in (0, \infty) \times \mathbb{R}^N \times \mathbb{R}^N$ is given by

$$p_t^{\bar{X},(2)}(x, y) = p_t^{\bar{X}}(x, y) + \frac{t^2}{2} \sum_{i=1}^2 [\mathscr{L}_0^z, \mathscr{L}_i^z] p_t^{\bar{X}^z}(x, y)|_{z=x} = \widetilde{\mathcal{M}}^{(2)}(t, x, y)\ p_t^{\bar{X}}(x, y) \tag{3.12}$$

with the weight on finite dimension $\widetilde{\mathcal{M}}^{(2)}(t,x,y)$ *given by*

$$\widetilde{\mathcal{M}}^{(2)}(t,x,y) = 1 + \frac{t^2}{2}\sum_{j=1}^{N}\mathcal{L}_0 b^j(x)\vartheta_{(j_1)}(t,x,y) + \frac{t^2}{2}\sum_{k=1}^{d}\sum_{j_1,j_2=1}^{N}\mathcal{L}_k b^{j_1}(x)\sigma_k^{j_2}(x)\vartheta_{(j_1,j_2)}(t,x,y) \tag{3.13}$$

$$+ \frac{t^2}{2}\sum_{k=1}^{d}\sum_{j_1,j_2=1}^{N}\mathcal{L}_0\sigma_k^{j_1}(x)\sigma_k^{j_2}(x)\vartheta_{(j_1,j_2)}(t,x,y) \tag{3.14}$$

$$+ \frac{t^2}{2}\sum_{k_1,k_2=1}^{d}\sum_{j_1,j_2,j_3=1}^{N}\mathcal{L}_{k_1}\sigma_{k_2}^{j_1}(x)\sigma_{k_1}^{j_2}(x)\sigma_{k_2}^{j_3}(x)\vartheta_{(j_1,j_2,j_3)}(t,x,y) \tag{3.15}$$

$$+ \frac{t^2}{4}\sum_{k_1,k_2=1}^{d}\sum_{j_1,j_2=1}^{N}\mathcal{L}_{k_1}\sigma_{k_2}^{j_1}(x)\mathcal{L}_{k_1}\sigma_{k_2}^{j_2}(x)\vartheta_{(j_1,j_2)}(t,x,y) \tag{3.16}$$

with $\vartheta_{(j_1)}(t,x,y) = \frac{1}{t}\sum_{k=1}^{N}A_k^{j_1}(x)(y_k - x_k - b^k(x)t)$,

$$\vartheta_{(j_1,j_2)}(t,x,y) = \frac{1}{t^2}\sum_{i_1,i_2=1}^{N}A_{i_1}^{j_1}(x)A_{i_2}^{j_2}(x)\left(y_{i_1} - x_{i_1} - b^{i_1}(x)t\right)\left(y_{i_2} - x_{i_2} - b^{i_2}(x)t\right) - \frac{1}{t}A_{j_2}^{j_1}(x)$$

and

$$\begin{aligned}\vartheta_{(j_1,j_2,j_3)}(t,x,y) = {} & \frac{1}{t^3}\sum_{i_1,i_2,i_3=1}^{N}A_{i_1}^{j_1}(x)A_{i_2}^{j_2}(x)A_{i_3}^{j_3}(x)\left(y_{i_1} - x_{i_1} - b^{i_1}(x)t\right)\left(y_{i_2} - x_{i_2} - b^{i_2}(x)t\right)\left(y_{i_3} - x_{i_3} - b^{i_3}(x)t\right)\\ & - \frac{A_{j_2}^{j_1}(x)}{t^2}\sum_{i=1}^{N}A_i^{j_3}(x)\left(y_i - x_i - b^i(x)t\right) - \frac{A_{j_3}^{j_1}(x)}{t^2}\sum_{i=1}^{N}A_i^{j_2}(x)\left(y_i - x_i - b^i(x)t\right)\\ & - \frac{A_{j_3}^{j_2}(x)}{t^2}\sum_{i=1}^{N}A_i^{j_1}(x)\left(y_i - x_i - b^i(x)t\right).\end{aligned}$$

Proof. By Theorem 3.1, we obtain

$$p_t^{\bar{X},(2)}(x,y) = p_t^{\bar{X}}(x,y) + \sum_{i=1}^{5}\frac{t^2}{2}[\mathscr{L}_0^z,\mathscr{L}_i^z]p_t^{\bar{X}^z}(x,y)|_{z=x} = p_t^{\bar{X}}(x,y) + \frac{t^2}{2}\sum_{i=1}^{2}[\mathscr{L}_0^z,\mathscr{L}_i^z]p_t^{\bar{X}^z}(x,y)|_{z=x}, \tag{3.17}$$

where on the second and third equations we used $[\mathscr{L}_0^z,\mathscr{L}_i^z]p_t^{\bar{X}^z}(x,y)|_{z=x} = 0$ for $i = 3,4,5$ which is due to the result that $[\mathscr{L}_0^z,\mathscr{L}_i^z]p_t^{\bar{X}^z}(x,y)|_{z=x}$ has the form $\psi(z)(x-z)^{\ell}p_t^{\bar{X}^z}(x,y)|_{z=x} = 0$ for some $\psi \in C_b^{\infty}(\mathbb{R}^N)$ and $\ell \in \mathbb{N}$. The term $\frac{t^2}{2}\sum_{i=1}^{2}[\mathscr{L}_0^z,\mathscr{L}_i^z]p_t^{\bar{X}^z}(x,y)|_{z=x}$ is calculated in the following way. Since we have

$$\begin{aligned}[\mathscr{L}_0^z,\mathscr{L}_1^z]p_t^{\bar{X}^z}(x,y)|_{z=x} = {} & \sum_{i,l=1}^{N}b^l(x)\partial_l b^i(x)\frac{\partial}{\partial x_i}p_t^{\bar{X}^z}(x,y)|_{z=x}\\ & + \sum_{k=1}^{d}\sum_{i,j,l=1}^{N}\left\{b^l(x)\partial_l\sigma_k^i(x)\sigma_k^j(x)\frac{\partial^2}{\partial x_i\partial x_j}p_t^{\bar{X}^z}(x,y)|_{z=x}\right.\\ & \left. + \sigma_k^l(x)\sigma_k^j(x)\partial_l b^i(x)\frac{\partial^2}{\partial x_i\partial x_j}p_t^{\bar{X}^z}(x,y)|_{z=x}\right\}\\ & + \sum_{k_1,k_2=1}^{d}\sum_{i_1,i_2,j,l=1}^{N}\sigma_{k_1}^l(x)\sigma_{k_1}^j(x)\partial_l\sigma_{k_2}^{i_1}(x)\sigma_{k_2}^{i_2}(x)\frac{\partial^3}{\partial x_j\partial x_{i_1}\partial x_{i_2}}p_t^{\bar{X}^z}(x,y)|_{z=x}\end{aligned}$$

and

$$[\mathscr{L}_0^z,\mathscr{L}_2^z]p_t^{\bar{X}^z}(x,y)|_{z=x} = \frac{1}{2}\sum_{k=1}^{d}\sum_{i,l_1.l_2=1}^{N}\sigma_k^{l_1}(x)\sigma_k^{l_2}(x)\partial_{l_1}\partial_{l_2}b^i(x)\frac{\partial}{\partial x_i}p_t^{\bar{X}^z}(x,y)|_{z=x}$$
$$+\frac{1}{2}\sum_{k_1,k_2=1}^{d}\sum_{i,j,l_1,l_2=1}^{N}\sigma_{k_1}^{l_1}(x)\sigma_{k_1}^{l_2}(x)\left\{\partial_{l_1}\partial_{l_2}\sigma_{k_2}^i(x)\sigma_{k_2}^j(x)+\partial_{l_1}\sigma_{k_2}^i(x)\partial_{l_2}\sigma_{k_2}^j(x)\right\}$$
$$\times\frac{\partial^2}{\partial x_i\partial x_j}p_t^{\bar{X}^z}(x,y)|_{z=x},$$

it holds that

$$\frac{t^2}{2}\sum_{i=1}^{2}[\mathscr{L}_0^z,\mathscr{L}_i^z]p_t^{\bar{X}^z}(x,y)|_{z=x} = \frac{t^2}{2}\sum_{j=1}^{N}\mathcal{L}_0 b^j(x)\frac{\partial}{\partial x_j}p_t^{\bar{X}^z}(x,y)|_{z=x}$$
$$+\frac{t^2}{2}\sum_{k=1}^{d}\sum_{j_1,j_2=1}^{N}\mathcal{L}_k b^{j_1}(x)\sigma_k^{j_2}(x)\frac{\partial^2}{\partial x_{j_1}\partial x_{j_2}}p_t^{\bar{X}^z}(x,y)|_{z=x}$$
$$+\frac{t^2}{2}\sum_{k=1}^{d}\sum_{j_1,j_2=1}^{N}\mathcal{L}_0\sigma_k^{j_1}(x)\sigma_k^{j_2}(x)\frac{\partial^2}{\partial x_{j_1}\partial x_{j_2}}p_t^{\bar{X}^z}(x,y)|_{z=x}$$
$$+\frac{t^2}{2}\sum_{k_1,k_2=1}^{d}\sum_{j_1,j_2,j_3=1}^{N}\mathcal{L}_{k_1}\sigma_{k_2}^{j_1}(x)\sigma_{k_1}^{j_2}(x)\sigma_{k_2}^{j_3}(x)\frac{\partial^3}{\partial x_{j_1}\partial x_{j_2}\partial x_{j_3}}p_t^{\bar{X}^z}(x,y)|_{z=x}$$
$$+\frac{t^2}{4}\sum_{k_1,k_2=1}^{d}\sum_{j_1,j_2=1}^{N}\mathcal{L}_{k_1}\sigma_{k_2}^{j_1}(x)\mathcal{L}_{k_1}\sigma_{k_2}^{j_2}(x)\frac{\partial^2}{\partial x_{j_1}\partial x_{j_2}}p_t^{\bar{X}^z}(x,y)|_{z=x}.$$

By the representation (3.9) with (3.8), we have the assertion. □

4. Numerical study of high order discretization of heat kernel

In this section, we show some numerical results of discretizing heat kernels for univariate and multivariate models using the proposed high order scheme in order to verify our assertion. We compare those with the results computed by the classical Euler–Maruyama method, the first order scheme. Here, the first order discretization is constructed as follows. Let $T \geq 1$ and $X(T,x)$ be the solution of SDE (2.6). We denote the density of $X(T,x)$ by $p_T^X(x,y),\ (x,y)\in\mathbb{R}^N\times\mathbb{R}^N$. In addition, the density of the Euler–Maruyama scheme $\bar{X}^{(n)}(T,x)$ is denoted by $p_T^{\bar{X}^{(n)}}(x,y)$. Then, $p_T^X(x,y)$ is approximated by the following convolution.

$$p_T^X(x,y)\approx p_T^{\bar{X}^{(n)}}(x,y)=\underbrace{\left(p_{T/n}^{\bar{X}}*\dots*p_{T/n}^{\bar{X}}\right)}_{n}(x,y)=E\left[p_{T/n}^{\bar{X}}\left(\bar{X}^{(n)}((n-1)T/n,x),y\right)\right], \quad (4.1)$$

where $p_t^{\bar{X}}(x,y),\ (t,x,y)\in(0,1]\times\mathbb{R}^N\times\mathbb{R}^N$ is the density of one step Euler–Maruyama scheme given by (3.1) with $z=x$. On the other hand, based on Theorem 3.1 using Malliavin calculus, we approximate $p_T^X(x,y)$ by m-th order scheme ($m\geq 2$)

$$p_T^X(x,y)\approx\underbrace{\left(p_{T/n}^{\bar{X},(m)}*\dots*p_{T/n}^{\bar{X},(m)}\right)}_{n}(x,y) \quad (4.2)$$
$$=E\left[p_{T/n}^{\bar{X},(m)}\left(\bar{X}^{(n)}((n-1)T/n,x),y\right)\prod_{j=1}^{n-1}\left\{1+\pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)T/n,x)}(B_{jT/n}-B_{(j-1)T/n})\right\}\right].$$

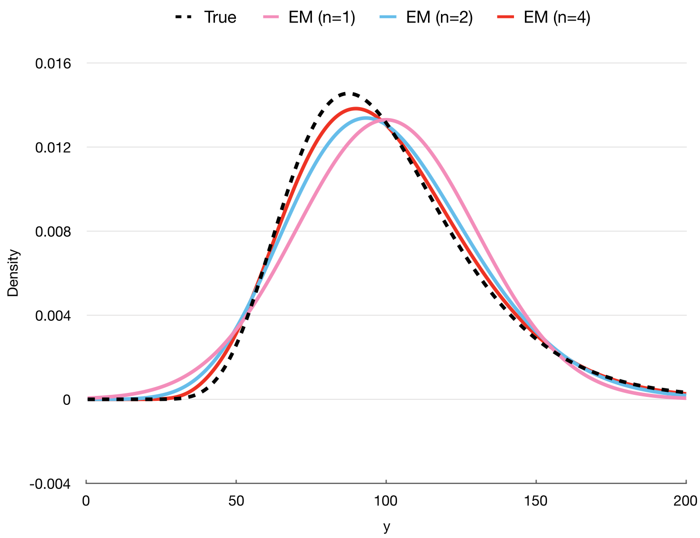

FIGURE 1. Kernel estimation (Euler–Maruyama scheme).

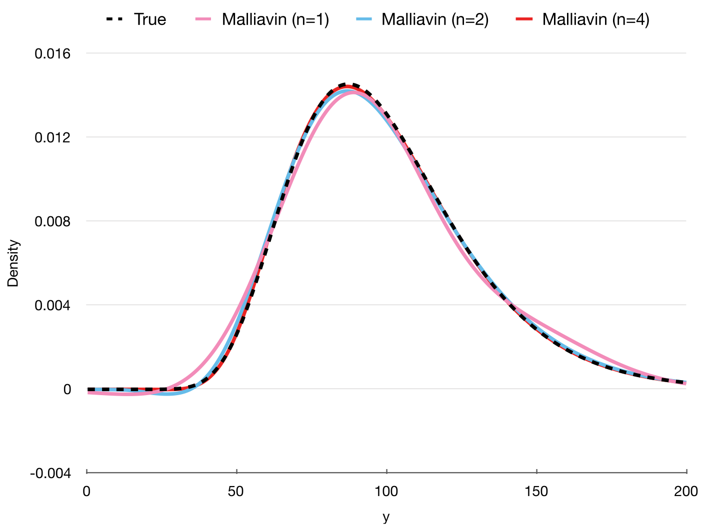

FIGURE 2. Kernel estimation (Second order scheme).

Note that for $n = 1$, we always have exact values for both approximations since $p^{\bar{X}}_{T/n}(x, y)$ and $p^{\bar{X},(m)}_{T/n}(x, y)$ are obtained in closed form. In the following numerical examples, we compute the approximation (4.1) and (4.2) with $m = 2$ by simulation. For the figures, the labels (EM) and (Malliavin) are used for (4.1) and (4.2).

4.1. Univariate model

Let us consider the following 1-dimensional SDE:

$$dX^x_t = \sigma X^x_t dW_t, \ X^x_0 = x \in \mathbb{R}, \tag{4.3}$$

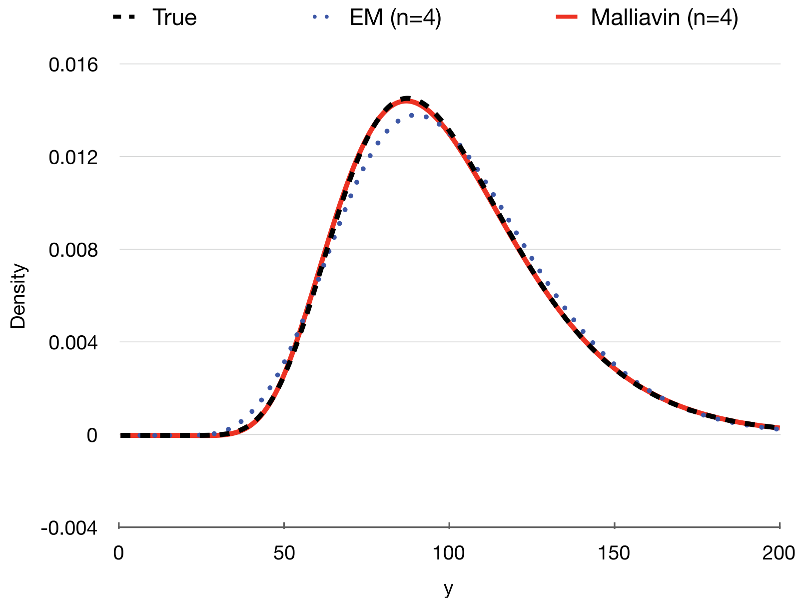

FIGURE 3. Comparison of approximated densities with time steps $n = 4$.

where $\sigma > 0$. We numerically compute the density of X_T^x, $T \geq 1$, $x \in \mathbb{R}$, using the Euler–Maruyama method and our proposed scheme ($m = 2$, *i.e.* second order scheme). Here, we apply the Quasi-Monte-Carlo (QMC) method to both schemes (4.1) and (4.2) with number of simulations $M = 10^4$ for each time step $n = 1, 2$ and 4. The parameters are given by $T = 1$, $x = 100$ and $\sigma = 0.3$. In this example, the exact density is obtained in closed form since the probability law of X_t^x is lognormal. We use the exact density as the benchmark in the following numerical tests.

