Accelerated Exponential Euler Scheme for Stochastic Heat Equation: Convergence Rate of Densities

CHUCHU CHEN†

Academy of Mathematics and Systems Science, Chinese Academy of Sciences, School of Mathematical Sciences, University of Chinese Academy of Sciences, 100049 Beijing, China

JIANBO CUI‡

Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

AND

JIALIN HONG§, DERUI SHENG¶

Academy of Mathematics and Systems Science, Chinese Academy of Sciences, School of Mathematical Sciences, University of Chinese Academy of Sciences, 100049 Beijing, China

[Received on 30 May 2023]

This paper studies the numerical approximation of the density of the stochastic heat equation driven by space-time white noise via the accelerated exponential Euler scheme. The existence and smoothness of the density of the numerical solution are proved by means of the Malliavin calculus. Based on a priori estimates of the numerical solution, we present a test function-independent weak convergence analysis, which is crucial to show the convergence of the density. The convergence order of the density in uniform convergence topology is shown to be exactly $1/2$ in nonlinear drift case and nearly 1 in affine drift case. As far as we know, this is the first result on the existence and convergence of density of the numerical solution to the stochastic partial differential equation.

Keywords: density, convergence order, accelerated exponential Euler scheme, stochastic heat equation, Malliavin calculus

AMS subject classifications: 65C30, 60H35, 60H15, 60H07

1. Introduction

In this paper, we consider the numerical approximation of the density of the stochastic heat equation driven by space-time white noise:

$$\partial_t u(t,x) = \partial_{xx} u(t,x) + b(u(t,x)) + \sigma \dot{W}(t,x),\ (t,x) \in (0,T] \times [0,1] \tag{1.1}$$

with a deterministic initial value $u(0,x) = u_0(x),\ x \in [0,1]$ and Neumann boundary conditions $\partial_x u(t,0) = \partial_x u(t,1) = 0,\ t \in [0,T]$. Here, $T > 0$ is a fixed number, $\sigma > 0$ is the noise intensity and $\{W(t,x);(t,x) \in [0,T] \times [0,1]\}$ is a Brownian sheet defined on some complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Notice that all the results in the paper also hold for the case $\sigma < 0$, due to the symmetry of the Brownian sheet. Eq. (1.1) arises in many physical problems, and characterizes the evolution of a scalar field in a space-time-dependent random medium. The choice of the white noise as random potential corresponds to considering those regimes with very

†Email: chenchuchu@lsec.cc.ac.cn
‡Email: jianbo.cui@polyu.edu.hk; jianbocui@lsec.cc.ac.cn
§Email: hjl@lsec.cc.ac.cn
¶Corresponding author. Email: sdr@lsec.cc.ac.cn

rapid variations, the type of turbulent flows (see Bertini & Cancrini (1995)). The density function of the solution $u(t,x)$ for any $(t,x) \in (0,T] \times [0,1]$ characterizes all relevant probabilistic information, whose existence, regularity and strictly positivity under suitable assumptions have been well studied (e.g. Bally & Pardoux (1998); Mueller & Nualart (2008); Nualart & Quer-Sardanyons (2009)).

In the research on numerical approximations of stochastic partial differential equations (SPDEs), the existing works mainly focus on the strong convergence analysis (e.g. Gyöngy (1998, 1999); Yan (2005); Jentzen & Kloeden (2009); Cox & van Neerven (2010); Jentzen *et al.* (2011)) and weak convergence analysis (e.g. Debussche (2011); Andersson & Larsson (2016); Bréhier & Debussche (2018); Cui *et al.* (2021); Hong & Wang (2019); Cui & Hong (2019); Bréhier (2020)) of numerical schemes. It is of interest to study further the density function of the numerical solution which is highly related to the convergence analysis of a numerical scheme and may provide an appropriate approximation of the density of the original equation (see e.g. Bally & Talay (1996); Cui *et al.* (2019) for the case of stochastic ordinary differential equations). However, to the best of our knowledge, there are few results concerning the density function of the numerical solution for SPDEs. It is natural to ask:

Problem 1. Does the density of the numerical solution exist, and further is it smooth?

Problem 2. If so, how to estimate the error between the density of the numerical solution and that of the exact solution in uniform convergence topology?