In Figures 1 and 2, the true and the approximated densities are illustrated. Figure 1 shows the comparison results for the Euler–Maruyama scheme and Figure 2 shows the estimated kernels with the proposed second order scheme. Through the experiments, we particularly observe that the convergence of the second order scheme is much faster than that of the Euler–Maruyama scheme. The approximated density of the second order scheme almost corresponds with the true density when $n = 4$, which can be checked in the following Figure 3.

4.2. Multivariate model

We consider the following 2-dimensional SDE:

$$dX_t^{1,x} = X_t^{2,x}(X_t^{1,x})^\beta dW_t^1, \ X_0^{1,x} = x_1 > 0 \tag{4.4}$$

$$dX_t^{2,x} = \nu\rho X_t^{2,x} dW_t^1 + \nu\sqrt{1-\rho^2} X_t^{2,x} dW_t^2, \ X_0^{2,x} = x_2 > 0, \tag{4.5}$$

where $\beta \in [0, 1]$, $\nu \geq 0$ and $\rho \in (-1, 1)$. In financial mathematics, the process $\{X_t^{1,x}, X_t^{2,x}\}_{t\geq 0}$ is known as the SABR model which is one of important classes of stochastic volatility model. Here, the first and second element of the process means underlying asset and its volatility, respectively. In this section, we especially investigate the marginal density of $X_T^{1,x}$ since estimating it is important in practice. As in the previous section, we apply QMC method to the Euler–Maruyama scheme and the proposed second order scheme ($m = 2$) with number of simulations $M = 10^5$ for each time step $n = 1, 2$ and 4. Since we do not have the analytical solution of the density of $X_T^{1,x}$, we numerically compute the benchmark value by the first order scheme (4.1) with the number of simulations $M = 10^6$ and the time step $n = 2^7$. We set $T = 1, x_1 = 100, x_2 = 0.3, \beta = 1, \nu = 0.1$ and $\rho = -0.5$.

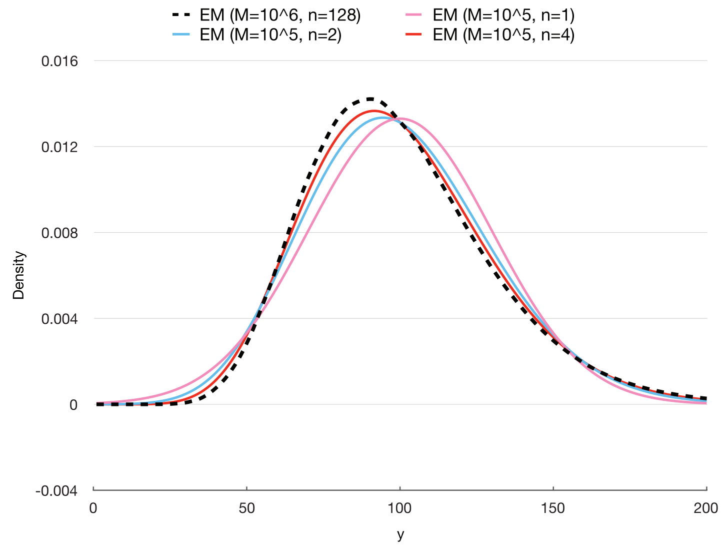

FIGURE 4. Kernel estimation (Euler–Maruyama scheme).

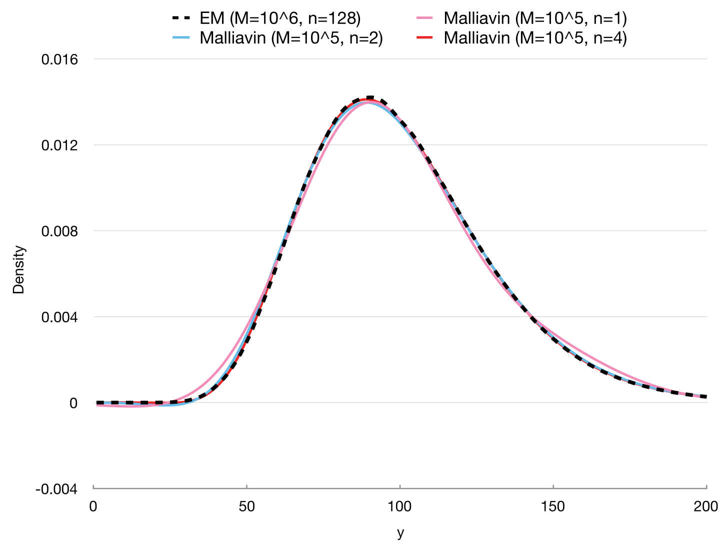

FIGURE 5. Kernel estimation (Second order scheme).

In Figures 4 and 5, the benchmark and the approximated densities are illustrated. Here, the benchmark density is represented by the dashed line. As in the previous section, Figures 4 and 5 show the comparison results for the Euler–Maruyama scheme and the proposed second order scheme, respectively. From these figures, it is also clear that the proposed scheme achieves faster convergence, compared to the Euler–Maruyama scheme. In particular, we observe from Figure 6 that the approximated density computed by the new scheme with $n = 4$ is almost the same with the benchmark density. Through these numerical studies, one can check that the new scheme has a superior efficiency compared to the classical Euler–Maruyama scheme.

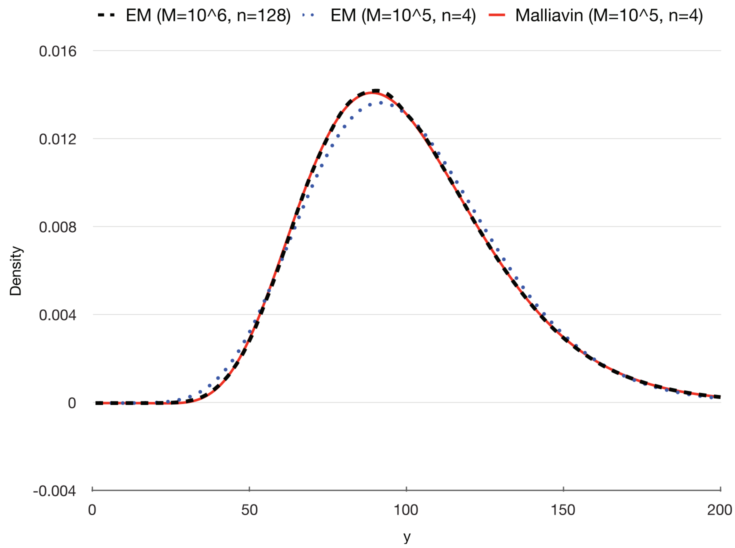

FIGURE 6. Comparison of approximated densities with time steps $n = 4$.

Remark 4.1. We give some remarks on numerical results for the density of the SABR model (including that of the Black–Scholes model). We consider the following system equivalent to the SABR model:

$$\begin{aligned} dX_t^1 &= \varepsilon X_t^2 (X_t^1)^\beta dW_t^1, \quad X_0^1 = x_0 > 0, \\ dX_t^2 &= \nu X_t^2 (\rho dW_t^1 + \sqrt{1-\rho^2} dW_t^2), \quad X_0^2 = 1, \end{aligned}$$

where $\varepsilon = x_2$ is the initial volatility. When $\beta = 1$ and $\nu = 0$, X^1 becomes the Black–Scholes model. For $\nu > 0$, the process X^2 takes positive values. We assume the condition such that at least X_T^1 has the density $p_T^{X^1}(x_0, \cdot)$ with the ellipticity at the starting point x_0. However, the model does not satisfy the uniformly elliptic condition in general and the coefficients are not sufficiently smooth. We approximate $p_T^{X^1}(x_0, \cdot)$ by the density of the following modified SABR model which satisfies the uniformly elliptic condition with smooth coefficients:

$$\begin{aligned} d\widetilde{X}_t^1 &= \varepsilon \psi_1(\widetilde{X}_t^2) \psi_2(\widetilde{X}_t^1) dW_t, \quad \widetilde{X}_0^1 = x_0 > 0, \\ d\widetilde{X}_t^2 &= \nu \psi_3(\widetilde{X}_t^2)(\rho dW_t^1 + \sqrt{1-\rho^2} dW_t^2), \quad \widetilde{X}_0^2 = 1. \end{aligned}$$

where $\psi_1(x) = h(x)x$, $\psi_2(x) = h(x)x^\beta$, $\psi_3(x) = h(x)x$ with

$$\begin{aligned} h(x) &= \frac{\gamma(x-a_2)}{\gamma(x-a_2) + \gamma(a_1 - x)}, \quad 0 < a_2 < a_1 < a := \frac{1}{2}x_0, \\ \gamma(x) &= e^{-1/x} \mathbf{1}_{x>0}. \end{aligned}$$

Note that the densities $p_T^{X^1}(x, \cdot)$ and $\widetilde{p}_T^{\widetilde{X}^1}(x, \cdot)$ have the following representations under their existence conditions:

$$p_T^{X^1}(x,y) = E\left[\mathbf{1}_{\{X_T^1 \geq y\}} \delta\left(DX_T^1 \|DX_T^1\|_H^{-2}\right)\right] \quad \text{and} \quad \widetilde{p}_T^{\widetilde{X}^1}(x,y) = E\left[\mathbf{1}_{\{\widetilde{X}_T^1 \geq y\}} \delta\left(D\widetilde{X}_T^1 \|D\widetilde{X}_T^1\|_H^{-2}\right)\right],$$

where $\delta(DX_T^1\|DX_T^1\|_H^{-2})\in L^2(\Omega)$ and $\delta(D\widetilde{X}_T^1\|D\widetilde{X}_T^1\|_H^{-2})\in L^2(\Omega)$. Then the difference between $p_T^{X^1}(x,y)$ and $\widetilde{p}_T^{\widetilde{X}^1}(x,y)$ is given by

$$\begin{aligned}
&|p_T^{X^1}(x,y)-\widetilde{p}_T^{\widetilde{X}^1}(x,y)|\\
&\quad=\left|E[\mathbf{1}_{\{X_T^1\geq y\}}\delta(DX_T^1\|DX_T^1\|_H^{-2})]-E[\mathbf{1}_{\{\widetilde{X}_T^1\geq y\}}\delta(D\widetilde{X}_T^1\|D\widetilde{X}_T^1\|_H^{-2})]\right|\\
&\quad\leq E\left[\left|\mathbf{1}_{\{X_T^1\geq y\}}\delta(DX_T^1\|DX_T^1\|_H^{-2})-\mathbf{1}_{\{\widetilde{X}_T^1\geq y\}}\delta(D\widetilde{X}_T^1\|D\widetilde{X}_T^1\|_H^{-2})\right|\right]\\
&\quad= E\left[\left|\mathbf{1}_{\{X_T^1\geq y\}}\delta(DX_T^1\|DX_T^1\|_H^{-2})-\mathbf{1}_{\{\widetilde{X}_T^1\geq y\}}\delta(D\widetilde{X}_T^1\|D\widetilde{X}_T^1\|_H^{-2})\right|\mathbf{1}_{\{X_\cdot^1\neq\widetilde{X}_\cdot^1\}}\right]\\
&\quad\leq C\{\|\delta(DX_T^1\|DX_T^1\|_H^{-2})\|_2+\|\delta(D\widetilde{X}_T^1\|D\widetilde{X}_T^1\|_H^{-2})\|_2\}P\left(\{\omega;\ X_\cdot^1(\omega)\neq\widetilde{X}_\cdot^1(\omega)\}\right)^{1/2}\\
&\quad\leq C\{\|\delta(DX_T^1\|DX_T^1\|_H^{-2})\|_2+\|\delta(D\widetilde{X}_T^1\|D\widetilde{X}_T^1\|_H^{-2})\|_2\}P\left(\{\omega;\ X_t^1(\omega)\geq a_1 \text{ for some } t\in[0,T]\}\right)^{1/2},
\end{aligned}$$

for some $C>0$. Since we have $\|\delta(DX_T^1\|DX_T^1\|_H^{-2})\|_2<\infty$ and $\|\delta(D\widetilde{X}_T^1\|D\widetilde{X}_T^1\|_H^{-2})\|_2<\infty$, furthermore it holds that

$$\begin{aligned}
&P\left(\{\omega;\ X_t^1(\omega)\geq a_1 \text{ for some } t\in[0,T]\}\right)\\
&\quad\leq P\left(\{\omega;\ \sup_t|X_t^1(\omega)-x|>a\}\right)+P\left(\{\omega;\ X_t^1(\omega)\geq a' \text{ for some } t\in[0,T]\}\cap\{\omega;\ \sup_t|X_t^1(\omega)-x|\leq a\}\right)\\
&\quad=P\left(\{\omega;\ \sup_t|X_t^1(\omega)-x|>a\}\right)\ =\ O(\varepsilon^k),\quad \text{for all } k=1,2,\dots,
\end{aligned}$$

where the large deviation estimate for small noise SDE ([20], Lem. 4) is applied in the above, then the difference $|p_T^{X^1}(x,y)-\widetilde{p}_T^{\widetilde{X}^1}(x,y)|$ is negligible, which implies that $p_T^{X^1}(x,\cdot)$ can be approximated by the proposed second order discretization.

5. Concluding remarks

In the paper, we showed a higher order operator splitting scheme for diffusion semigroups using the Baker–Campbell–Hausdorff expansion around the Euler–Maruyama scheme and combined with Malliavin calculus. The heat kernel approximation was given with the new algorithm as the extension of Bally and Talay [4] and Pedersen [19]. We illustrated the numerical experiments for the scheme and the effectiveness was checked.

Although the uniformly elliptic condition is assumed in the paper, we believe that the proposed scheme can work in weaker conditions. In other words, a higher order scheme for the density of hypoelliptic diffusions will be constructed and we are able to prove the conjecture in [7] by different approach. Furthermore, construction of a higher order scheme for density of hypoelliptic diffusions will lead to various applications such as parametric inference as discussed in [16]. The higher order discretization of hypoelliptic heat kernels should be developed as future work.

Appendix A. Proof of Proposition 2.1

Let

$$U(s,x):=e^{s\mathscr{L}_0^z}\widehat{\mathscr{L}_i}e^{(t-s)\mathscr{L}_0^z}\varphi(x)=\int_{\mathbb{R}^N}\widehat{\mathscr{L}_i}\left(\int_{\mathbb{R}^N}\varphi(y)p^{\bar{X}^z}(t-s,\xi,y)\,\mathrm{d}y\right)p^{\bar{X}^z}(s,x,\xi)\,\mathrm{d}\xi.$$

Then, we have

$$U(s,x)=U(0,x)+\sum_{k=1}^m\frac{s^k}{k!}\frac{\partial^k}{\partial s^k}U(s,x)\Big|_{s=0}+\int_0^s\frac{(s-\xi)^m}{m!}\frac{\partial^{m+1}}{\partial\xi^{m+1}}U(\xi,x)\,\mathrm{d}\xi.$$

Note that

$$\frac{\partial}{\partial s}U(s,x) = e^{s\mathscr{L}_0^z}\mathscr{L}_0^z\widehat{\mathscr{L}_i}e^{(t-s)\mathscr{L}_0^z}\varphi(x) - e^{s\mathscr{L}_0^z}\widehat{\mathscr{L}_i}\mathscr{L}_0^z e^{(t-s)\mathscr{L}_0^z}\varphi(x)$$
$$= e^{s\mathscr{L}_0^z}[\mathscr{L}_0^z,\widehat{\mathscr{L}_i}]e^{(t-s)\mathscr{L}_0^z}\varphi(x),$$

and

$$\frac{\partial}{\partial s}U(s,x)\Big|_{s=0} = [\mathscr{L}_0^z,\widehat{\mathscr{L}_i}]e^{t\mathscr{L}_0^z}\varphi(x).$$

Furthermore, it holds that for $k \in \mathbb{N}$,

$$\frac{\partial^k}{\partial s^k}U(s,x) = e^{s\mathscr{L}_0^z}\underbrace{\Big[\mathscr{L}_0^z,\Big[\dots\Big[\mathscr{L}_0^z,\Big[\mathscr{L}_0^z,\widehat{\mathscr{L}_i}\Big]\Big]\dots\Big]\Big]}_{k\text{-times}}e^{(t-s)\mathscr{L}_0^z}\varphi(x),$$

and

$$\frac{\partial^k}{\partial s^k}U(s,x)\Big|_{s=0} = \underbrace{\Big[\mathscr{L}_0^z,\Big[\dots\Big[\mathscr{L}_0^z,\Big[\mathscr{L}_0^z,\widehat{\mathscr{L}_i}\Big]\Big]\dots\Big]\Big]}_{k\text{-times}}e^{t\mathscr{L}_0^z}\varphi(x).$$

We prove the explicit truncation in the formula (2.45). Let us define by $\mathcal{D}_i \subset \mathcal{DO}$ $(i = 0,1,2,\dots)$ a set of differential operators acting on $\varphi(x), \varphi \in C^\infty(\mathbb{R}^N)$ with at most i-th degree polynomial of x in the coefficients. Then, it is obvious that $\mathcal{D}_0$ is commutative sub-algebra, that is, for any $X, Y \in \mathcal{D}_0$, $[X,Y] = 0$. Moreover, for all $X \in \mathcal{D}_0$, $Y \in \mathcal{D}_i$, the commutator $[X,Y]$ is an element of $\mathcal{D}_{i-1}$. Since $\mathscr{L}_0^z \in \mathcal{D}_0$ and $\widehat{\mathscr{L}_i} \in \mathcal{D}_i$, we have $[\mathscr{L}_0^z,\widehat{\mathscr{L}_i}] \in \mathcal{D}_{i-1}$ and then,

$$\underbrace{\Big[\mathscr{L}_0^z,\Big[\mathscr{L}_0^z,\dots,\Big[\mathscr{L}_0^z,\widehat{\mathscr{L}_i}\Big]\Big]\dots\Big]}_{i\text{-times}} \in \mathcal{D}_0. \tag{A.1}$$

By the commutativity of $\mathcal{D}_0$, we have for every $i \in \mathbb{N}$,

$$\underbrace{\Big[\mathscr{L}_0^z,\Big[\mathscr{L}_0^z,\dots,\Big[\mathscr{L}_0^z,\widehat{\mathscr{L}_i}\Big]\Big]\dots\Big]}_{n\text{-times}} = 0 \tag{A.2}$$

for any integers $n \geq i+1$. Then the assertions are obtained. □

Appendix B. Proof of Proposition 2.2

First, we introduce the result on $P_s^{i,z}\varphi(x)$ which plays an important role in the proof of the proposition. We recall that for $i \in \mathbb{N}$, $P_s^{i,z}\varphi(x)$ with a bounded and measurable function $\varphi : \mathbb{R}^N \to \mathbb{R}$ is defined as

$$P_s^{i,z}\varphi(x) = \sum_{k=0}^{i-1}\int_0^s P_{s-u}^{0,z}\mathscr{L}_{i-k}^z P_u^{k,z}\varphi(x)\,\mathrm{d}u, \quad s \in (0,1],\ x,z \in \mathbb{R}^N. \tag{B.1}$$

Lemma B.1. *For $i \in \mathbb{N}$, each term of $P_s^{i,z}\varphi(x)$, $s \in (0,1]$, $x, z \in \mathbb{R}^N$ is given in the form*

$$s^l\psi(z)\prod_{j=1}^{p}(x_{l_j} - z_{l_j})\partial^\gamma(P_s^{0,z}\varphi)(x), \tag{B.2}$$

where $l \geq 1, p \geq 0$ *and multi-index* $\gamma \in \{1, 2, \dots, N\}^{|\gamma|}$ *satisfy*

$$2l + p - |\gamma| \geq i, \tag{B.3}$$

and with $\psi \in C_b^{\infty}(\mathbb{R}^N)$, $l_j = 1, \dots, N$, $j = 1, \dots, p$.