Aiming to solve the above problems, we study the accelerated exponential Euler (AEE) scheme of Eq. (1.1). Introducing a uniform partition of $[0,T]$ with the temporal stepsize $\delta = T/N, N \in \mathbb{N}_0 := \mathbb{N} - \{0\}$, the numerical solution $U^{\delta,i+1}(x)$ of the AEE scheme is given by

$$\begin{aligned} U^{\delta,i+1}(x) = \int_0^1 G_\delta(x,y)U^{\delta,i}(y)\mathrm{d}y + \int_{t_i}^{t_{i+1}} \int_0^1 G_{t_{i+1}-s}(x,y)b(U^{\delta,i}(y))\mathrm{d}y\mathrm{d}s \\ + \int_{t_i}^{t_{i+1}} \int_0^1 G_{t_{i+1}-s}(x,y)\sigma W(\mathrm{d}s,\mathrm{d}y), \end{aligned} \tag{1.2}$$

for any $i \in \{0,1,\ldots,N-1\}$, and $U^{\delta,0}(x) = u_0(x), x \in [0,1]$, where $t_i = i\delta$ and $G_t(x,y)$ is the Green function associated with Neumann boundary conditions (see (2.3) for its expression). The strong convergence order of the AEE scheme in $L^2(\Omega;L^2(0,1))$ has been investigated in Jentzen & Kloeden (2009); Wang & Qi (2015). It is shown in Jentzen & Kloeden (2009) that the order is nearly 1 if the drift coefficient b is linear and in Wang & Qi (2015) that the order is nearly $1/2$ under less restrictive assumptions on b. The main contributions of this work are to prove the existence, smoothness of the density of the numerical solution $U^{\delta,N}(x)$ and to derive its convergence order.

First, we establish the non-degeneracy of the numerical solution, and prove the existence and smoothness of the corresponding density by means of Malliavin calculus. The major obstacle of this non-degeneracy lies in the estimates of the negative moments of the determinant of the corresponding Malliavin covariance matrix, which is overcome by proving a discrete comparison principle. To obtain the convergence order of density of the numerical solution, we use a test function-independent weak convergence result in the sense that

$$|\mathbb{E}[f(U^{\delta,N}(x)) - f(u(T,x))]| \leqslant C\delta^\mu$$

holds for some C independent of $f \in \Psi$ (see (3.3) for the definition of Ψ). One key ingredient for this test function-independent weak convergence analysis is the application of the Malliavin integration by parts formula (see Lemma 3.5). Another issue is that the moments of the Gateaux derivatives, as well as the Malliavin derivatives, of both $u(T,x)$ and $U^{\delta,N}(x)$ are dominated by the multiples of the Green function associated to Neumann boundary conditions, instead of being bounded by a constant in the case of stochastic ordinary differential equations (see e.g. Bally & Talay (1996)). Based on the technical estimates on the Green function, we obtain the weak convergence order $\mu = 1/2$, which removes the infinitesimal factor in the weak convergence order of the numerical scheme (see e.g. Debussche & Printems (2009)).

Combining the existence of a smooth density of the numerical solution and the test function-independent weak convergence analysis, we deduce that there exists $C > 0$ such that for any $x \in [0,1]$,

$$\|q^{\delta}_{N,x} - q_{T,x}\|_{L^{\infty}(\mathbb{R})} \leqslant C\delta^{\frac{1}{2}},$$

where $q^{\delta}_{N,x}$ and $q_{T,x}$ are the densities of $U^{\delta,N}(x)$ and $u(T,x)$, respectively. When b is affine, the above convergence order $1/2$ of density in uniform convergence topology can be improved to $1-\varepsilon$ for arbitrarily small $\varepsilon > 0$, which coincides with the strong convergence order $1-\varepsilon$ in Jentzen & Kloeden (2009). As far as we know, this is the first result on the convergence of density for numerical approximations to that of SPDEs. By further taking into account the uniform boundedness of $q^{\delta}_{N,x}$ in $L^1(\mathbb{R})$, it is concluded that $q^{\delta}_{N,x}$ converges to $q_{T,x}$ in $L^1(\mathbb{R})$, as δ tends to 0. This implies that the distribution of $U^{\delta,N}(x)$ converges to the distribution of $u(T,x)$ in total variation distance.

The paper is structured as follows. In Section 2, some notations and useful properties of the Green function and some elements of Malliavin calculus are introduced briefly. We present the main results and approach on the existence, smoothness, and convergence rate of the density of the numerical solution associated with the AEE scheme in Section 3. In Sections 4 and 5, some technique estimates concerning the regularity of the numerical solution are obtained. Finally, Section 6 is devoted to the proof of Proposition 3.3, based on which, the main result on the convergence order of the density follows.

2. Preliminaries

Notation: For any $x,y \in \mathbb{R}$, we denote $x \wedge y = \min\{x,y\}$ and $x \vee y = \max\{x,y\}$. Denote by E the Banach space $\mathbf{C}([0,1])$ endowed with the norm $\|h\|_E = \sup_{x\in[0,1]} |h(x)|$. For $v \in E$, we set $(G_t * v)(x) := \int_0^1 G_t(x,y)v(y)\mathrm{d}y$, $\forall\, t > 0, x \in [0,1]$. Let $\mathbf{C}^k_b$ be the set of all k times continuously differentiable functions with bounded derivatives from $\mathbb{R}$ to $\mathbb{R}$, and $\mathbf{C}^{\infty}_b := \bigcap_{k\geqslant 1} \mathbf{C}^k_b$. For $f \in \mathbf{C}^k_b$, denote $|f|_i := \sup_{x\in\mathbb{R}} |f^{(i)}(x)|$, $i \in \{1,\dots,k\}$. Hereafter, we use C to denote a generic positive constant that may change from one place to another and depend on several parameters but never on the stepsize δ.

In the sequel, without pointing it out explicitly, all equations hold almost surely (a.s.) or almost everywhere (a.e.). For $0 \leqslant t \leqslant T$, let $\mathscr{F}_t$ be the σ-field generated by $\{W(s,x);(s,x) \in [0,t]\times[0,1]\}$ and the $\mathbb{P}$-null sets. For $0 \leqslant s \leqslant t \leqslant T, x \in [0,1]$ and $v : \Omega \to E$ being $\mathscr{F}_s$-measurable, we denote by $\varphi^x_t(s,v)$ (resp. $\Phi^x_t(s,v)$) the exact flow of Eq. (1.1) (resp. numerical flow of the scheme (1.2)) at (t,x) starting from v at time s. More precisely,

$$\varphi^x_t(s,v) = (G_{t-s} * v)(x) + \int_s^t \int_0^1 G_{t-r}(x,z)b(\varphi^z_r(s,v))\mathrm{d}z\mathrm{d}r + \int_s^t \int_0^1 G_{t-r}(x,z)\sigma W(\mathrm{d}r,\mathrm{d}z), \tag{2.1}$$

and

$$\Phi^x_t(s,v) = (G_{t-s} * v)(x) + \int_s^t \int_0^1 G_{t-r}(x,z)b(\Phi^z_{\lfloor r\rfloor}(s,v))\mathrm{d}z\mathrm{d}r + \int_s^t \int_0^1 G_{t-r}(x,z)\sigma W(\mathrm{d}r,\mathrm{d}z) \tag{2.2}$$

with $\lfloor r\rfloor = t_k$, if $t_k < r \leqslant t_{k+1}$, $k \in \{0,\dots,N-1\}$.

In this section, we will also introduce several useful properties of the Green function, as well as some notations in the Malliavin calculus.

2.1 *Properties of Green function*

Recall the explicit expression of the Green function associated with Neumann boundary conditions

$$G_t(x,y) = \frac{1}{\sqrt{4\pi t}} \sum_{n=-\infty}^{+\infty} \left(e^{-\frac{(x-y-2n)^2}{4t}} + e^{-\frac{(x+y-2n)^2}{4t}} \right). \tag{2.3}$$

For any $x,y \in [0,1]$, the following properties will be used frequently (see (Bally & Pardoux, 1998, Appendix)):

$$(1) \qquad G_t(x,y) > 0 \quad \text{and} \quad (G_t * 1)(x) \equiv 1, \quad \forall\, t > 0. \tag{2.4}$$

$$(2) \qquad \langle G_t(x,\cdot), G_s(\cdot,y)\rangle_{L^2(0,1)} = G_{s+t}(x,y), \quad \forall\, s,t>0. \tag{2.5}$$

$$(3) \qquad \text{For some } K:=K(T)>0, \quad P_t(x,y)/K \leqslant G_t(x,y) \leqslant KP_t(x,y), \quad \forall\, t\in(0,T]. \tag{2.6}$$

Here, $P_t(x,y) = (4\pi t)^{-\frac{1}{2}}\exp\left(-\frac{(x-y)^2}{4t}\right)$ is the heat kernel on $\mathbb{R}$, and property (2.5) is the semigroup property of G. For any $x,y\in\mathbb{R}$, it is obvious that $P_t^2(x,y) = (8\pi t)^{-\frac{1}{2}}P_{t/2}(x,y)$ and $P_s(x,y) \leqslant \sqrt{t/s}P_t(x,y)$ provided $0<s\leqslant t$. The explicit formula of $G_t(x,y)$ is complicated, whose estimations could be converted into those of $P_t(x,y)$ thanks to (2.6). For instance, there exists $C=C(T)>0$ such that for any $x,y\in[0,1]$, $0<s\leqslant t\leqslant T$,

$$G_t^2(x,y) \leqslant Ct^{-\frac{1}{2}}G_{t/2}(x,y), \qquad G_s(x,y) \leqslant C\sqrt{t/s}G_t(x,y). \tag{2.7}$$

In particular, there is some $C(T)>0$ such that

$$\int_s^t\int_0^1 G_r^2(x,y)\mathrm{d}y\mathrm{d}r \leqslant C(t-s)^{\frac{1}{2}}, \quad \forall\, 0\leqslant s<t\leqslant T. \tag{2.8}$$

The following lemma gives the regularity in time of G.