Proof. We show this by induction with respect to the integer i. When $i = 1$, it is easy to check that the assertion holds. Then, we assume that for all $k = 1, 2, \dots, n$, $n \in \mathbb{N}$, each term of $P_s^{k,z}\varphi(x)$ is given by

$$s^l \psi(z) \prod_{j=1}^{p} (x_{l_j} - z_{l_j}) \partial^{\gamma} (P_s^{0,z}\varphi)(x), \tag{B.4}$$

with $\psi \in C_b^{\infty}(\mathbb{R}^N)$ and $l, p \in \mathbb{N}$ and a multi-index γ such that $2l + p - |\gamma| \geq k$. From now on, we will show the assertion holds for $k = n + 1$. Because of the above assumption, terms in $P_s^{n+1,z}\varphi(x)$ are given by

$$\int_0^s P_{s-u}^{0,z} \left\{ u^l \psi(z) \mathscr{L}_{n+1-k}^{z} \prod_{j=1}^{p} (x_{l_j} - z_{l_j}) \partial^{\gamma} \right\} P_u^{0,z}\varphi(x)\, \mathrm{d}u. \tag{B.5}$$

Here, we note that $\mathscr{L}_{n+1-k}^{z} \prod_{j=1}^{p} (x_{l_j} - z_{l_j}) \partial^{\gamma} (P_u^{0,z}\varphi)(x)$ has the form of

$$\begin{aligned} &\sum_{r=0}^{2} \chi_r(z) \prod_{d=1}^{n+1-k} (x_{\iota_d} - z_{\iota_d}) \prod_{e=1}^{p-r} (x_{\iota_e} - z_{\iota_e}) \partial^{\beta^r} \partial^{\gamma} (P_u^{0,z}\varphi)(x) \\ &\qquad =: \sum_{r=0}^{2} \overline{\mathscr{L}}_{n+1-k+p-r}^{z,(\gamma)} (P_u^{0,z}\varphi)(x) \end{aligned}$$

with some function $\chi_r \in C_b^{\infty}(\mathbb{R}^N)$, integers $1 \leq \iota_d, \iota_e \leq N$ and a multi-index β^r satisfying $|\beta^r| \leq 2 - r$ $(0 \leq r \leq 2)$. Then, each term of (B.5) takes the form of

$$\begin{aligned} &\int_0^s u^l \psi(z) P_{s-u}^{0,z} \overline{\mathscr{L}}_{n+1-k+p-r}^{z,(\gamma)} P_u^{0,z}\varphi(x)\, \mathrm{d}u \\ &\qquad = \psi(z) \int_0^s (s-u)^l P_u^{0,z} \overline{\mathscr{L}}_{n+1-k+p-r}^{z,(\gamma)} P_{s-u}^{0,z}\varphi(x)\, \mathrm{d}u \end{aligned} \tag{B.6}$$

and we have

$$P_u^{0,z} \overline{\mathscr{L}}_{n+1-k+p-r}^{z,(\gamma)} P_{s-u}^{0,z}\varphi(x) \tag{B.7}$$

$$= e^{u\mathscr{L}_0^z} \overline{\mathscr{L}}_{n+1-k+p-r}^{z,(\gamma)} e^{(s-u)\mathscr{L}_0^z} \varphi(x)$$

$$= \sum_{j=0}^{n+1-k+p-r} \frac{u^j}{j!} \underbrace{\left[\mathscr{L}_0^z, \left[\mathscr{L}_0^z, \left[\dots, \left[\mathscr{L}_0^z, \overline{\mathscr{L}}_{n+1-k+p-r}^{z,(\gamma)}\right] \dots\right]\right]\right]}_{j\text{-times}} (P_s^{0,z}\varphi)(x). \tag{B.8}$$

where the truncation in the above summation is justified by the same argument of Appendix A.

Note that we are able to show that for $j = 0, 1, 2, \dots, n + 1 - k + p - r$, the term

$$\frac{u^j}{j!} \underbrace{\left[\mathscr{L}_0^z, \left[\mathscr{L}_0^z, \left[\dots, \left[\mathscr{L}_0^z, \overline{\mathscr{L}}_{n+1-k+p-r}^{z,(\gamma)}\right] \dots\right]\right]\right]}_{j\text{-times}} (P_s^{0,z}\varphi)(x)$$

is expressed as

$$\frac{u^j}{j!}\chi_r(z)g_j(z)\times \sum_{\substack{1\le a_1,\dots,a_j\le 2\\ \sum_{i=1}^j a_i\le n+1-k+p-r}} \sum_{\substack{d_1\le 2-a_1\\ \dots\\ d_j\le 2-a_j}} \sum_{\substack{\lambda_1\in\{1,\dots N\}^{d_1}\\ \dots\\ \lambda_j\in\{1,\dots N\}^{d_j}}} \prod_{e=1}^{n+1-k+p-r-\sum_{i=1}^j a_j} (x_{\iota_e}-z_{\iota_e})\partial^{\beta^r}\partial^{\gamma}\partial^{\lambda_1}\dots\partial^{\lambda_j}(P_s^{0,z}\varphi)(x)$$

where $g_j\in C_b^\infty(\mathbb{R}^N)$ and we assume $\prod_{i=1}^0(x_i-z_i)=1$. Then, terms in (B.6) are given by

$$\psi_j(z)\int_0^s (s-u)^l u^j \prod_{e=1}^{n+1-k+p-r-\sum_{i=1}^j a_i} (x_{\iota_e}-z_{\iota_e})\partial^{\beta^r}\partial^{\gamma}\partial^{\lambda_1}\dots\partial^{\lambda_j}(P_s^{0,z}\varphi)(x)\,\mathrm{d}u, \tag{B.9}$$

with $j=0,1,2,\dots,n+1-k+p-r$, $\psi_j\in C_b^\infty(\mathbb{R}^N)$, integers $1\le\iota_e\le N$ and $1\le a_1,\dots,a_j\le 2$ satisfying $\sum_{i=1}^j a_i\le n+1-k+p-r$ and the multi-indices β^r,γ,λ_j whose elements take values in $\{1,2,\dots,N\}$. Moreover, the multi-indices β^r and $\lambda_i, i=1,2,\dots,j$ are defined so as to satisfy $|\beta^r|\le 2-r$ $(0\le r\le 2)$ and $|\lambda_i|\le 2-a_i$. Changing the variable $u\mapsto su$ in (B.9), we obtain

$$\begin{aligned}\psi_j(z)&\int_0^s (s-u)^l u^j \prod_{e=1}^{n+1-k+p-r-\sum_{i=1}^j a_i} (x_{\iota_e}-z_{\iota_e})\partial^{\beta^r}\partial^{\gamma}\partial^{\lambda_1}\dots\partial^{\lambda_j}(P_s^{0,z}\varphi)(x)\,\mathrm{d}u\\ &=\psi_j(z)s^{l+j+1}\int_0^1(1-u)^l u^j\,\mathrm{d}u \prod_{e=1}^{n+1-k+p-r-\sum_{i=1}^j a_i} (x_{\iota_e}-z_{\iota_e})\partial^{\beta^r}\partial^{\gamma}\partial^{\lambda_1}\dots\partial^{\lambda_j}(P_s^{0,z}\varphi)(x).\end{aligned}$$

In particular, it follows that

$$\begin{aligned}n+1-k+p-r-\sum_{i=1}^j a_i+2l+2j+2-|\beta^r|-|\gamma|-\sum_{i=1}^j|\lambda_i|&\\ \ge n+1-k+p-r+2l+2-|\beta^r|-|\gamma|&\\ \ge n+1-k+p+2l-|\gamma|&\\ \ge n+1,&\end{aligned}$$

where we applied $|\lambda_i|\le 2-a_i,\ i=1,2,\dots,j$, $|\beta^r|\le 2-r$ and the assumption of induction $p+2l-|\gamma|\ge k$. This implies that $P_s^{n+1,z}\varphi(x)$ is given as the summation of the terms

$$s^l\psi(z)\prod_{j=1}^p(x_{l_j}-z_{l_j})\partial^\gamma(P_s^{0,z}\varphi)(x), \tag{B.10}$$

with $\psi\in C_b^\infty(\mathbb{R}^N)$ $l\ge 1, p\ge 0$ and $\gamma\in\{1,2,\dots,N\}^{|\gamma|}$ such that $2l+p-|\gamma|\ge n+1$. □

Proof of Proposition 2.2. In the proof, let $t\in(0,1]$, $x\in\mathbb{R}^N$ and C be a generic function independent of t and x. By the perturbation method, we have

$$P_t\varphi(x)=P_t^{0,z}\varphi(x)|_{z=x}+\int_0^t P_{t-s}(\mathscr{L}-\mathscr{L}_0^z)P_s^{0,z}\varphi(x)\,\mathrm{d}s|_{z=x}$$

and expand it through the expansion of $\mathscr{L}$. We can see $\mathscr{L}-\mathscr{L}_0^z=\sum_{i=1}^{2m+1}\mathscr{L}_i^z+\widetilde{\mathscr{L}^z}$, $z\in\mathbb{R}^N$ by Taylor's formula where $\mathscr{L}_i^z$, $i=1,\dots,2m+1$ are given in (2.14) and $\widetilde{\mathscr{L}^z}$ is defined as

$$\widetilde{\mathscr{L}^z}g(x)=\sum_{l_1,\dots,l_{2m+2}=1}^N\prod_{k=1}^{2m+2}(x_{l_k}-z_{l_k})\left\{\sum_{r_1=1}^N h_{r_1}^{l_1,\dots,l_{2m+2}}(x,z)\frac{\partial}{\partial x_{r_1}}g(x)\right.$$
$$\left.+\sum_{r_1,r_2=1}^N h_{r_1,r_2}^{l_1,\dots,l_{2m+2}}(x,z)\frac{\partial^2}{\partial x_{r_1}\partial x_{r_2}}g(x)\right\},\quad g\in C_b^\infty(\mathbb{R}^N),\ x\in\mathbb{R}^N,$$

for some bounded functions $h_{r_1,\dots,r_k}^{l_1,\dots,l_{2m+2}}(\cdot,z)$, $l_1,\dots,l_{2m+2}=1,\dots,N$, $k=1,2$. We note that $s\mapsto P_{t-s}(\mathscr{L}-\mathscr{L}_0^z)P_s^{0,z}\varphi(x)\in L^1([0,t])$ holds since we have

$$\int_0^t|P_{t-s}(\mathscr{L}-\mathscr{L}_0^z)P_s^{0,z}\varphi(x)|\,\mathrm{d}s\le C\left\{\int_0^{t/2}\frac{\|\varphi\|_\infty}{t-s}\,\mathrm{d}s+\int_{t/2}^t\frac{\|\varphi\|_\infty}{s}\,\mathrm{d}s\right\}=2\log 2\times C\|\varphi\|_\infty,\tag{B.11}$$

through the estimates $\|\partial^\alpha P_s^{0,z}\varphi\|_\infty\le C\frac{\|\varphi\|_\infty}{s}$ and $\|H_\alpha(X(t-s,x),h_{\alpha,\beta}(X(t-s,x))p_{\alpha,\beta}(X(t-s,x)-z))\|_p\le C\frac{\|\varphi\|_\infty}{t-s}$, $p\ge1$, $|\alpha|=2$, for the terms appearing in the following representation through the integration by parts:

$$\begin{aligned}&P_{t-s}(\mathscr{L}-\mathscr{L}_0^z)P_s^{0,z}\varphi(x)\\&=\sum_{\substack{\alpha\in\{1,\dots,N\}^{|\alpha|},\\|\alpha|=1,2}}\sum_{\substack{\beta\in\{1,\dots,N\}^{|\beta|},\\\beta\le\nu}}E[\partial^\alpha P_s^{0,z}\varphi(X(t-s,x))h_{\alpha,\beta}(X(t-s,x))p_{\alpha,\beta}(X(t-s,x)-z)]\qquad\text{(B.12)}\\&=\sum_{\substack{\alpha\in\{1,\dots,N\}^{|\alpha|},\\|\alpha|=1,2}}\sum_{\substack{\beta\in\{1,\dots,N\}^{|\beta|},\\\beta\le\nu}}E[P_s^{0,z}\varphi(X(t-s,x))H_\alpha(X(t-s,x),h_{\alpha,\beta}(X(t-s,x))p_{\alpha,\beta}(X(t-s,x)-z))],\qquad\text{(B.13)}\end{aligned}$$

for some $\nu\in\mathbb{N}$, $h_{\alpha,\beta}\in C_b^\infty(\mathbb{R}^N)$ and polynomial functions $p_{\alpha,\beta}:\mathbb{R}^N\to\mathbb{R}$, which is the same argument as in (2.44).