Lemma 2.1 For any $\nu\in(1/3,1)$, there is $C=C(T,\nu)$ such that for any $0<s<t\leqslant T$,

$$\max\left(\int_0^1|G_t(x,y)-G_s(x,y)|\mathrm{d}x, \int_0^1|G_t(x,y)-G_s(x,y)|\mathrm{d}y\right) \leqslant Cs^{-\nu}(t-s)^{\nu}.$$

Proof. Similar to (Walsh, 1986, Corollary 3.4), the series expansion in (2.3) shows that $G_t(x,y) = P_t(x,y) + H_t(x,y)$ with $H_t(x,y)\in C^\infty([0,T]\times(0,1)^2)$. From (Mishura *et al.*, 2021, Corollary 2.2), we have

$$\max\left(\int_{\mathbb{R}}|P_t(x,y)-P_s(x,y)|\mathrm{d}x, \int_{\mathbb{R}}|P_t(x,y)-P_s(x,y)|\mathrm{d}y\right) \leqslant C(\nu)s^{-\nu}(t-s)^{\nu},$$

for any $\nu\in(1/3,1)$. Finally, the proof is completed by the facts that $H_t(x,y)\in C^\infty([0,T]\times(0,1)^2)$ and $|G_t(x,y)-G_s(x,y)| \leqslant |P_t(x,y)-P_s(x,y)| + |H_t(x,y)-H_s(x,y)|$. □

2.2 *Malliavin calculus*

Now we turn to a brief introduction of the Malliavin calculus (see e.g. Nualart (2006)). In the context of the Malliavin calculus, the isonormal Gaussian family $\{W(h), h\in\mathbb{H}\}$ corresponding to $\mathbb{H}:=L^2([0,T]\times[0,1])$ is given by the Wiener integral $W(h)=\int_0^T\int_0^1 h(s,y)W(\mathrm{d}s,\mathrm{d}y)$. We denote by $\mathscr{S}$ the class of smooth real-valued random variables of the form

$$X = g(W(h_1),\ldots,W(h_n)), \tag{2.9}$$

where $g\in\mathbf{C}_p^\infty(\mathbb{R}^n)$, $h_i\in\mathbb{H}$, $i=1,\ldots,n$, $n\geqslant 1$. Here, $\mathbf{C}_p^\infty(\mathbb{R}^n)$ is the space of all real-valued smooth functions on $\mathbb{R}^n$ whose partial derivatives have at most polynomial growths. The Malliavin derivative of $X\in\mathscr{S}$ of the form (2.9) is an $\mathbb{H}$-valued random variable given by $DX=\sum_{i=1}^n\partial_i g(W(h_1),\ldots,W(h_n))h_i$, which is also a random field $DX=\{D_{\theta,\xi}X,(\theta,\xi)\in[0,T]\times[0,1]\}$ with $D_{\theta,\xi}X=\sum_{i=1}^n\partial_i g(W(h_1),\ldots,W(h_n))h_i(\theta,\xi)$. Here, $D_{\theta,\xi}X$ is defined for a.e. $(\theta,\xi,\omega)\in[0,T]\times[0,1]\times\Omega$. For any $p\geqslant 1$, we denote the domain of D in $L^p(\Omega)$ by $\mathbb{D}^{1,p}$, meaning that $\mathbb{D}^{1,p}$ is the closure of $\mathscr{S}$ with respect to the norm $\|X\|_{1,p}=\left(\mathbb{E}\left[|X|^p+\|DX\|_{\mathbb{H}}^p\right]\right)^{\frac{1}{p}}$.

We define the iteration of the operator D in such a way that for $X\in\mathscr{S}$, the iterated derivative D^kX is a random variable with values in $\mathbb{H}^{\otimes k}$. More precisely, for $k\in\mathbb{N}_0$, $D^kX=\{D_{r_1,\theta_1}\cdots D_{r_k,\theta_k}X,(r_i,\theta_i)\in[0,T]\times[0,1]\}$ is a measurable function on the product space $([0,T]\times[0,1])^k\times\Omega$. Then for $p\geqslant 1$ and $k\in\mathbb{N}$, denote $\mathbb{D}^{k,p}$ the completion of $\mathscr{S}$ with respect to the norm

$$\|X\|_{k,p} = \left(\mathbb{E}\left[|X|^p+\sum_{j=1}^k\|D^jX\|_{\mathbb{H}^{\otimes j}}^p\right]\right)^{\frac{1}{p}}. \tag{2.10}$$

In particular, for $p\geqslant 1$, we simply write $\|X\|_p$ as an abbreviation for $\|X\|_{0,p}$. Define $L^{\infty-}(\Omega):=\bigcap_{p\geqslant 1}L^p(\Omega)$, $\mathbb{D}^{k,\infty}:=\bigcap_{p\geqslant 1}\mathbb{D}^{k,p}$, $\mathbb{D}^\infty:=\bigcap_{k\geqslant 1}\mathbb{D}^{k,\infty}$ to be topological projective limits. The following proposition is a Hölder inequality for the $\|\cdot\|_{k,p}$ norms, which implies that $\mathbb{D}^\infty$ is closed under multiplication.