We expand $\int_0^t P_{t-s}(\mathscr{L}-\mathscr{L}_0^z)P_s^{0,z}\varphi(x)|_{z=x}\,\mathrm{d}s$ around $\sum_{i=1}^{2m+1}P_t^{i,z}\varphi(x)|_{z=x}$. Since it holds $\partial_tP_t^{0,z}\varphi(x)=\mathscr{L}_0^zP_t^{0,z}\varphi(x)$, we have

$$\begin{aligned}\partial_tP_t^{i,z}\varphi(x)&=\sum_{k=0}^{i-1}\mathscr{L}_{i-k}^zP_t^{k,z}\varphi(x)+\mathscr{L}_0^z\sum_{k=0}^{i-1}\int_0^tP_{t-s}^{0,z}\mathscr{L}_{i-k}^zP_s^{k,z}\varphi(x)\,\mathrm{d}s\\&=\sum_{k=0}^{i}\mathscr{L}_{i-k}^zP_t^{k,z}\varphi(x),\quad i=1,2,\dots,2m+1.\end{aligned}$$

Then we get

$$(\partial_t-\mathscr{L})\left\{P_t\varphi(x)-\sum_{i=0}^{2m+1}P_t^{i,z}\varphi(x)\right\}=\sum_{k=0}^{2m+1}\left(\mathscr{L}-\sum_{i=0}^{2m+1-k}\mathscr{L}_i^z\right)P_t^{k,z}\varphi(x),$$

by considering $\partial_tP_t\varphi(x)=\mathscr{L}P_t\varphi(x)$. Also note that

$$\begin{aligned}&\partial_t\int_0^tP_{t-s}\left(\mathscr{L}-\sum_{i=0}^{2m+1-k}\mathscr{L}_i^z\right)P_s^{k,z}\varphi(x)\,\mathrm{d}s\\&=\left(\mathscr{L}-\sum_{i=0}^{2m+1-k}\mathscr{L}_i^z\right)P_t^{k,z}\varphi(x)+\mathscr{L}\int_0^tP_{t-s}\left(\mathscr{L}-\sum_{i=0}^{2m+1-k}\mathscr{L}_i^z\right)P_s^{k,z}\varphi(x)\,\mathrm{d}s.\end{aligned}$$

Therefore, when we put $u(t,x) = P_t\varphi(x) - \sum_{i=0}^{2m+1} P_t^{i,z}\varphi(x)$, $\hat{u}(t,x) = \sum_{k=0}^{2m+1}\int_0^t P_{t-s}(\mathscr{L} - \sum_{i=0}^{2m+1-k}\mathscr{L}_i^z) P_s^{k,z}\varphi(x)\,\mathrm{d}s$ and $v(t,x) = \sum_{k=0}^{2m+1}(\mathscr{L} - \sum_{i=0}^{2m+1-k}\mathscr{L}_i^z)P_t^{k,z}\varphi(x)$, we have

$$(\partial_t - \mathscr{L})u(t,x) = v(t,x), \quad (\partial_t - \mathscr{L})\hat{u}(t,x) = v(t,x).$$

Since $\lim_{t\to 0} u(t,x) = \lim_{t\to 0}\hat{u}(t,x) = 0$, it holds that $u(t,x) = \hat{u}(t,x)$ by the uniqueness of the solution to the PDE. Then we have the representation of $\mathscr{R}_t^\varphi(x)$ as follows:

$$\mathscr{R}_t^\varphi(x) = P_t\varphi(x) - \sum_{i=0}^{2m+1} P_t^{i,z}\varphi(x)|_{z=x} = \sum_{k=0}^{2m+1}\int_0^t P_{t-s}\left(\mathscr{L} - \sum_{i=0}^{2m+1-k}\mathscr{L}_i^z\right)P_s^{k,z}\varphi(x)\,\mathrm{d}s|_{z=x} \tag{B.14}$$

$$= \int_0^t P_{t-s}\widetilde{\mathscr{L}^z}P_s^{0,z}\varphi(x)\,\mathrm{d}s|_{z=x} + \sum_{i=1}^{2m+1}\int_0^t P_{t-s}\left(\mathscr{L}_{2m+1}^z + \ldots + \mathscr{L}_{2m+1-(i-1)}^z + \widetilde{\mathscr{L}^z}\right)P_s^{i,z}\varphi(x)\,\mathrm{d}s|_{z=x}.$$

Then, in order to estimate $\mathscr{R}_t^\varphi(x)$, let us consider the following terms:

$$\int_0^t P_{t-s}\widetilde{\mathscr{L}^z}P_s^{0,z}\varphi(x)\,\mathrm{d}s|_{z=x} \quad\text{and}\quad \int_0^t P_{t-s}\mathscr{L}_{2m+1-q}^z P_s^{i,z}\varphi(x)\,\mathrm{d}s|_{z=x},\ i = 1,2,\ldots,2m+1, \tag{B.15}$$

where $q = 0,1,\ldots,i-1$. Due to the definition of the semi-group $\{P_t\}_{t\geq 0}$ and the differential operators $\widetilde{\mathscr{L}^z}$ and $\mathscr{L}_{2m+1-q}^z$, the integrands of the above terms are given by the sum of the form:

$$E\left[\partial^\alpha(P_s^{i,z}\varphi)(X(t-s,x))g\left(X(t-s,x)\right)\prod_{j=1}^{2m+2-q}\left(X^{l_j}(t-s,x) - z_{l_j}\right)\right]|_{z=x}, \tag{B.16}$$

where $\alpha \in \{1,2,\ldots N\}^{|\alpha|}$ is a multi-index of the length $|\alpha| = 1,2$, $q = 0,1,\ldots,i$, $l_j = 1,\ldots,N$ $(j = 1,\ldots,2m+2-q)$ and a function $g \in C_b^\infty(\mathbb{R}^N)$.

For $q = 0,1,\ldots,i$ and multi-index α such that $|\alpha| = 1,2$, let us define

$$\begin{aligned}\Gamma_q^\alpha(s,z) :=&\ E\left[\partial^\alpha\left(P_s^{i,z}\varphi\right)(X(t-s,x))g\left(X(t-s,x)\right)\prod_{j=1}^{2m+2-q}\left(X^{l_j}(t-s,x) - z_{l_j}\right)\right]\\ =&\ E\left[(P_s^{i,z}\varphi)(X(t-s,x))H_\alpha\left(X(t-s,x), g(X(t-s,x))\prod_{j=1}^{2m+2-q}\left(X^{l_j}(t-s,x) - z_{l_j}\right)\right)\right],\end{aligned} \tag{B.17}$$

for $s < t$, $z \in \mathbb{R}^N$. Then, we have

$$|\Gamma_q^\alpha(s,z)| \leq \left\|(P_s^{i,z}\varphi)(X(t-s,x))\right\|_2\left\|H_\alpha(X(t-s,x), g(X(t-s,x))\prod_{j=1}^{2m+2-q}(X^{l_j}(t-s,x) - x_{l_j}))\right\|_2. \tag{B.18}$$

To show the bound of $|\Gamma_q^\alpha(s,x)|$, we need the following lemma of which proof is given after the proof of Proposition 2.2.

Lemma B.2. *Let $q = 0,1,\ldots,i$ and α be a multi-index such that $|\alpha| = 1,2$. Then, when $z = x$, we have*

$$\left\|(P_s^{i,z}\varphi)(X(t-s,x))\right\|_2 \leq C\|\varphi\|_\infty s^{\frac{2l-r}{2}}(t-s)^{\frac{p}{2}}, \tag{B.19}$$

where $l, r \geq 1$ $p \geq 0$ satisfy $2l - r + p \geq i$, and

$$\left\|H_\alpha(X(t-s,x), g(X(t-s,x))\prod_{j=1}^{2m+2-q}(X^{l_j}(t-s,x) - z_{l_j}))\right\|_2 \leq C(t-s)^{\frac{2m-i}{2}}, \tag{B.20}$$

where $C > 0$ is a constant independent of x, s and t.

Using (B.19) and (B.20) in the above lemma, we have for any $q = 0, 1, \dots, i$ and α be a multi-index such that $|\alpha| = 1, 2$,

$$\left| \int_0^t \Gamma_q^\alpha(s,x)\,\mathrm{d}s \right| \le C\|\varphi\|_\infty \int_0^t (t-s)^{\frac{2m-i+p}{2}} s^{\frac{2l-r}{2}}\,\mathrm{d}s$$
$$\le C\|\varphi\|_\infty t^{\frac{2m-i+(2l-r+p)+2}{2}} \le C\|\varphi\|_\infty t^{\frac{2m-i+i+2}{2}} = C\|\varphi\|_\infty t^{m+1},$$

since $2l - r + p \ge i$. Therefore, the proof of Proposition 2.2 is completed. ☐

Proof of Lemma B.2. By Lemma B.1 and Hölder's estimate, we have $\|(P_s^{i,z}\varphi)(X(t-s,x))\|_2 \le C\|\partial^\gamma P_s^{0,z}\varphi\|_\infty s^l \prod_{i=1}^p \|X^{l_j}(t-s,x) - z_{l_j}\|_{r_i}$ for some $p, l \in \mathbb{N}$, $r_i > 1$, $i = 1, \dots, p$ and a multi-index γ of the length $|\gamma| = r$ satisfying $2l - r + p \ge i$. Then when $z = x$ we have

$$\|(P_s^{i,z}\varphi)(X(t-s,x))\|_2 \le C\|\partial^\gamma P_s^{0,z}\varphi\|_\infty s^l (t-s)^{\frac{p}{2}} \le C'\|\varphi\|_\infty s^{\frac{2l-|\gamma|}{2}}(t-s)^{\frac{p}{2}},$$

where we used the basic estimates $\|\partial^\gamma P_s^{0,z}\varphi\|_\infty \le C\|\varphi\|_\infty s^{-\frac{r}{2}}$ and $\|X^l(t-s,x) - x_l\|_\kappa \le C(t-s)^{1/2}$, $\kappa > 1$. Furthermore, we have $\|H_\alpha(X(t-s,x), g(X(t-s,x))\prod_{j=1}^{2m+2-q}(X^{l_j}(t-s,x) - x_{l_j}))\|_2 \le C(t-s)^{-|\alpha|/2}\|\prod_{j=1}^{2m+2-q}(X^{l_j}(t-s,x) - x_{l_j}))\|_{|\alpha|,\kappa}$ for some $\kappa > 1$. Then, when $z = x$,

$$\left\| H_\alpha\left(X(t-s,x), g(X(t-s,x)) \prod_{j=1}^{2m+2-q} (X^{l_j}(t-s,x) - z_{l_j}) \right) \right\|_2$$
$$\le C(t-s)^{\frac{2m+2-q-|\alpha|}{2}} \le C(t-s)^{\frac{2m+2-i-2}{2}} \le C(t-s)^{\frac{2m-i}{2}}.$$

☐

Appendix C. Proof of Proposition 2.3

From the definition of $\{P_t^{i,z}\}_t$ given in (B.1) it holds that for a bounded and measurable $\varphi : \mathbb{R}^N \to \mathbb{R}$,

$$\sum_{i=1}^{2m+1} P_t^{i,z}\varphi(x)|_{z=x} \tag{C.1}$$
$$= \sum_{i=1}^{2m+1} \sum_{i \le \sum_{j=1}^i k_j \le 2m+1} \int_0^t \int_{t_i}^t \cdots \int_{t_2}^t P_{t_i}^{0,z} \mathscr{L}_{k_1}^z P_{t_{i-1}-t_i}^{0,z} \mathscr{L}_{k_2}^z \dots \mathscr{L}_{k_i}^z P_{t-t_1}^{0,z}\varphi(x)\,\mathrm{d}t_1 \dots \mathrm{d}t_i|_{z=x}, \quad t > 0, x \in \mathbb{R}^N.$$

Using the Baker–Campbell–Hausdorff formula for $P_{t_i}^{0,z} \mathscr{L}_{k_1}^z P_{t_{i-1}-t_i}^{0,z} \mathscr{L}_{k_2}^z \dots \mathscr{L}_{k_i}^z P_{t-t_1}^{0,z}\varphi(x)$ in (C.1), we obtain

$$P_{t_i}^{0,z} \mathscr{L}_{k_1}^z P_{t_{i-1}-t_i}^{0,z} \mathscr{L}_{k_2}^z \dots \mathscr{L}_{k_i}^z P_{t-t_1}^{0,z}\varphi(x) \tag{C.2}$$
$$= e^{t_i \mathscr{L}_0^z} \mathscr{L}_{k_1} e^{(t_{i-1}-t_i)\mathscr{L}_0^z} \mathscr{L}_{k_2} \dots \mathscr{L}_{k_i} e^{(t-t_1)\mathscr{L}_0^z} \varphi(x)$$
$$= \prod_{l=1}^i \left(\sum_{\alpha=0}^{k_l} \frac{(t_{i+1-l})^\alpha}{\alpha!} \underbrace{[\mathscr{L}_0^z, [\mathscr{L}_0^z, \dots, [\mathscr{L}_0^z, \mathscr{L}_{k_l}^z]] \dots]]}_{\alpha\text{-times}} \right) P_t^{0,z}\varphi(x)$$
$$= \sum_{\substack{0 \le \alpha_1 \le k_1 \\ \cdots \\ 0 \le \alpha_i \le k_i}} \frac{(t_1)^{\alpha_i}(t_2)^{\alpha_{i-1}} \dots (t_i)^{\alpha_1}}{\alpha_1!\alpha_2! \dots \alpha_i!} \prod_{l=1}^i \left(\underbrace{[\mathscr{L}_0^z, [\mathscr{L}_0^z, \dots, [\mathscr{L}_0^z, \mathscr{L}_{k_l}^z]] \dots]]}_{\alpha_l\text{-times}} \right) P_t^{0,z}\varphi(x).$$

On the second equation of (C.2), the summation is truncated for all $j > k_l$ due to (A.2). Also

$$\underbrace{[\mathscr{L}_0^z, [\mathscr{L}_0^z, \dots, [\mathscr{L}_0^z, \mathscr{L}_{k_l}^z]] \dots]]}_{\alpha\text{-times}} = \mathscr{L}_{k_l} \tag{C.3}$$

is used in (C.2) when $\alpha = 0$. Hence, substituting the equation (C.2) into (C.1), we have

$$\begin{aligned}
&\sum_{i=0}^{2m+1} P_t^{i,z}\varphi(x)|_{z=x} \qquad\qquad (C.4)\\
&= P_t^{0,z}\varphi(x)|_{z=x} + \sum_{i=1}^{2m+1} \sum_{i \le \sum_{j=1}^i k_j \le 2m+1} \sum_{\substack{0\le\alpha_1\le k_1\\ \cdots \\ 0\le\alpha_i\le k_i}} \int_0^t\int_{t_i}^t \cdots \int_{t_2}^t \frac{(t_1)^{\alpha_i}(t_2)^{\alpha_{i-1}}\dots(t_i)^{\alpha_1}}{\alpha_1!\alpha_2!\dots\alpha_i!}\,\mathrm{d}t_1\dots\mathrm{d}t_i\\
&\quad\times \prod_{l=1}^i \left(\underbrace{[\mathscr{L}_0^z, [\mathscr{L}_0^z, \dots, [\mathscr{L}_0^z, \mathscr{L}_{k_l}^z]] \dots]]}_{\alpha_l\text{-times}} \right) P_t^{0,z}\varphi(x)|_{z=x}\\
&= P_t^{0,z}\varphi(x)|_{z=x}\\
&\quad + \sum_{i=1}^{2m+1} \sum_{i \le \sum_{j=1}^i k_j \le 2m+1} \sum_{\substack{1\le\alpha_1\le k_1\\ 0\le\alpha_2\le k_2\\ \cdots \\ 0\le\alpha_i\le k_i}} \frac{t^{\sum_{l=1}^i \alpha_l + i}}{\alpha!} I(\alpha) \prod_{l=1}^i \left(\underbrace{[\mathscr{L}_0^z, [\mathscr{L}_0^z, \dots, [\mathscr{L}_0^z, \mathscr{L}_{k_l}^z]] \dots]]}_{\alpha_l\text{-times}} \right) P_t^{0,z}\varphi(x)|_{z=x},
\end{aligned}$$

where $I(\alpha)$ is given by (2.20). Note that we obtained the second equation through the changing variables in the multiple time integral: $t_k \mapsto tt_k,\ k = 1, 2, \dots, i$, and we took the summation for $\alpha_1 \ge 1$ since whenever $\alpha_1 = 0$,

$$\mathscr{L}_{k_1}^z \prod_{l=2}^i \left(\underbrace{[\mathscr{L}_0^z, [\mathscr{L}_0^z, \dots, [\mathscr{L}_0^z, \mathscr{L}_{k_l}^z]] \dots]]}_{\alpha_l\text{-times}} \right) P_t^{0,z}\varphi(x)|_{z=x} = 0.$$

Then, we decompose the term (C.4) as follows:

$$P_t^{0,z}\varphi(x)|_{z=x} + \sum_{i=1}^{m-1} \sum_{i \le \sum_{j=1}^i k_j \le 2m+1} \sum_{\substack{1\le\alpha_1\le k_1\\ 0\le\alpha_2\le k_2\\ \cdots \\ 0\le\alpha_i\le k_i\\ \sum_{l=1}^i \alpha_l + i \le m}} \frac{t^{\sum_{l=1}^i \alpha_l + i}}{\alpha!} I(\alpha) \prod_{l=1}^i \left(\underbrace{[\mathscr{L}_0^z, [\mathscr{L}_0^z, \dots, [\mathscr{L}_0^z, \mathscr{L}_{k_l}^z]] \dots]]}_{\alpha_l\text{-times}} \right) P_t^{0,z}\varphi(x)|_{z=x} \tag{C.5}$$

$$+ \sum_{i=1}^{m-1} \sum_{i \le \sum_{j=1}^i k_j \le 2m+1} \sum_{\substack{1\le\alpha_1\le k_1\\ 0\le\alpha_2\le k_2\\ \cdots \\ 0\le\alpha_i\le k_i\\ \sum_{l=1}^i \alpha_l + i > m}} \frac{t^{\sum_{l=1}^i \alpha_l + i}}{\alpha!} I(\alpha) \prod_{l=1}^i \left(\underbrace{[\mathscr{L}_0^z, [\mathscr{L}_0^z, \dots, [\mathscr{L}_0^z, \mathscr{L}_{k_l}^z]] \dots]]}_{\alpha_l\text{-times}} \right) P_t^{0,z}\varphi(x)|_{z=x} \tag{C.6}$$

$$+ \sum_{i=m}^{2m+1} \sum_{i \le \sum_{j=1}^i k_j \le 2m+1} \sum_{\substack{1\le\alpha_1\le k_1\\ 0\le\alpha_2\le k_2\\ \cdots \\ 0\le\alpha_i\le k_i}} \frac{t^{\sum_{l=1}^i \alpha_l + i}}{\alpha!} I(\alpha) \prod_{l=1}^i \left(\underbrace{[\mathscr{L}_0^z, [\mathscr{L}_0^z, \dots, [\mathscr{L}_0^z, \mathscr{L}_{k_l}^z]] \dots]]}_{\alpha_l\text{-times}} \right) P_t^{0,z}\varphi(x)|_{z=x}. \tag{C.7}$$

Therefore, it is easy to see that $\sum_{i=0}^{2m+1} P_t^{i,z}\varphi(x)|_{z=x}$ is given as the sum of terms (2.47), (2.48) and (2.49). Furthermore, it is clear that all terms in (C.6) and (C.7) are given as

$$t^{m+1}a(t)b(x)\partial^\beta P_t^{0,z}\varphi(x)|_{z=x},$$

for some multi-index $\beta \in \{1,2,\dots,N\}^{|\beta|}$, a non-decreasing function $a(\cdot)$ and $b \in C_b^\infty(\mathbb{R}^N)$. □

Appendix D. Proof of Lemma 2.1

We will use an abbreviate notation $B_{(j-1)\frac{T}{n},j\frac{T}{n}}$ for the Brownian increment $B_{jT/n}-B_{(j-1)T/n}$ for $j=1,\dots,n$. Let $p=2e,\ e\in\mathbb{N}$. We show that there exist constants $C,c>0$ such that

$$\begin{aligned}E\left[\left|\prod_{j=1}^{J}\left\{1+\pi_{T/n}^{(m),\bar{X}^{(n)}\left((j-1)\frac{T}{n},x\right)}(B_{(j-1)\frac{T}{n},j\frac{T}{n}})\right\}\right|^p\right]&\\ +\sum_{i=1}^{K}E\left[\left\|D^i\prod_{j=1}^{J}\left\{1+\pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)\frac{T}{n},x)}(B_{(j-1)\frac{T}{n},j\frac{T}{n}})\right\}\right\|_{H^{\otimes i}}^p\right]&\le C\left(1+cT/n\right)^J,\end{aligned}\tag{D.1}$$

for $J=1,\dots,n$ and $K\in\mathbb{N}$. If the inequality (D.1) holds, we reach to the conclusion:

$$\left\|\prod_{j=1}^{J}\left\{1+\pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)T/n,x)}(B_{(j-1)\frac{T}{n},j\frac{T}{n}})\right\}\right\|_{K,p}\le C\left(1+cT/n\right)^{J/p}\le C\left(1+cT/n\right)^{n}\le Ce^{cT}.\tag{D.2}$$

We first note that $\pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)\frac{T}{n},x)}(B_{(j-1)\frac{T}{n},j\frac{T}{n}})$ is given by the sum of Wiener functionals of the form:

$$g(\bar{X}^{(n)}((j-1)T/n,x))(T/n)^r\mathscr{P}^i(B_{(j-1)\frac{T}{n},j\frac{T}{n}}),\qquad r+i\ge 2,\ r\in\mathbb{Z},\ i\in\mathbb{N}\cup\{0\},$$

where $g\in C_b^\infty(\mathbb{R}^N)$ and $\mathscr{P}^i:\mathbb{R}^d\ni\xi\mapsto\mathscr{P}^i(\xi)=\prod_{j=1}^d\xi_j^{i_j}$ with $i_1+\dots+i_d=i$.