Proposition 2.2 (Nualart, 2006, Proposition 1.5.6) Let $X \in \mathbb{D}^{k,p}$, $H \in \mathbb{D}^{k,q}$ for $k \in \mathbb{N}$, $1 < p < q < \infty$, and let r be such that $1/p + 1/q = 1/r$. Then, $XH \in \mathbb{D}^{k,r}$ and

$$\|XH\|_{k,r} \leqslant C(p,q,k)\|X\|_{k,p}\|H\|_{k,q}.$$

Consider the adjoint operator D^* of D. If an $\mathbb{H}$-valued random variable $\phi \in L^2(\Omega,\mathbb{H})$ satisfies $|\mathbb{E}[\langle \phi, DX\rangle_{\mathbb{H}}]| \leqslant C(\phi)\|X\|_{L^2(\Omega)}$, $\forall\, X \in \mathbb{D}^{1,2}$, then $\phi \in Dom(D^*)$ and $D^*(\phi) \in L^2(\Omega)$ is characterized by

$$\mathbb{E}[\langle \phi, DX\rangle_{\mathbb{H}}] = \mathbb{E}[XD^*(\phi)], \quad \forall\, X \in \mathbb{D}^{1,2}. \tag{2.11}$$

In the particular case that $\phi \in \mathbb{H}$ is deterministic, it holds that $\phi \in Dom(D^*)$ and $D^*\phi = W(\phi)$ (Nualart, 2006, Proposition 1.3.11). For a real-valued random variable $X \in \mathbb{D}^{1,2}$, we denote by $\Gamma_X := \langle DX, DX\rangle_{\mathbb{H}} = \|DX\|^2_{\mathbb{H}}$ the Malliavin covariance matrix of X.

Proposition 2.3 (Nualart, 2006, Theorem 2.1.4) Let X be a non-degenerate real-valued random variable, i.e., $X \in \mathbb{D}^\infty$, Γ_X is invertible almost surely (a.s.), and $\Gamma_X^{-1} \in L^{\infty-}(\Omega)$. Then X possesses a smooth density.

3. Main results and approach

As is shown in Bally & Pardoux (1998), for any $x \in [0,1]$, $u(T,x)$ admits a smooth density, and we refer interested readers to Sanz-Solé (2008); Cui & Hong (2020) and references therein for the study of densities of other SPDEs. However, as far as we know, there exists few result on the density of the numerical solution of the SPDE. In this section, we present our main results and approach on the existence, smoothness, and the convergence rate of the density of the numerical solution associated with the AEE scheme.

3.1 *Main results*

Our first main result is about the existence and smoothness of the density of the numerical solution $U^{\delta,N}(x)$.

Theorem 3.1 Assume that $b \in \mathbf{C}_b^\infty$, $\sigma > 0$. Then for every $x \in [0,1]$, $U^{\delta,N}(x)$ admits a smooth density $q^\delta_{N,x}$.

Based on Proposition 2.3, the proof of Theorem 3.1 boils down to showing the non-degeneracy of $U^{\delta,N}(x)$. The major obstacle of this non-degeneracy lies in the estimates of the negative moments of the determinant of the corresponding Malliavin covariance matrix, which is overcome by proving a discrete comparison principle. See Proposition 4.3 and Lemma 4.5 for the proofs that $U^{\delta,N}(x) \in \mathbb{D}^\infty$ and $\Gamma^{-1}_{U^{\delta,N}(x)} \in L^{\infty-}(\Omega)$, respectively.

The second main result of this paper is the convergence rate of the density of the numerical solution associated with the AEE scheme (1.2) for Eq. (1.1) in uniform convergence topology.