First, we show the bound for the first term of the left-hand side of (D.1). Let $\mathcal{F}_t:=\sigma(B_s;\ s\le t)$, $t\le T$. Due to the tower property of conditional expectation, we obtain

$$\begin{aligned}&E\left[\left|\prod_{j=1}^{J}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}\left((j-1)\frac{T}{n},x\right)}(B_{(j-1)\frac{T}{n},j\frac{T}{n}})\right)\right|^p\right]\qquad\text{(D.3)}\\&=E\left[\left|\prod_{j=1}^{J-1}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}\left((j-1)\frac{T}{n},x\right)}(B_{(j-1)\frac{T}{n},j\frac{T}{n}})\right)\right|^pE\left[\left|1+\pi_{T/n}^{(m),\bar{X}^{(n)}((J-1)T/n,x)}(B_{(J-1)\frac{T}{n},J\frac{T}{n}})\right|^p\middle|\mathcal{F}_{(J-1)\frac{T}{n}}\right]\right]\\&=E\left[\left|\prod_{j=1}^{J-1}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}\left((j-1)\frac{T}{n},x\right)}(B_{(j-1)\frac{T}{n},j\frac{T}{n}})\right)\right|^pE\left[\left|1+\pi_{T/n}^{(m),\eta}(B_{(J-1)\frac{T}{n},J\frac{T}{n}})\right|^p\right]\Big|_{\eta=\bar{X}^{(n)}\left((J-1)\frac{T}{n},x\right)}\right].\end{aligned}$$

Noting that $B_t^1,B_t^2,\dots,B_t^d$ are independent with each other and for $t>0$ and $k=1,2,\dots,d$,

$$E[(B_t^k)^r]=\begin{cases}0 & (r:\text{odd})\\ \frac{r!}{2^{r/2}(r/2)!}t^{r/2} & (r:\text{even})\end{cases},\tag{D.4}$$

it follows that there exists a constant $c > 0$ such that $E[|1 + \pi_{T/n}^{(m),\eta}(B_{(J-1)\frac{T}{n},J\frac{T}{n}})|^p] \le 1 + cT/n$ for all $\eta \in \mathbb{R}^N$ under the assumption that the coefficients b, σ and their derivatives are bounded. Then we have

$$E\left[\left|\prod_{j=1}^{J}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)\frac{T}{n},x)}(B_{(j-1)\frac{T}{n},j\frac{T}{n}})\right)\right|^p\right] \le (1+cT/n)^J \tag{D.5}$$

for some positive constant $c > 0$.

Next, we assume $K = 1$ and estimate the upper bound of

$$\begin{aligned}&E\left[\left\|D\prod_{j=1}^{J}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)\frac{T}{n},x)}(B_{(j-1)\frac{T}{n},j\frac{T}{n}})\right)\right\|_H^p\right]\\ &\quad = E\left[\left(\sum_{k=1}^{d}\int_0^T\left(D_{k,t}\left\{\prod_{j=1}^{J}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}\left((j-1)\frac{T}{n},x\right)}(B_{(j-1)\frac{T}{n},j\frac{T}{n}})\right)\right\}\right)^2 \mathrm{d}t\right)^e\right]\end{aligned}$$

with $p = 2e$. The chain rule of Malliavin derivative gives

$$\begin{aligned}&\int_0^T\left(D_{k,t}\left\{\prod_{j=1}^{J}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}\left((j-1)\frac{T}{n},x\right)}\left(B_{(j-1)\frac{T}{n},j\frac{T}{n}}\right)\right)\right\}\right)^2 \mathrm{d}t \qquad \text{(D.6)}\\
&= \int_0^T\left|\left\{D_{k,t}\prod_{j=1}^{J-1}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}\left((j-1)\frac{T}{n},x\right)}\left(B_{(j-1)\frac{T}{n},j\frac{T}{n}}\right)\right)\right\}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}\left((J-1)\frac{T}{n},x\right)}\left(B_{(J-1)\frac{T}{n},J\frac{T}{n}}\right)\right)\right.\\
&\quad \left.+\left\{\prod_{j=1}^{J-1}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}\left((j-1)\frac{T}{n},x\right)}\left(B_{(j-1)\frac{T}{n},j\frac{T}{n}}\right)\right)\right\}D_{k,t}\pi_{T/n}^{(m),\bar{X}^{(n)}\left((J-1)\frac{T}{n},x\right)}\left(B_{(J-1)\frac{T}{n},J\frac{T}{n}}\right)\right|^2 \mathrm{d}t\\
&= \sum_{l_1,l_2=1}^{J}\left(\prod_{\substack{j_1\in\{1,\ldots,J\}\setminus\{l_1\}\\ j_2\in\{1,\ldots,J\}\setminus\{l_2\}}}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}\left((j_1-1)\frac{T}{n},x\right)}\left(B_{(j_1-1)\frac{T}{n},j_1\frac{T}{n}}\right)\right)\left(1+\pi^{(m),\bar{X}^{(n)}\left((j_2-1)\frac{T}{n},x\right)}\left(B_{(j_2-1)\frac{T}{n},j_2\frac{T}{n}}\right)\right)\right)\\
&\quad \times\int_0^T D_{k,t}\left\{\pi_{T/n}^{(m),\bar{X}^{(n)}\left((l_1-1)\frac{T}{n},x\right)}\left(B_{(l_1-1)\frac{T}{n},l_1\frac{T}{n}}\right)\right\}D_{k,t}\left\{\pi_{T/n}^{(m),\bar{X}^{(n)}\left((l_2-1)\frac{T}{n},x\right)}\left(B_{(l_2-1)\frac{T}{n},l_2\frac{T}{n}}\right)\right\}\mathrm{d}t.\end{aligned}$$

In particular, $D_{k,t}\{\pi_{T/n}^{(m),\bar{X}^{(n)}((l-1)\frac{T}{n},x)}(B_{(l-1)\frac{T}{n},l\frac{T}{n}})\}$ is reduced to the following term:

$$\begin{aligned}&D_{k,t}g\left(\bar{X}^{(n)}((l-1)T/n,x)\right)(T/n)^r\mathscr{P}^i\left(B_{(l-1)\frac{T}{n},l\frac{T}{n}}\right) \qquad \text{(D.7)}\\
&= g\left(\bar{X}^{(n)}((l-1)T/n,x)\right)(T/n)^r\partial_k\mathscr{P}^i\left(B_{(l-1)\frac{T}{n},l\frac{T}{n}}\right)\mathbf{1}_{((l-1)T/n,lT/n]}(t)\\
&\quad + \sum_{q=1}^{N}\partial_q g\left(\bar{X}^{(n)}((l-1)T/n,x)\right)D_{k,t}\bar{X}^{(n),q}((l-1)T/n,x)\mathbf{1}_{(0,(l-1)T/n]}(t)(T/n)^r\mathscr{P}^i\left(B_{(l-1)\frac{T}{n},l\frac{T}{n}}\right).\end{aligned}$$

Using again the argument of conditional expectation in (D.3) with (D.4) inductively, we obtain

$$E\left[\left(\int_0^T\left(D_t\left\{\prod_{j=1}^{J}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}\left((j-1)\frac{T}{n},x\right)}\left(B_{(j-1)\frac{T}{n},j\frac{T}{n}}\right)\right)\right\}\right)^2\mathrm{d}t\right)^e\right] \le C\left(1+c\frac{T}{n}\right)^J, \tag{D.8}$$

by applying the estimate $\sup_{1\le j\le n}\| \int_0^T |D_t\bar{X}^{(n)}((j-1)T/n,x)|^2\,\mathrm{d}t\|_p \le c'$ (by Bally and Talay [3]), where constants $C,c,c'>0$ are independent of J and n. Therefore, there exist $C,c>0$ (which are independent of J and n) such that

$$E\left[\left(\int_0^T\left(D_t\left\{\prod_{j=1}^{J}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)\frac{T}{n},x)}\left(B_{j\frac{T}{n},(j-1)\frac{T}{n}}\right)\right)\right\}\right)^2\mathrm{d}t\right)^{p/2}\right]\le Ce^{cT}, \tag{D.9}$$

for all $J=1,\dots,n$. For $K\ge 2$, we proceed in the same way and obtain the bound (D.1). □

Appendix E. Proof of Lemma 2.2

We prepare some notation and basic facts on Watanabe distributions on Wiener space. Let us denote the Dirac delta function mass at $y\in\mathbb{R}^N$ by δ_y which is an element of the space of Schwartz tempered distributions $\mathcal{S}'(\mathbb{R}^N)$, the dual of the space of Schwartz rapidly decreasing functions $\mathcal{S}(\mathbb{R}^N)$. We define the space of Watanabe distributions $\mathbb{D}^{-\infty}$ as the dual of $\mathbb{D}^{\infty}$ and denote by ${}_{-\infty}\langle\Phi,G\rangle_{\infty}$ the coupling between $\Phi\in\mathbb{D}^{-\infty}$ and $G\in\mathbb{D}^{\infty}$. Note that the composition $\delta_y(F)$ of δ_y and a nondegenerate $F\in(\mathbb{D}^{\infty})^N$ is well-defined as an element of $\mathbb{D}^{-\infty}$ and we have for any bounded and measurable function $f:\mathbb{R}^N\to\mathbb{R}$ and $G\in\mathbb{D}^{\infty}$,

$$E[f(F)G]=\int_{\mathbb{R}^N} f(y)\,{}_{-\infty}\langle\delta_y(F),G\rangle_{\infty}\,\mathrm{d}y. \tag{E.1}$$

Furthermore, for $S\in\mathcal{S}'(\mathbb{R}^N)$, it holds that

$${}_{-\infty}\langle\partial_i S(F),G\rangle_{\infty}={}_{-\infty}\langle S(F),H_{(i)}(F,G)\rangle_{\infty}={}_{\mathcal{S}'}\langle S,E[H_{(i)}(F,G)|F=\cdot]p^F(\cdot)\rangle_{\mathcal{S}}, \tag{E.2}$$

where ${}_{\mathcal{S}'}\langle\cdot,\cdot\rangle_{\mathcal{S}}$ is the coupling on $\mathcal{S}'(\mathbb{R}^N)\times\mathcal{S}(\mathbb{R}^N)$, and p^F is the density of F.

Hereafter we write $\bar{X}^x_t$ for $\bar{X}(t,x)=\bar{X}^z(t,x)|_{z=x}$ for $t>0, x\in\mathbb{R}^N$. We give a representation of $(Q^{(m)}_{T/n})^{n-1}\Psi^f_{T/n}(x)$ where $\Psi^f_t(\cdot)=\sum_{l\le\nu}h_l(t)g_l(\cdot)\partial^{\beta^{(l)}}P^{0,z}_t f(\cdot)|_{z=\cdot}$. Since we have for $\beta=(\beta_1,\dots,\beta_e)$, $e\in\mathbb{N}$,

$$\begin{aligned}\partial^{\beta}P^{0,z}_t f(x)|_{z=x}&=\frac{\partial^e}{\partial x_{\beta_1}\dots\partial x_{\beta_e}}\int_{\mathbb{R}^N}f(y)\,{}_{-\infty}\langle\delta_y(\bar{X}^z(t,x)),1\rangle_{\infty}\,\mathrm{d}y|_{z=x}\\&=\frac{\partial^{e-1}}{\partial x_{\beta_1}\dots\partial x_{\beta_{e-1}}}\int_{\mathbb{R}^N}f(y)\sum_{k=1}^N{}_{-\infty}\langle\partial_k\delta_y(\bar{X}^z(t,x)),\mathbf{1}_{k=\beta_e}\rangle_{\infty}\,\mathrm{d}y|_{z=x}\\&=\frac{\partial^{e-1}}{\partial x_{\beta_1}\dots\partial x_{\beta_{e-1}}}\int_{\mathbb{R}^N}f(y)\,{}_{-\infty}\langle\partial_{\beta_e}\delta_y(\bar{X}^z(t,x)),1\rangle_{\infty}\,\mathrm{d}y|_{z=x}\\&=\int_{\mathbb{R}^N}f(y)\,{}_{-\infty}\langle\partial^{\beta}\delta_y(\bar{X}^z(t,x)),1\rangle_{\infty}\,\mathrm{d}y|_{z=x}\\&=\int_{\mathbb{R}^N}f(y)\,{}_{-\infty}\langle\delta_y(\bar{X}^z(t,x)),H_{\beta}(\bar{X}^z(t,x),1)\rangle_{\infty}\,\mathrm{d}y|_{z=x},\end{aligned} \tag{E.3}$$

one gets

$$\begin{aligned}&(Q^{(m)}_{T/n})^{n-1}\Psi^f_{T/n}(x)\\&=E\left[\Psi^f_{T/n}\left(\bar{X}^{(n)}((n-1)T/n,x)\right)\prod_{j=1}^{n-1}\left(1+\pi^{(m),\bar{X}^{(n)}((j-1)T/n,x)}_{T/n}(B_{jT/n}-B_{(j-1)T/n})\right)\right]\\&=\sum_{l\le\nu}h_l(T/n)E\left[\int_{\mathbb{R}^N}f(y)_{-\infty}\left\langle\delta_y\left(\bar{X}^{\xi}_{T/n}\right),H_{\beta^{(l)}}\left(\bar{X}^{\xi}_{T/n},1\right)\right\rangle_{\infty}\mathrm{d}y\Big|_{\xi=\bar{X}^{(n)}((n-1)T/n,x)}G^{(n-1)}_l\right],\end{aligned} \tag{E.4}$$

where the $\mathcal{F}_{(n-1)T/n}$-measurable random variable $G_l^{(n-1)} \in \mathbb{D}^\infty$ is defined as

$$G_l^{(n-1)} = g_l(\bar{X}^{(n)}((n-1)T/n, x)) \prod_{j=1}^{n-1} \left\{ 1 + \pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)T/n,x)}(B_{jT/n} - B_{(j-1)T/n}) \right\}. \tag{E.5}$$

Furthermore, using Fubini's theorem, we obtain

$$\begin{aligned} &E\left[\int_{\mathbb{R}^N} f(y) {}_{-\infty}\left\langle \delta_y\left(\bar{X}^\xi_{T/n}\right), H_{\beta^{(l)}}\left(\bar{X}^\xi_{T/n}, 1\right) \right\rangle_\infty \mathrm{d}y \Big|_{\xi=\bar{X}^{(n)}((n-1)T/n,x)} G_l^{(n-1)} \right] \\ &\quad = \int_{\mathbb{R}^N} f(y) E\left[{}_{-\infty}\left\langle \delta_y\left(\bar{X}^\xi_{T/n}\right), H_{\beta^{(l)}}\left(\bar{X}^\xi_{T/n}, 1\right) \right\rangle_\infty \Big|_{\xi=\bar{X}^{(n)}((n-1)T/n,x)} G_l^{(n-1)} \right] \mathrm{d}y. \end{aligned}$$

If we show that it holds

$$\begin{aligned} &E\left[{}_{-\infty}\left\langle \delta_y\left(\bar{X}^\xi_{T/n}\right), H_{\beta^{(l)}}\left(\bar{X}^\xi_{T/n}, 1\right) \right\rangle_\infty \Big|_{\xi=\bar{X}^{(n)}((n-1)T/n,x)} G_l^{(n-1)} \right] \\ &\quad = E\left[\delta_y\left(\bar{X}^{(n)}(T,x)\right) H_{\beta^{(l)}}\left(\bar{X}^{(n)}(T,x), G_l^{(n-1)}\right) \right], \end{aligned} \tag{E.6}$$

then from the equations (E.4) and (E.6) we get

$$\begin{aligned} (Q^{(m)}_{T/n})^{n-1} \Psi^f_{T/n}(x) &= \sum_{l \le \nu} h_l(T/n) \int_{\mathbb{R}^N} f(y) {}_{-\infty}\left\langle \delta_y\left(\bar{X}^{(n)}(T,x)\right), H_{\beta^{(l)}}\left(\bar{X}^{(n)}(T,x), G_l^{(n-1)}\right) \right\rangle_\infty \mathrm{d}y \\ &= \sum_{l \le \nu} h_l(T/n) E\left[f\left(\bar{X}^{(n)}(T,x)\right) H_{\beta^{(l)}}\left(\bar{X}^{(n)}(T,x), G_l^{(n-1)}\right) \right] \end{aligned}$$

and easily reach to the conclusion with the same argument we gave in the last part of proof of Theorem 2.1.

In order to complete the proof, let us show that the equation (E.6) holds. We note that the distribution δ_y is represented as the weak derivative of Heaviside function, namely, $\delta_y(\cdot) = \partial^\gamma T_y(\cdot)$, where $\gamma = (1, 2, \dots, N)$ and for $x \in \mathbb{R}^N$, $T_y(x) = 1$ if $x_i \ge y_i$ for all $i = 1, 2, \dots, N$ and $T_y(x) = 0$ otherwise. Then, we introduce the mollifier of Heaviside function T_y given by $T_y^\varepsilon := T_y * \psi_\varepsilon \in C^\infty(\mathbb{R}^N)$ with some suitable smooth function ψ_ε, $\varepsilon > 0$ on $\mathbb{R}^N$ such that $T_y^\varepsilon \to T_y$ $(\varepsilon \to 0)$ where the limit is understood in the space of Schwartz distributions. We consider the following function on $\mathbb{R}^N$ depending on $\varepsilon > 0$:

$$F^\varepsilon(y) := E\left[{}_{-\infty}\left\langle \partial^\gamma T_y^\varepsilon\left(\bar{X}^\xi_{T/n}\right), H_{\beta^{(l)}}\left(\bar{X}^\xi_{T/n}, 1\right) \right\rangle_\infty \Big|_{\xi=\bar{X}^{(n)}((n-1)T/n,x)} G_l^{(n-1)} \right], \quad y \in \mathbb{R}^N. \tag{E.7}$$

From now on, we transform $F^\varepsilon(y)$ in two ways. First, we have

$$\begin{aligned} F^\varepsilon(y) &= E\left[{}_{-\infty}\left\langle \partial^{\beta^{(l)}} \partial^\gamma T_y^\varepsilon\left(\bar{X}^\xi_{T/n}\right), 1 \right\rangle_\infty \Big|_{\xi=\bar{X}^{(n)}((n-1)T/n,x)} G_l^{(n-1)} \right] \\ &= E\left[\partial^{\beta^{(l)}} \partial^\gamma T_y^\varepsilon\left(\bar{X}^{(n)}(T,x)\right) G_l^{(n-1)} \right] \\ &= E\left[T_y^\varepsilon\left(\bar{X}^{(n)}(T,x)\right) H_{\beta^{(l)} * \gamma}\left(\bar{X}^{(n)}(T,x), G_l^{(n-1)}\right) \right], \end{aligned} \tag{E.8}$$

where $\beta^{(l)} * \gamma := (\beta^{(l)}_1, \dots, \beta^{(l)}_l, 1, 2, \dots, N)$. On the other hand, we get

$$F^\varepsilon(y) = E\left[{}_{-\infty}\left\langle T_y^\varepsilon\left(\bar{X}^\xi_{T/n}\right), H_{\beta^{(l)} * \gamma}\left(\bar{X}^\xi_{T/n}, 1\right) \right\rangle_\infty \Big|_{\xi=\bar{X}^{(n)}((n-1)T/n,x)} G_l^{(n-1)} \right]. \tag{E.9}$$

Now, we take the limit $\varepsilon \downarrow 0$ for both terms (E.8) and (E.9). For (E.8), since T_y^ε is a bounded function and there exists a constant $M > 0$ such that $|T_y^\varepsilon(\bar{X}^{(n)}(T,x)(\omega))| \le M$ for all $\omega \in \Omega$, $\varepsilon > 0$, the dominated convergence theorem allows us to exchange the limit and integration and get

$$\begin{aligned}
&\lim_{\varepsilon\downarrow 0} E\left[T_y^\varepsilon\left(\bar{X}^{(n)}(T,x)\right)H_{\beta^{(l)}*\gamma}\left(\bar{X}^{(n)}(T,x),G_l^{(n-1)}\right)\right] \\
&\quad= E\left[T_y\left(\bar{X}^{(n)}(T,x)\right)H_{\beta^{(l)}*\gamma}\left(\bar{X}^{(n)}(T,x),G_l^{(n-1)}\right)\right] \\
&\quad= E\left[\delta_y\left(\bar{X}^{(n)}(T,x)\right)H_{\beta^{(l)}}\left(\bar{X}^{(n)}(T,x),G_l^{(n-1)}\right)\right].
\end{aligned} \tag{E.10}$$

For (E.9), let $h_y^\varepsilon(\xi) := {}_{-\infty}\langle T_y^\varepsilon(\bar{X}_{T/n}^\xi), H_{\beta^{(l)}*\gamma}(\bar{X}_{T/n}^\xi,1)\rangle_\infty = E[T_y^\varepsilon(\bar{X}_{T/n}^\xi)H_{\beta^{(l)}*\gamma}(\bar{X}_{T/n}^\xi,1)]$, then there exists constants $C,q>0$ (which do not depend on ξ and ε) such that

$$\sup_{\xi\in\mathbb{R}^N,\varepsilon>0}\left|h_y^\varepsilon(\xi)\right| \le \sup_{\xi\in\mathbb{R}^N,\varepsilon>0}\left\{\|T_y^\varepsilon\|_\infty\|H_{\beta^{(l)}*\gamma}\left(\bar{X}_{T/n}^\xi,1\right)\|_1\right\} \le C\frac{1}{(T/n)^q}. \tag{E.11}$$

Therefore when $\varepsilon\downarrow 0$ we have

$$\lim_{\varepsilon\downarrow 0}E[h_y^\varepsilon(Z)G] = E[\lim_{\varepsilon\downarrow 0}h_y^\varepsilon(Z)G] \tag{E.12}$$

for $Z,G\in\mathbb{D}^\infty$. Also, since $T_y^\varepsilon\to T_y$ in $\mathcal{S}'(\mathbb{R}^N)$ $(\varepsilon\to 0)$ we have

$$\begin{aligned}
h_y^\varepsilon(\xi) &= {}_{\mathcal{S}'}\left\langle T_y^\varepsilon, E[H_{\beta^{(l)}*\gamma}\left(\bar{X}_{T/n}^\xi,1\right)|\bar{X}_{T/n}^\xi=\cdot\,]p^{\bar{X}_{T/n}^\xi}(\cdot)\right\rangle_{\mathcal{S}} \\
&\to {}_{\mathcal{S}'}\left\langle T_y, E[H_{\beta^{(l)}*\gamma}\left(\bar{X}_{T/n}^\xi,1\right)|\bar{X}_{T/n}^\xi=\cdot\,]p^{\bar{X}_{T/n}^\xi}(\cdot)\right\rangle_{\mathcal{S}}
\end{aligned} \tag{E.13}$$

by (E.2) and by the basic argument on the Schwartz distribution theory. Therefore, we get

$$\begin{aligned}
&\lim_{\varepsilon\downarrow 0}E\left[{}_{-\infty}\left\langle T_y^\varepsilon\left(\bar{X}_{T/n}^\xi\right),H_{\beta^{(l)}*\gamma}\left(\bar{X}_{T/n}^\xi,1\right)\right\rangle_\infty\Big|_{\xi=\bar{X}^{(n)}((n-1)T/n,x)}G_l^{(n-1)}\right] \\
&\quad=\lim_{\varepsilon\downarrow 0}E\left[h_y^\varepsilon\left(\bar{X}^{(n)}((n-1)T/n,x)\right)G_l^{(n-1)}\right] \\
&\quad=E\left[\lim_{\varepsilon\downarrow 0}h_y^\varepsilon\left(\bar{X}^{(n)}((n-1)T/n,x)\right)G_l^{(n-1)}\right] \\
&\quad=E\left[\lim_{\varepsilon\downarrow 0}{}_{-\infty}\left\langle T_y^\varepsilon\left(\bar{X}_{T/n}^\xi\right),H_{\beta^{(l)}*\gamma}\left(\bar{X}_{T/n}^\xi,1\right)\right\rangle_\infty\Big|_{\xi=\bar{X}^{(n)}((n-1)T/n,x)}G_l^{(n-1)}\right] \\
&\quad=E\left[{}_{-\infty}\left\langle T_y\left(\bar{X}_{T/n}^\xi\right),H_{\beta^{(l)}*\gamma}\left(\bar{X}_{T/n}^\xi,1\right)\right\rangle_\infty\Big|_{\xi=\bar{X}^{(n)}((n-1)T/n,x)}G_l^{(n-1)}\right] \\
&\quad=E\left[{}_{-\infty}\left\langle \delta_y\left(\bar{X}_{T/n}^\xi\right),H_{\beta^{(l)}}\left(\bar{X}_{T/n}^\xi,1\right)\right\rangle_\infty\Big|_{\xi=\bar{X}^{(n)}((n-1)T/n,x)}G_l^{(n-1)}\right].
\end{aligned} \tag{E.14}$$

In conclusion, we obtain

$$\begin{aligned}
\lim_{\varepsilon\downarrow 0}F^\varepsilon(y) &= E\left[\delta_y\left(\bar{X}^{(n)}(T,x)\right)H_{\beta^{(l)}}\left(\bar{X}^{(n)}(T,x),G_l^{(n-1)}\right)\right] \\
&= E\left[{}_{-\infty}\left\langle \delta_y\left(\bar{X}_{T/n}^\xi\right),H_{\beta^{(l)}}\left(\bar{X}_{T/n}^\xi,1\right)\right\rangle_\infty\Big|_{\xi=\bar{X}^{(n)}((n-1)T/n,x)}G_l^{(n-1)}\right], \quad y\in\mathbb{R}^N.
\end{aligned}$$

□

Appendix F. Proof of Theorem 3.1

We prepare the following three results which play an important role in the proof of Theorem 3.1.

Lemma F.1. *It holds that there exist constants $C, c > 0$ such that*

$$\frac{1}{r^{\frac{N}{2}}}\int_{\mathbb{R}^{\mathbb{N}}} \exp\left(-c_1\frac{|x-z|^2}{r}\right)\exp\left(-c_2\frac{|y-z|^2}{s-r}\right)\mathrm{d}z \leq C\left(\frac{s-r}{s}\right)^{\frac{N}{2}}\exp\left(-c\frac{|x-y|^2}{s}\right) \tag{F.1}$$

for any constants $c_1, c_2 > 0$ and $(r,x,y,z) \in (0,s)\times\mathbb{R}^N\times\mathbb{R}^N\times\mathbb{R}^N$. Furthermore, it holds for any $(r,x,y) \in [0,s)\times\mathbb{R}^N\times\mathbb{R}^N$ and $c > 0$,

$$E\left[\exp\left(-c\frac{|y-\bar{X}^{(n)}(r,x)|^2}{s-r}\right)\right] \leq K(T)\left(\frac{s-r}{s}\right)^{\frac{N}{2}}\exp\left(-c'\frac{|x-y|^2}{s}\right), \tag{F.2}$$

where $c' > 0$ and $K(\cdot)$ is a non decreasing function.

Proof. See Gobet and Labart [9]. □

Lemma F.2 (Small time expansion of heat kernel)**.** *It holds for any $(t,x,y) \in (0,1]\times\mathbb{R}^N\times\mathbb{R}^N$*

$$p_t^X(x,y) - p_t^{\bar{X},(m)}(x,y) = \mathscr{R}_t(x,y) + t^{m+1}\Psi_t(x,y), \tag{F.3}$$

where

$$|\mathscr{R}_t(x,y)| \leq Ct^{m+1}\frac{1}{t^{\frac{N}{2}}}\exp\left(-c\frac{|y-x|^2}{t}\right) \tag{F.4}$$

with some constants $C, c > 0$ and $\Psi_t(x,y)$ has the following representation:

$$\Psi_t(x,y) = \sum_{l\leq\nu} h_l(t)g_l(x)\partial^{\beta_l}p_t^{\bar{X}^z}(x,y)|_{z=x} \tag{F.5}$$

for some $\nu \in \mathbb{N}$, functions h_l, $l \leq \nu$ at most polynomial growth, $g_l \in C_b^\infty(\mathbb{R}^N)$, $l \leq \nu$ and multi-indices $\beta^{(l)} \in \{1,2,\ldots,N\}^l$, $l \leq \nu$.

Proof. To give the upper bound of (F.4), it is enough to estimate

$$\begin{aligned}&\left|\int_0^t s^k E\left[\partial^\alpha\left(p_s^{\bar{X}^x}(X(t-s,x),y)\right)g(X(t-s,x))\prod_{i=1}^e\left(X^{l_i}(t-s,x)-x_{l_i}\right)\right]\mathrm{d}s\right| \\ &= \left|\int_0^t s^k E\Big[p_s^{\bar{X}^x}(X(t-s,x),y)H_\alpha\Big(X(t-s,x), g(X(t-s,x))\prod_{i=1}^e(X^{l_i}(t-s,x)-x_{l_i})\Big)\Big]\,\mathrm{d}s\right|,\end{aligned} \tag{F.6}$$

for $k \geq 1, e \geq 0$ and a multi-index α satisfying $k + \frac{e-|\alpha|}{2} \geq m$, by (B.16) with Lemma 2.3. We note that Hölder's inequality and the estimate in Lemma B.2 give

$$\begin{aligned}&\left|E\left[p_s^{\bar{X}^x}(X(t-s,x),y)H_\alpha\left(X(t-s,x), g(X(t-s,x))\prod_{i=1}^e(X^{l_i}(t-s,x)-x_{l_i})\right)\right]\right| \\ &\leq C\left\{E\left[\left|p_s^{\bar{X}^x}(X(t-s,x),y)\right|^q\right]\right\}^{1/q}(t-s)^{\frac{e-|\alpha|}{2}}\end{aligned}$$

for some $q > 1$ and $C > 0$. Now, due to the explicit formula of $p_s^{\bar{X}}(x,y)$, $(s,x,y) \in (0,\infty) \times \mathbb{R}^N \times \mathbb{R}^N$, there exist constants $K, c > 0$ such that

$$\begin{aligned}\left\{E\left[\left|p_s^{\bar{X}^x}(X(t-s,x),y)\right|^q\right]\right\}^{1/q} &\leq \left\{E\left[\frac{K^q}{s^{\frac{Nq}{2}}}\exp\left(-c'q\frac{|y-X(t-s,x)|^2}{s}\right)\right]\right\}^{1/q}\\ &\leq \frac{K}{s^{\frac{N}{2}}}\left\{E\left[\exp\left(-c\frac{|y-X(t-s,x)|^2}{s}\right)\right]\right\}^{1/q}\\ &\leq \frac{K}{s^{\frac{N}{2}}}\left(\frac{s}{t}\right)^{\frac{N}{2q}}\exp\left(-c\frac{|y-x|^2}{t}\right),\end{aligned}$$

where on the last inequality we used the following result which is derived from the similar argument in the proof of Lemma F.1: There exists some constant $C > 0$ such that for $c > 0$

$$E\left[\exp\left(-c\frac{|y-X(t-s,x)|^2}{s}\right)\right] \leq C\left(\frac{s}{t}\right)^{\frac{N}{2}}\exp\left(-c\frac{|y-x|^2}{t}\right). \tag{F.7}$$

Hence, we have

$$\begin{aligned}&\left|\int_0^t s^k E[\partial^\alpha p_s^{\bar{X}^x}(X(t-s,x),y)g(X(t-s,x))\prod_{i=1}^{e}(X^{l_i}(t-s,x)-x_{l_i})]\,\mathrm{d}s\right|\\ &\quad\leq C\int_0^t s^{k+\frac{N}{2q}-\frac{N}{2}}(t-s)^{\frac{e-|\alpha|}{2}}\,\mathrm{d}s \times \frac{1}{t^{\frac{N}{2q}}}\exp\left(-c\frac{|y-x|^2}{t}\right)\\ &\quad\leq Ct^{k+\frac{e-|\alpha|}{2}+1}\frac{1}{t^{\frac{N}{2}}}\exp\left(-c\frac{|y-x|^2}{t}\right)\\ &\quad\leq Ct^{m+1}\frac{1}{t^{\frac{N}{2}}}\exp\left(-c\frac{|y-x|^2}{t}\right),\end{aligned}$$

since $k + \frac{e-|\alpha|}{2} \geq m$. The representation (F.5) is immediately obtained from (2.50) of Lemma 2.3 by replacing $P_t^{0,z}\varphi(y)$ with the kernel $p_t^{\bar{X}^z}(x,y)$. ☐

Lemma F.3. *Let $(s,x,y) \in (0,T] \times \mathbb{R}^N \times \mathbb{R}^N$. Also let T_y be the Heaviside function defined in Appendix E. Then, for a multi-index $\alpha = (\alpha_1,\dots,\alpha_\ell) \in \{1,\dots,N\}^\ell$, where ℓ is an integer satisfying $\ell = |\alpha| \geq 0$ and $\gamma = (1,\dots,N)$, it holds*

$$\begin{aligned}&\left|E\left[T_y(\bar{X}^{(n)}(s,x))H_{\gamma*\alpha}\left(\bar{X}^{(n)}(s,x),\prod_{j=1}^{J}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)T/n,x)}\left(B_{jT/n}-B_{(j-1)T/n}\right)\right)\right)\right]\right|\\ &\quad\leq \frac{K(T)}{s^{\frac{N+|\alpha|}{2}}}\exp\left(-c\frac{|y-x|^2}{s}\right)\end{aligned}$$

with some non-decreasing function $K(\cdot)$ and constant $c > 0$ both of which are independent of $J = 1,2,\dots,n$. Here, a notation $\gamma * \alpha := (1,\dots,N,\alpha_1,\dots,\alpha_\ell)$ is used.