Theorem 3.2 Assume that $b \in \mathbf{C}_b^\infty$ and $\sigma > 0$. Then for any $x \in [0,1]$,

$$\|q^\delta_{N,x} - q_{T,x}\|_{L^\infty(\mathbb{R})} \leqslant C(T,\sigma,\|u_0\|_E)\delta^{\frac{1}{2}}.$$

In addition, if b is affine, then for any $\mu \in (\frac{1}{2},1)$ and $x \in [0,1]$,

$$\|q^\delta_{N,x} - q_{T,x}\|_{L^\infty(\mathbb{R})} \leqslant C(T,\sigma,\|u_0\|_E,\mu)\delta^{\mu}.$$

Recall that the total variation distance of probability measures $\tilde{\mu}_1$ and $\tilde{\mu}_2$ on a σ-algebra Π is defined by $d_{TV}(\tilde{\mu}_1,\tilde{\mu}_2) = 2\sup\{|\tilde{\mu}_1(A) - \tilde{\mu}_2(A)| : A \in \Pi\}$. Because $u(T,x)$ and $U^{\delta,N}(x)$ have smooth densities $q_{T,x}$ and $q^\delta_{N,x}$, respectively, it is readily to verify that the set $A = \{\mathrm{z} \in \mathbb{R} : q_{T,x}(\mathrm{z}) > q^\delta_{N,x}(\mathrm{z})\}$ attains the supremum of $\sup\{|\mathbb{P}(u(T,x) \in A) - \mathbb{P}(U^{\delta,N}(x) \in A)| : A \in \mathscr{B}(\mathbb{R})\}$, which leads to

$$d_{TV}(u(T,x)\circ\mathbb{P}^{-1}, U^{\delta,N}(x)\circ\mathbb{P}^{-1}) = \int_{\mathbb{R}} |q^\delta_{N,x}(\mathrm{z}) - q_{T,x}(\mathrm{z})|\mathrm{dz}.$$

Theorem 3.2, together with Scheffé lemma (Serfling, 1980, Chapter 1.5, Theorem C), also indicates that

$$\int_{\mathbb{R}} |q^\delta_{N,x}(\mathrm{z}) - q_{T,x}(\mathrm{z})|\mathrm{dz} \to 0, \quad \text{as } \delta \to 0.$$

Hence, the distribution of $U^{\delta,N}(x)$ converges to the distribution of $u(T,x)$ in total variation distance.

3.2 *Main approach of the proof of Theorem 3.2*

Our strategy is as follows. The existence and smoothness of densities of both the exact solution $u(T,x)$ and the numerical solution $U^{\delta,N}(x)$ implies

$$q^{\delta}_{N,x}(\mathrm{z}) = \lim_{n\to\infty}\int_{\mathbb{R}} g_{n^{-1}}(\mathrm{y}-\mathrm{z})q^{\delta}_{N,x}(\mathrm{y})\mathrm{dy} = \lim_{n\to\infty}\mathbb{E}[g_{n^{-1}}(U^{\delta,N}(x)-\mathrm{z})], \tag{3.1}$$

$$q_{T,x}(\mathrm{z}) = \lim_{n\to\infty}\int_{\mathbb{R}} g_{n^{-1}}(\mathrm{y}-\mathrm{z})q_{T,x}(\mathrm{y})\mathrm{dy} = \lim_{n\to\infty}\mathbb{E}[g_{n^{-1}}(u(T,x)-\mathrm{z})], \tag{3.2}$$

where $g_\zeta(\mathrm{y}-\mathrm{z}) = (2\pi\zeta)^{-\frac{1}{2}}\exp(-\frac{|\mathrm{y}-\mathrm{z}|^2}{2\zeta})$. It follows from (3.1) and (3.2) that for every $\mathrm{z}\in\mathbb{R}$ and $x\in[0,1]$,

$$|q^{\delta}_{N,x}(\mathrm{z}) - q_{T,x}(\mathrm{z})| = \lim_{n\to\infty}\left|\mathbb{E}[g_{n^{-1}}(U^{\delta,N}(x)-\mathrm{z})] - \mathbb{E}\left[g_{n^{-1}}(u(T,x)-\mathrm{z})\right]\right|.$$

In order to estimate the error between $q^{\delta}_{N,x}(\mathrm{z})$ and $q_{T,x}(\mathrm{z})$, we notice that $\{g_{n^{-1}}(\cdot-\mathrm{z})\}_{n\geqslant 1,\mathrm{z}\in\mathbb{R}}$ belongs to

$$\Psi := \{f:\mathbb{R}\to\mathbb{R} | f\in\mathbf{C}^\infty_p(\mathbb{R}),\ \exists\, F:\mathbb{R}\to\mathbb{R} \text{ such that } 0\leqslant F\leqslant 1 \text{ and } F'=f\}. \tag{3.3}$$

Then we consider the test function space Ψ and establish the following test function-independent weak convergence result of the AEE scheme (1.2), which yields that the error between $q^{\delta}_{N,x}$ and $q_{T,x}$ possesses the same convergence order.