Proof. Hölder's inequality gives

$$\begin{aligned}&\left|E\left[T_y\left(\bar{X}^{(n)}(s,x)\right)H_{\gamma*\alpha}\left(\bar{X}^{(n)}(s,x),\prod_{j=1}^{J}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)T/n,x)}\left(B_{jT/n}-B_{(j-1)T/n}\right)\right)\right)\right]\right|\\ &\quad\leq \|T_y\left(\bar{X}^{(n)}(s,x)\right)\|_2\left\|H_{\gamma*\alpha}\left(\bar{X}^{(n)}(s,x),\prod_{j=1}^{J}\left(1+\pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)T/n,x)}\left(B_{jT/n}-B_{(j-1)T/n}\right)\right)\right)\right\|_2.\end{aligned} \tag{F.8}$$

Using the upper bound of the density of the Euler–Maruyama scheme given in [9], we obtain

$$E[|T_y(\bar{X}^{(n)}(s,x))|^2] = E[T_y(\bar{X}^{(n)}(s,x))] \le \frac{K(T)}{s^{\frac{N}{2}}} \int_{\mathbb{R}^N} T_y(\xi) \exp\left(-c\frac{|\xi - x|^2}{s}\right) \mathrm{d}\xi$$
$$\le K(T) \prod_{i=1}^N \frac{1}{\sqrt{s}} \int_{\mathbb{R}} T_{y_i}(\xi_i) \exp\left(-c\frac{|\xi_i - x_i|^2}{s}\right) \mathrm{d}\xi_i$$

for some non-decreasing function $K(\cdot)$. When $y_i - x_i \ge \sqrt{s},\ i = 1, 2, \ldots, N$, we have

$$\frac{1}{\sqrt{s}} \int_{\mathbb{R}} T_{y_i}(\xi_i) \exp\left(-c\frac{|\xi_i - x_i|^2}{s}\right) \mathrm{d}\xi_i = \frac{1}{\sqrt{s}} \int_{y_i - x_i}^{\infty} \exp\left(-c\frac{z^2}{s}\right) \mathrm{d}z$$
$$\le \frac{1}{\sqrt{s}} \int_{y_i - x_i}^{\infty} \frac{z}{y_i - x_i} \exp\left(-c\frac{z^2}{s}\right) \mathrm{d}z \le C_1 \exp\left(-c\frac{|y_i - x_i|^2}{s}\right)$$

for some $C_1 > 0$. On the other hand, for $y_i - x_i \le \sqrt{s},\ i = 1, 2, \ldots, N$,

$$\frac{1}{\sqrt{s}} \int_{\mathbb{R}} T_{y_i}(\xi_i) \exp\left(-c\frac{|\xi_i - x_i|^2}{s}\right) \mathrm{d}\xi_i \le \frac{1}{\sqrt{s}} \int_{\mathbb{R}} \exp\left(-c\frac{|\xi_i - x_i|^2}{s}\right) \mathrm{d}\xi_i \ \le C$$
$$\le C \exp\left(c\frac{(y_i - x_i)^2}{s}\right) \exp\left(-c\frac{(y_i - x_i)^2}{s}\right) \le C_2 \exp\left(-c\frac{(y_i - x_i)^2}{s}\right)$$

for some constant $C_2 > 0$. Therefore, we have

$$E[|T_y(\bar{X}^{(n)}(s,x))|^2]^{1/2} \le K(T) \exp\left(-c_1\frac{|y - x|^2}{s}\right). \tag{F.9}$$

Furthermore, by Kusuoka–Stroock's integration by parts, there exists a constant $c > 0$ such that

$$\left\| H_{\gamma * \alpha}\left(\bar{X}^{(n)}(s,x), \prod_{j=1}^J \left(1 + \pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)T/n,x)}\left(B_{jT/n} - B_{(j-1)T/n}\right)\right)\right)\right\|_2$$
$$\le \frac{c}{s^{\frac{N+|\alpha|}{2}}} \left\| \prod_{j=1}^J \left(1 + \pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)T/n,x)}\left(B_{jT/n} - B_{(j-1)T/n}\right)\right)\right\|_{N+|\alpha|,4}.$$

By Lemma 2.1, we have the bound $K(T)$ independent of $J = 1, \ldots, n$ such that

$$\left\| H_{\gamma * \alpha}\left(\bar{X}^{(n)}(s,x), \prod_{j=1}^J \left(1 + \pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)T/n,x)}\left(B_{jT/n} - B_{(j-1)T/n}\right)\right)\right)\right\|_2 \le \frac{c}{s^{\frac{N+|\alpha|}{2}}} K(T). \tag{F.10}$$

Applying (F.9) and (F.10) to (F.8), we have the assertion. □

Proof of Theorem 3.1. The global approximation is written using $Q_t^{(m)}$, $t > 0$ given in (2.19) of Theorem 2.1 as follows:

$$p_T^X(x,y) - E\left[p_{T/n}^{\bar{X},(m)}\left(\bar{X}^{(n)}((n-1)T/n,x), y\right) \prod_{j=1}^{n-1} \left\{1 + \pi_{T/n}^{(m),\bar{X}^{(n)}((j-1)T/n,x)}\left(B_{jT/n} - B_{(j-1)T/n}\right)\right\}\right]$$
$$= p_T^X(x,y) - \left(Q_{T/n}^{(m)}\right)^{n-1} p_{T/n}^{\bar{X},(m)}(\cdot, y)|_{\cdot = x}$$
$$= \sum_{k=0}^{n-2} \left(Q_{T/n}^{(m)}\right)^k \left(P_{T/n} - Q_{T/n}^{(m)}\right) p_{T-(k+1)T/n}^X(\cdot, y)|_{\cdot = x} + \left(Q_{T/n}^{(m)}\right)^{n-1} \left(p_{T/n}^X(\cdot, y) - p_{T/n}^{\bar{X},(m)}(\cdot, y)\right)|_{\cdot = x}.$$

By Lemma F.2, we have for every $k = 0, 1, 2, \ldots, n-2$,

$$\begin{aligned}
&\left(P_{T/n} - Q^{(m)}_{T/n}\right) p^X_{T-(k+1)T/n}(\cdot, y)|_{\cdot=x} \\
&\quad = \int_{\mathbb{R}^N} p^X_{T-(k+1)T/n}(z, y) \left(p^X_{T/n}(x, z) - p^{\bar{X},(m)}_{T/n}(x, z)\right) \mathrm{d}z \\
&\quad = \int_{\mathbb{R}^N} p^X_{T-(k+1)T/n}(z, y) \left(\mathscr{R}_{T/n}(x, z) + (T/n)^{m+1} \Psi_{T/n}(x, z)\right) \mathrm{d}z \\
&\quad = \left(\mathscr{R}_{T/n} * p^X_{T-(k+1)T/n}\right)(x, y) + (T/n)^{m+1} \left(\Psi_{T/n} * p^X_{T-(k+1)T/n}\right)(x, y).
\end{aligned} \tag{F.11}$$

Hence, we obtain

$$\begin{aligned}
&\left(Q^{(m)}_{T/n}\right)^k \left(P_{T/n} - Q^{(m)}_{T/n}\right) p^X_{T-(k+1)T/n}(\cdot, y)|_{\cdot=x} \\
&\quad = \Bigg(\underbrace{p^{\bar{X},(m)}_{T/n} * \ldots * p^{\bar{X},(m)}_{T/n}}_{k\text{-times}} * \mathscr{R}_{T/n} * p^X_{T-(k+1)T/n}\Bigg)(x, y) + (T/n)^{m+1} \left(Q^{(m)}_{T/n}\right)^k \left(\Psi_{T/n} * p^X_{T-(k+1)T/n}\right)(x, y) \\
&\quad =: \mathcal{M}_1(x, y) + (T/n)^{m+1} \mathcal{M}_2(x, y), \quad (x, y) \in \mathbb{R}^N \times \mathbb{R}^N.
\end{aligned}$$

Moreover, we define a function $\mathcal{M}_3 : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ by

$$\mathcal{M}_3(x, y) = \left(Q^{(m)}_{T/n}\right)^{n-1} \left(p^X_{T/n}(\cdot, y) - p^{\bar{X},(m)}_{T/n}(\cdot, y)\right)|_{\cdot=x}, \quad (x, y) \in \mathbb{R}^N \times \mathbb{R}^N. \tag{F.12}$$

Since we are able to show that the terms $\mathcal{M}_1(x, y)$, $(T/n)^{m+1}\mathcal{M}_2(x, y)$ and $\mathcal{M}_3(x, y)$ are bounded by

$$\left(\frac{T}{n}\right)^{m+1} \frac{K(T)}{T^Q} \exp\left(-c\frac{|y-x|^2}{T}\right)$$

for some non-decreasing function $K(\cdot)$ and constants $c > 0$, $Q \geq N/2$, we have

$$\begin{aligned}
&\left| p^X_T(x, y) - E\left[p^{\bar{X},(m)}_{T/n}(\bar{X}^{(n)}((n-1)T/n, x), y) \prod_{j=1}^{n-1} \left\{1 + \pi^{(m),\bar{X}^{(n)}((j-1)T/n,x)}_{T/n} \left(B_{jT/n} - B_{(j-1)T/n}\right)\right\}\right]\right| \\
&\quad \leq \sum_{k=1}^{n} \left(\frac{T}{n}\right)^{m+1} \frac{K(T)}{T^Q} \exp\left(-c\frac{|y-x|^2}{T}\right) \\
&\quad = \left(\frac{T}{n}\right)^{m} \frac{K(T)}{T^Q} \exp\left(-c\frac{|y-x|^2}{T}\right).
\end{aligned}$$

Then, in what follows, we shall estimate the terms $\mathcal{M}_1(x, y)$, $\mathcal{M}_2(x, y)$ and $\mathcal{M}_3(x, y)$, using the key results; Lemmas F.1–F.3. □

F.1. Upper bound for $\mathcal{M}_1$

At first, we will give the upper bound for the first term of (F.11). Taking advantage of Lemma F.1 and Lemma F.2, we have

$$\begin{aligned}
&\left|\left(\mathscr{R}_{T/n} * p^X_{T-(k+1)T/n}\right)(x,y)\right| \qquad \text{(F.13)}\\
&\le C\frac{(T/n)^{m+1}}{(T/n)^{\frac{N}{2}}}\,\frac{1}{(T-(k+1)T/n)^{\frac{N}{2}}}\int_{\mathbb{R}^N}\exp\left(-c_1\frac{|z-x|^2}{T/n}\right)\exp\left(-c_2\frac{|y-z|^2}{T-(k+1)T/n}\right)\mathrm{d}z\\
&\le C\frac{(T/n)^{m+1}}{(T/n)^{\frac{N}{2}}}\,\frac{(T/n)^{\frac{N}{2}}}{(T-kT/n)^{\frac{N}{2}}}\exp\left(-c\frac{|y-x|^2}{T-kT/n}\right)\\
&= C\frac{(T/n)^{m+1}}{(T-kT/n)^{\frac{N}{2}}}\exp\left(-c\frac{|y-x|^2}{T-kT/n}\right),
\end{aligned}$$

where c_1, c_2, c and C are some positive constants.

Since we have for $k = 1,\dots,n-2$,

$$\begin{aligned}
&\Bigg(\underbrace{p^{\bar{X},(m)}_{T/n} * \dots * p^{\bar{X},(m)}_{T/n}}_{k\text{-times}}\Bigg)(x,\xi)\\
&={}_{-\infty}\left\langle \delta_\xi\left(\bar{X}^{(n)}(kT/n,x)\right), \prod_{j=1}^{k}\left(1+\pi^{(m),\bar{X}^{(n)}((j-1)T/n,x)}_{T/n}\left(B_{jT/n}-B_{(j-1)T/n}\right)\right)\right\rangle_\infty\\
&= E\left[T_\xi(\bar{X}^{(n)}(kT/n,x))H_\gamma\left(\bar{X}^{(n)}(kT/n,x), \prod_{j=1}^{k}\left(1+\pi^{(m),\bar{X}^{(n)}((j-1)T/n,x)}_{T/n}\left(B_{jT/n}-B_{(j-1)T/n}\right)\right)\right)\right],
\end{aligned} \qquad \text{(F.14)}$$

where $\gamma = (1,\dots,N)$ and $T_\xi(\cdot)$ is the Heaviside function introduced in Appendix E, it follows that

$$\left|\Bigg(\underbrace{p^{\bar{X},(m)}_{T/n} * \dots * p^{\bar{X},(m)}_{T/n}}_{k\text{-times}}\Bigg)(x,\xi)\right| \le \frac{C(T)}{(kT/n)^{\frac{N}{2}}}\exp\left(-c\frac{|\xi-x|^2}{kT/n}\right), \qquad \text{(F.15)}$$

where $C(T) > 0$ is a constant which does not depend on $k = 1,\dots,n-2$ by applying Lemma F.3 with $|\alpha| = 0$ to (F.14).

Then, using the bounds (F.13) and (F.15), we obtain for $k = 1,\dots,n-2$,

$$\begin{aligned}
|\mathcal{M}_1(x,y)| &= \left|\Bigg(\underbrace{p^{\bar{X},(m)}_{T/n} * \dots * p^{\bar{X},(m)}_{T/n}}_{k\text{-times}} * \mathscr{R}_{T/n} * p^X_{T-(k+1)T/n}\Bigg)(x,y)\right|\\
&\le \int_{\mathbb{R}^N}\left|\Bigg(\underbrace{p^{\bar{X},(m)}_{T/n} * \dots * p^{\bar{X},(m)}_{T/n}}_{k\text{-times}}\Bigg)(x,z)\left(\mathscr{R}_{T/n} * p^X_{T-(k+1)T/n}\right)(z,y)\right|\mathrm{d}z\\
&\le \frac{C(T)}{(kT/n)^{N/2}}\,\frac{(T/n)^{m+1}}{(T-kT/n)^{\frac{N}{2}}}\int_{\mathbb{R}^N}\exp\left(-c_1\frac{|z-x|^2}{kT/n}\right)\exp\left(-c_2\frac{|y-z|^2}{T-kT/n}\right)\mathrm{d}z
\end{aligned}$$

$$\le C(T)\frac{(T/n)^{m+1}}{(T-kT/n)^{\frac{N}{2}}}\left(\frac{T-kT/n}{T}\right)^{\frac{N}{2}}\exp\left(-c_3\frac{|y-x|^2}{T}\right)$$
$$= C(T)\frac{(T/n)^{m+1}}{T^{\frac{N}{2}}}\exp\left(-c_3\frac{|y-x|^2}{T}\right), \tag{F.16}$$

where $C(T)$, c_1, c_2 and c_3 are some positive constants (which are independent of k and n) and on the third inequality we applied Lemma F.1 again. When $k=0$, it is easy to see from the inequality (F.13) that $|\mathcal{M}_1(x,y)|$ has the same bound as (F.16).

F.2. Upper bound for $\mathcal{M}_2$

We will give the upper bound for $\left(Q^{(m)}_{T/n}\right)^k(\Psi_{T/n}*p^X_{T-(k+1)T/n})(\cdot,y)|_{\cdot=x}$, $k=0,1,2,\dots,n-2$. In particular, we proceed the derivation differently for the cases k is small or large. Due to the explicit form of $\Psi_{T/n}$ and the operator $Q_{T/n}$, we have

$$\left|\left(Q^{(m)}_{T/n}\right)^k\left(\Psi_{T/n}*p^X_{T-(k+1)T/n}\right)(\cdot,y)|_{\cdot=x}\right|$$
$$\le\sum_{l\le\nu}h_l\left(\frac{T}{n}\right)\left|E\left[\partial^{\alpha^{(l)}}p^X_{T-(k+1)T/n}\left(\bar{X}^{(n)}((k+1)T/n,x),y\right)G^{(k)}_l\right)\right]\right|$$

with some multi-indices $\alpha^{(l)}$, $l\le\nu\in\mathbb{N}$ and the random variable $G^{(k)}_l$ is given by

$$G^{(k)}_l=g_l\left(\bar{X}^{(n)}(kT/n,x)\right)\prod_{j=1}^{k}\left(1+\pi^{(m),\bar{X}^{(n)}((j-1)T/n,x)}_{T/n}\left(B_{jT/n}-B_{(j-1)T/n}\right)\right), \tag{F.17}$$

for some $g_l\in C^\infty_b(\mathbb{R}^N)$, which is the same one we defined in (2.59). Then it suffices to give the estimate for the above expectation.