Proposition 3.3 Let $b\in\mathbf{C}^\infty_b$ and $\sigma>0$. Then for any $x\in[0,1]$ and $f\in\Psi$,

$$\left|\mathbb{E}[f(U^{\delta,N}(x))] - \mathbb{E}\left[f(u(T,x))\right]\right| \leqslant C(T,\sigma,\|u_0\|_E)\delta^{\frac{1}{2}}. \tag{3.4}$$

In addition, if b is affine, then for any $\mu\in(\frac{1}{2},1)$, $x\in[0,1]$ and $f\in\Psi$,

$$\left|\mathbb{E}[f(U^{\delta,N}(x))] - \mathbb{E}\left[f(u(T,x))\right]\right| \leqslant C(T,\sigma,\|u_0\|_E,\mu)\delta^{\mu}. \tag{3.5}$$

Here, the constants C in (3.4) and (3.5) are independent of the test function f and the variable x.

Remark 3.4 In the study of convergence of densities of numerical methods, one usually needs to transform the convergence of densities into weak convergence or strong convergence of numerical methods. In the present work, our methods mainly based on the above test function-independent weak convergence result. We would like to mention that following the strategy used in Bally & Caramellino (2014), the estimate of $q^{\delta}_{N,x}-q_{T,x}$ can be also converted to that of $\|u(T,x)-U^{\delta,N}(x)\|_{3,p}$.

The proof of Proposition 3.3 relies on the weak error decomposition and an integration by parts formula. We first give a decomposition for the weak error $\mathbb{E}[f(u(T,x))]-\mathbb{E}[f(U^{\delta,N}(x))]$. Notice that by (2.2),

$$\begin{aligned}\Phi^y_{t_{i-1}}(0,u_0) =&(G_{t_{i-1}}*u_0)(y) + \int_0^{t_{i-1}}\int_0^1 G_{t_{i-1}-\theta}(y,\xi)b(\Phi^{\xi}_{\lfloor\theta\rfloor}(0,u_0))\mathrm{d}\xi\mathrm{d}\theta\\ &+\int_0^{t_{i-1}}\int_0^1 G_{t_{i-1}-\theta}(y,\xi)\sigma W(\mathrm{d}\theta,\mathrm{d}\xi),\quad i\in\{1,\dots,N\},\end{aligned} \tag{3.6}$$

and by (2.1), for $r>t_{i-1}$,

$$\begin{aligned}&\varphi^y_r(t_{i-1},\Phi_{t_{i-1}}(0,u_0)) = (G_{r-t_{i-1}}*\Phi_{t_{i-1}}(0,u_0))(y)\\ &+\int_{t_{i-1}}^r\int_0^1 G_{r-\theta}(y,\xi)b(\varphi^{\xi}_{\theta}(t_{i-1},\Phi_{t_{i-1}}(0,u_0)))\mathrm{d}\xi\mathrm{d}\theta + \int_{t_{i-1}}^r\int_0^1 G_{r-\theta}(y,\xi)\sigma W(\mathrm{d}\theta,\mathrm{d}\xi)\\ =&(G_r*u_0)(y) + \int_0^{t_{i-1}}\int_0^1 G_{r-\theta}(y,\xi)b(\Phi^{\xi}_{\lfloor\theta\rfloor}(0,u_0))\mathrm{d}\xi\mathrm{d}\theta + \int_0^{t_{i-1}}\int_0^1 G_{r-\theta}(y,\xi)\sigma W(\mathrm{d}\theta,\mathrm{d}\xi)\\ &+\int_{t_{i-1}}^r\int_0^1 G_{r-\theta}(y,\xi)b(\varphi^{\xi}_{\theta}(t_{i-1},\Phi_{t_{i-1}}(0,u_0)))\mathrm{d}\xi\mathrm{d}\theta + \int_{t_{i-1}}^r\int_0^1 G_{r-\theta}(y,\xi)\sigma W(\mathrm{d}\theta,\mathrm{d}\xi),\end{aligned} \tag{3.7}$$

where in the last step we have used (2.5) and (3.6). Then the one-step error between Eq. (1.1) and the AEE scheme (1.2) is divided into