F.2.1. The case $(k+1)T/n\in(0,T/2)$

First, it holds,

$$E\left[\partial^{\alpha^{(l)}}p^X_{T-(k+1)T/n}\left(\bar{X}^{(n)}((k+1)T/n,x),y\right)G^{(k)}_l\right]$$
$$=\int_{\mathbb{R}^N}\partial^{\alpha^{(l)}}p^X_{T-(k+1)T/n}(\xi,y)\,{}_{-\infty}\left\langle\delta_\xi\left(\bar{X}^{(n)}((k+1)T/n,x)\right),G^{(k)}_l\right\rangle_\infty\mathrm{d}\xi. \tag{F.18}$$

Applying the result in [8], we obtain

$$|\partial^{\alpha^{(l)}}p^X_{T-(k+1)T/n}(\xi,y)|\le\frac{C}{(T-(k+1)T/n)^{\frac{N+|\alpha^{(l)}|}{2}}}\exp\left(-c_1\frac{|y-\xi|^2}{T-(k+1)T/n}\right)$$

and by Lemma F.3 we also have

$$\left|{}_{-\infty}\left\langle\delta_\xi\left(\bar{X}^{(n)}((k+1)T/n,x)\right),G^{(k)}_l\right\rangle_\infty\right|=\left|E\left[T_\xi\left(\bar{X}^{(n)}((k+1)T/n,x)\right)H_\gamma\left(\bar{X}^{(n)}((k+1)T/n,x),G^{(k)}_l\right)\right]\right|$$
$$\le\frac{K(T)}{((k+1)T/n)^{\frac{N}{2}}}\exp\left(-c_2\frac{|\xi-x|^2}{(k+1)T/n}\right)$$

with some non-decreasing function $K(\cdot)$ and positive constants C, c_1 and c_2 which are independent of n. Hence, it follows that

$$\begin{aligned}
&\left|E\left[\partial^{\alpha^{(l)}} p^X_{T-(k+1)T/n}\left(\bar{X}^{(n)}((k+1)T/n,x),y\right) G^{(k)}_l\right]\right| \\
&\quad\le \frac{C}{(T-(k+1)T/n)^{\frac{N+|\alpha^{(l)}|}{2}}}\frac{K(T)}{((k+1)T/n)^{\frac{N}{2}}}\int_{\mathbb{R}^N}\exp\left(-c_1\frac{|y-\xi|^2}{T-(k+1)T/n}\right)\exp\left(-c_2\frac{|\xi-x|^2}{(k+1)T/n}\right)\mathrm{d}\xi \\
&\quad\le \frac{K'(T)}{(T-(k+1)T/n)^{\frac{N+|\alpha^{(l)}|}{2}}}\left(\frac{T-(k+1)T/n}{T}\right)^{\frac{N}{2}}\exp\left(-c\frac{|y-x|^2}{T}\right) \\
&\quad= \frac{K'(T)}{(T-(k+1)T/n)^{\frac{|\alpha^{(l)}|}{2}}}\frac{1}{T^{\frac{N}{2}}}\exp\left(-c\frac{|y-x|^2}{T}\right)
\end{aligned}$$

with some non-decreasing function $K'(\cdot)$ and constant $c>0$. In particular, we applied Lemma F.1 on the second inequality.

Since we assume $T/2 \le T-(k+1)T/n < T$, we conclude that

$$\left|E\left[\partial^{\alpha^{(l)}} p^X_{T-(k+1)T/n}\left(\bar{X}^{(n)}((k+1)T/n,x),y\right) G^{(k)}_l\right]\right| \le \frac{K(T)}{T^Q}\exp\left(-c\frac{|y-x|^2}{T}\right),$$

with some non-decreasing function $K(\cdot)$, some constants $c>0$ and $Q \ge \frac{N}{2}$ which are independent of n.

F.2.2. the case $(k+1)T/n \in [T/2, T)$

Applying the integration by parts on Wiener space, we obtain

$$\begin{aligned}
&\left|E\left[\partial^{\alpha^{(l)}} p^X_{T-(k+1)T/n}\left(\bar{X}^{(n)}((k+1)T/n,x),y\right) G^{(k)}_l\right]\right| \\
&\quad= \left|E\left[p^X_{T-(k+1)T/n}\left(\bar{X}^{(n)}((k+1)T/n,x),y\right) H_{\alpha^{(l)}}\left(\bar{X}^{(n)}((k+1)T/n,x),G^{(k)}_l\right)\right]\right| \\
&\quad= \left|\int_{\mathbb{R}^N} p^X_{T-(k+1)T/n}(\xi,y)\, {}_{-\infty}\left\langle \delta_\xi\left(\bar{X}^{(n)}((k+1)T/n,x)\right), H_{\alpha^{(l)}}\left(\bar{X}^{(n)}((k+1)T/n,x),G^{(k)}_l\right)\right\rangle_\infty \mathrm{d}\xi\right|.
\end{aligned}$$

Again, due to Lemma F.3 we have

$$\begin{aligned}
&\left|{}_{-\infty}\left\langle \delta_\xi\left(\bar{X}^{(n)}((k+1)T/n,x)\right), H_{\alpha^{(l)}}\left(\bar{X}^{(n)}((k+1)T/n,x),G^{(k)}_l\right)\right\rangle_\infty\right| \\
&\quad\le \frac{K(T)}{((k+1)T/n)^{\frac{N+|\alpha^{(l)}|}{2}}}\exp\left(-c\frac{|\xi-x|^2}{(k+1)T/n}\right)
\end{aligned}$$

and since $T/2 \le (k+1)T/n < T$, we obtain

$$\begin{aligned}
&\left|E\left[\partial^{\alpha^{(l)}} p^X_{T-(k+1)T/n}\left(\bar{X}^{(n)}((k+1)T/n,x),y\right) G^{(k)}_l\right]\right| \\
&\quad\le \frac{C}{(T-(k+1)T/n)^{\frac{N}{2}}}\frac{K(T)}{((k+1)T/n)^{\frac{N+|\alpha^{(l)}|}{2}}}\int_{\mathbb{R}^N}\exp\left(-c_1\frac{|y-\xi|^2}{T-(k+1)T/n}\right)\exp\left(-c_2\frac{|\xi-x|^2}{(k+1)T/n}\right)d\xi \\
&\quad\le \frac{K'(T)}{((k+1)T/n)^{\frac{N+|\alpha^{(l)}|}{2}}}\left(\frac{(k+1)T/n}{T}\right)^{\frac{N}{2}}\exp\left(-c\frac{|y-x|^2}{T}\right) \\
&\quad= \frac{K'(T)}{((k+1)T/n)^{\frac{|\alpha^{(l)}|}{2}}}\frac{1}{T^{\frac{N}{2}}}\exp\left(-c\frac{|y-x|^2}{T}\right) \\
&\quad\le \frac{K'(T)}{T^Q}\exp\left(-c\frac{|y-x|^2}{T}\right)
\end{aligned}$$

with some non-decreasing function $K'(\cdot)$ and constants $c>0,\ Q\geq N/2$ which are independent of n.

F.3. Upper bound for $\mathcal{M}_3$

By the small time expansion formula (F.3), we get

$$\begin{aligned}&\left(Q^{(m)}_{T/n}\right)^{n-1}(p^X_{T/n}(\cdot,y)-p^{\bar{X},(m)}_{T/n}(\cdot,y))|_{\cdot=x}\\&=\left(\underbrace{p^{\bar{X},(m)}_{T/n}*\ldots*p^{\bar{X},(m)}_{T/n}}_{(n-1)\text{-times}}*\mathscr{R}_{T/n}\right)(x,y)+(T/n)^{m+1}\left(Q^{(m)}_{T/n}\right)^{n-1}\Psi_{T/n}(\cdot,y)|_{\cdot=x}.\end{aligned}\tag{F.19}$$

For the first term of (F.19), using the bound (F.15) with $k=n-1$ and Lemma F.1, we obtain

$$\begin{aligned}&\left|\left(\underbrace{p^{\bar{X},(m)}_{T/n}*\ldots*p^{\bar{X},(m)}_{T/n}}_{(n-1)\text{-times}}*\mathscr{R}_{T/n}\right)(x,y)\right|\\&\leq\frac{K(T)}{((n-1)T/n)^{\frac{N}{2}}}\frac{1}{(T/n)^{\frac{N}{2}}}\left(\frac{T}{n}\right)^{m+1}\int_{\mathbb{R}^N}\exp\left(-c_1\frac{|\xi-x|^2}{(n-1)T/n}\right)\exp\left(-c_2\frac{|y-\xi|^2}{T/n}\right)\mathrm{d}\xi\\&\leq\left(\frac{T}{n}\right)^{m+1}\frac{K(T)}{((n-1)T/n)^{\frac{N}{2}}}\left(\frac{(n-1)T/n}{T}\right)^{\frac{N}{2}}\exp\left(-c\frac{|y-x|^2}{T}\right)\\&=\left(\frac{T}{n}\right)^{m+1}\frac{K(T)}{T^{\frac{N}{2}}}\exp\left(-c\frac{|y-x|^2}{T}\right),\end{aligned}\tag{F.20}$$

for some non-decreasing function $K(\cdot)$ and a constant $c>0$ which are independent of n.

Finally we consider the second term of (F.19). We notice that

$$\begin{aligned}\frac{\partial^l}{\partial x_{\alpha_1}\ldots\partial x_{\alpha_l}}p^{\bar{X}^z}_t(x,y)|_{z=x}&=\frac{\partial^l}{\partial x_{\alpha_1}\ldots\partial x_{\alpha_l}}{}_{-\infty}\left\langle\delta_y(\bar{X}^z(t,x)),1\right\rangle_\infty|_{z=x}\\&={}_{-\infty}\left\langle\delta_y\left(\bar{X}^z(t,x)\right),H_\alpha\left(\bar{X}^z(t,x),1\right)\right\rangle_\infty|_{z=x}\end{aligned}\tag{F.21}$$

for any multi-index $\alpha\in\{1,\ldots,N\}^l, l\in\mathbb{N}$ and $(t,x,y)\in[0,\infty)\times\mathbb{R}^N\times\mathbb{R}^N$. From now on, we briefly write $\bar{X}^z_t$ for $\bar{X}^z(t,x)$. Then, $\Psi_{T/n}(x,y)$ is given by

$$\Psi_{T/n}(x,y)=\sum_{l\leq\nu}h_l\left(\frac{T}{n}\right)g_l(x)_{-\infty}\left\langle\delta_y\left(\bar{X}^z_{T/n}\right),H_{\beta^{(l)}}\left(\bar{X}^z_{T/n},1\right)\right\rangle_\infty|_{z=x}\tag{F.22}$$

and hence we have

$$\begin{aligned}&\left(Q^{(m)}_{T/n}\right)^{n-1}\Psi_{T/n}(x,y)\\&=\sum_{l\leq\nu}h_l\left(\frac{T}{n}\right)E\left[{}_{-\infty}\left\langle\delta_y\left(\bar{X}^z_{T/n}\right),H_{\beta^{(l)}}(\bar{X}^z_{T/n},1)\right\rangle_\infty|_{z=\bar{X}^{(n)}((n-1)T/n,x)}G^{(n-1)}_l\right]\end{aligned}$$

where $\mathcal{F}_{(n-1)T/n}$-measurable random variable $G_l^{(n-1)}$ is the same one we defined in (E.5) on Appendix E. Since we have seen the equation (E.6) holds, we obtain

$$
\begin{aligned}
| \left(Q_{T/n}^{(m)}\right)^{n-1} \Psi_{T/n}(x,y)| &\leq \sum_{l\leq\nu} h_l\left(\frac{T}{n}\right)\left|E\left[\delta_y\left(\bar{X}^{(n)}(T,x)\right)H_{\beta^{(l)}}\left(\bar{X}^{(n)}(T,x),G_l^{(n-1)}\right)\right]\right| \\
&= \sum_{l\leq\nu} h_l\left(\frac{T}{n}\right)\left|E\left[T_y\left(\bar{X}^{(n)}(T,x)\right)H_{\beta^{(l)}*\gamma}\left(\bar{X}^{(n)}(T,x),G_l^{(n-1)}\right)\right]\right|
\end{aligned}
$$

where $\gamma = (1,2,\ldots,N)$. Applying Lemma F.3, we easily obtain

$$
\left|\left(Q_{T/n}^{(m)}\right)^{n-1} \Psi_{T/n}(x,y)\right| \leq \frac{K(T)}{T^Q}\exp\left(-c\frac{|y-x|^2}{T}\right) \tag{F.23}
$$

for some non-decreasing function $K(\cdot)$ and constants $c>0$, $Q\geq N/2$ which are independent of n.

Acknowledgements. We thank Prof. Shigeo Kusuoka (University of Tokyo) for advices on the approximation scheme. We are also grateful to the Editor, the Associate Editor and two anonymous reviewers for their valuable comments and suggestions. This work is supported by JSPS KAKENHI (Grant Number 16K13773, 19K13736), MEXT, Japan and a research fund from *Tokio Marine Kagami Memorial Foundation*.

References

[1] H.F. Baker, On the exponential theorem for a simply transitive continuous group, and the calculation of finite equations from the constants of structure. *Lond. M. S. Proc.* **34** (1901) 91–127.
[2] V. Bally and A. Kohatsu-Higa, A probabilistic interpretation of the parametrix method. *Ann. Appl. Probab.* **25** (2015) 3095–3138.
[3] V. Bally and D. Talay, The law of the Euler scheme for stochastic differential equations I. Convergence rate of the distribution function. *Probab. Theory Relat. Fields* **104** (1996) 43–60.
[4] V. Bally and D. Talay, The law of the Euler scheme for stochastic differential equations II. Convergence rate of the density, *Monte Carlo Methods App.* **2** (1996) 93–128.
[5] A. Bonfiglioli and R. Fulci, Topics in Noncommutative Algebra "The Theorem of Campbell, Baker, Hausdorff and Dynkin". Springer (2012).
[6] J.E. Campbell, On a law of combination of operators bearing on the theory of transformation groups. *Lond. M. S. Proc.* **28** (1897) 381–390.
[7] S. Ditlevsen and A. Samson, Hypoelliptic diffusions: discretization, filtering and inference from complete and partial observations. *J. R. Stat. Soc. B* **81** (2019) 361–384.
[8] A. Friedman, Partial Differential Equations of Parabolic Type. Prentice-Hall Inc (1964).
[9] E. Gobet and C. Labart, Sharp estimates for the convergence of the density of the Euler scheme in small time. *Electron. Commun. Probab.* **13** (2008) 352–363.
[10] J. Guyon, Euler schemes and tempered distributions. *Stochastic Processes App.* **116** (2006) 877–904.
[11] F. Hausdorff, Die symbolische Exponentialformel in der Gruppentheorie. *Leipz. Ber.* **58** (1906) 19–48.
[12] N. Ikeda and S. Watanabe Stochastic Differential Equations and Diffusion Processes, 2nd edition. North-Holland Mathematical Library (1989).
[13] V. Konakov and E. Mammen, Edgeworth type expansions for Euler schemes for stochastic differential equations. *Monte Carlo Methods App.* **8** (2002) 271–285.
[14] S. Kusuoka and D. Stroock, Applications of the Malliavin calculus Part I. In Vol. 32 of *Stochastic Analysis.* Katata/Kyoto 1982 (1984) 271–306.
[15] P.H. Labordére, X. Tan and N. Touzi, Unbiased simulation of stochastic differential equations. *Ann. Appl. Probab.* **27** (2017) 1–37.
[16] A. Melnykova, Parametric inference for multi-dimensional hypoelliptic ergodic diffusion with full observations. Preprint. Hal : hal-01704010, version 2 (2019).
[17] G. Maruyama, Continuous Markov processes and stochastic equations. *Rend. Circ. Mat. Palermo* **4** (1955) 48–90.
[18] D. Nualart, The Malliavin Calculus and Related Topics. Springer (2006).
[19] A.R. Pedersen, A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observation. *Scand. J. Stat.* **22** (1995) 55–71.

[20] A. Takahashi, Asymptotic expansion approach in finance, edited by P. Friz, J. Gatheral, A. Gulisashvili, A. Jacquier and J. Teichmann. In: Large Deviations and Asymptotic Methods in Finance. *Springer Proceedings in Mathematics & Statistics* (2015).

[21] A. Takahashi and T. Yamada, An asymptotic expansion with push-down of Malliavin weights. *SIAM J. Financial Math.* **3** (2012) 95–136.

[22] A. Takahashi and T. Yamada, A weak approximation with asymptotic expansion and multidimensional Malliavin weights. *Ann. Appl. Probab.* **26** (2016) 818–856.

[23] T. Yamada, An arbitrary high order weak approximation of SDE and Malliavin Monte Carlo: analysis of probability distribution functions. *SIAM J. Numer. Anal.* **57** (2019) 563–591.

[24] T. Yamada and K. Yamamoto, Second order discretization of Bismut-Elworthy-Li formula: application to sensitivity analysis. *SIAM/ASA J. Uncertainty Quantification* **7** (2019) 143–173.