$$
\begin{aligned}
&\varphi_{t_i}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))-\Phi_{t_i}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\\
&=\int_{t_{i-1}}^{t_i}\int_0^1 G_{t_i-r}(\cdot,y)\{b(\varphi_r^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0)))-b(\Phi_{t_{i-1}}^y(0,u_0))\}\mathrm{d}y\mathrm{d}r\\
&=\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1 G_{t_i-r}(\cdot,y)b'(Z_i^\beta(r,y))\left(\varphi_r^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))-\Phi_{t_{i-1}}^y(0,u_0)\right)\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\\
&=\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1 G_{t_i-r}(\cdot,y)b'(Z_i^\beta(r,y))\underbrace{\int_0^1\{G_r(y,\xi)-G_{t_{i-1}}(y,\xi)\}u_0(\xi)\mathrm{d}\xi}_{=:E^i_{initial\_u}(r,y)}\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\\
&\quad+\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1 G_{t_i-r}(\cdot,y)b'(Z_i^\beta(r,y))\underbrace{\int_0^{t_{i-1}}\int_0^1\{G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)\}b(\Phi^\xi_{\lfloor\theta\rfloor}(0,u_0))\mathrm{d}\xi\mathrm{d}\theta}_{=:E^i_{initial\_b}(r,y)}\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\\
&\quad+\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1 G_{t_i-r}(\cdot,y)b'(Z_i^\beta(r,y))\underbrace{\int_0^{t_{i-1}}\int_0^1\{G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)\}\sigma W(\mathrm{d}\theta,\mathrm{d}\xi)}_{=:E^i_{initial\_\sigma}(r,y)}\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\\
&\quad+\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1 G_{t_i-r}(\cdot,y)b'(Z_i^\beta(r,y))\underbrace{\int_{t_{i-1}}^{r}\int_0^1 G_{r-\theta}(y,\xi)b(\varphi^\xi_\theta(t_{i-1},\Phi_{t_{i-1}}(0,u_0)))\mathrm{d}\xi\mathrm{d}\theta}_{=:E^i_b(r,y)}\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\\
&\quad+\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1 G_{t_i-r}(\cdot,y)b'(Z_i^\beta(r,y))\underbrace{\int_{t_{i-1}}^{r}\int_0^1 G_{r-\theta}(y,\xi)\sigma W(\mathrm{d}\theta,\mathrm{d}\xi)}_{=:E^i_\sigma(r,y)}\mathrm{d}y\mathrm{d}r\mathrm{d}\beta,
\end{aligned}
\tag{3.8}
$$

where $Z_i^\beta(r,y):=\beta\varphi_r^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))+(1-\beta)\Phi^y_{t_{i-1}}(0,u_0)$. In the third identity of (3.8), we have used (3.7) and (3.6) to deal with $\varphi_r^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))$ and $\Phi^y_{t_{i-1}}(0,u_0)$, respectively. Introduce

$$
Y_i^\tau(y):=(1-\tau)\Phi^y_{t_i}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))+\tau\varphi^y_{t_i}(t_{i-1},\Phi_{t_{i-1}}(0,u_0)),\quad \tau\in[0,1],\tag{3.9}
$$

which is the convex combination of the numerical and exact flows at (t_i,y) starting from $\Phi_{t_{i-1}}(0,u_0)$ at time t_{i-1}.

Then we have the following telescoping sum

$$
\begin{aligned}
&\mathbb{E}[f(u(T,x))]-\mathbb{E}[f(U^{\delta,N}(x))]=\mathbb{E}[f(\varphi_T^x(0,u_0))]-\mathbb{E}[f(\Phi_T^x(0,u_0))]\\
&=\sum_{i=1}^N\left\{\mathbb{E}\left[f(\varphi_T^x(t_{i-1},\Phi_{t_{i-1}}(0,u_0)))\right]-\mathbb{E}[f(\varphi_T^x(t_i,\Phi_{t_i}(0,u_0)))]\right\}\\
&=\sum_{i=1}^N\mathbb{E}\left[\mathbb{E}\left[f(\varphi_T^x(t_i,\varphi_{t_i}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))))-f(\varphi_T^x(t_i,\Phi_{t_i}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))))\right]|\mathscr{F}_{t_i}\right]\\
&=\sum_{i=1}^N\mathbb{E}\left[\mathbb{E}[f(\varphi_T^x(t_i,\eta_i))-f(\varphi_T^x(t_i,\zeta_i))]\,|_{\eta_i=\varphi_{t_i}(t_{i-1},\Phi_{t_{i-1}}(0,u_0)),\zeta_i=\Phi_{t_i}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))}\right],
\end{aligned}
\tag{3.10}
$$

where $f\in\Psi$. The Gateaux derivative of $\varphi_t^x(s,\cdot)$ at $v\in E$ in the direction $h\in E$ is defined by

$$
\mathscr{D}^h\varphi_t^x(s,v):=\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\varphi_t^x(s,v+\varepsilon h)\big|_{\varepsilon=0}=\lim_{\varepsilon\to0}\frac{\varphi_t^x(s,v+\varepsilon h)-\varphi_t^x(s,v)}{\varepsilon}.
$$

Besides, the map $v\mapsto\varphi_t^x(s,v)$ is called Fréchet differentiable at $v\in E$ (see e.g. (Kesavan, 2020, Chapter 1)) if

$$
\lim_{\varepsilon\to0}\frac{|\varphi_t^x(s,v+h)-\varphi_t^x(s,v)-\langle\mathscr{D}\varphi_t^x(s,v),h\rangle|}{\|h\|_E}=0.
$$

It is known that if the Fréchet derivative exists, then $\mathscr{D}^h\varphi_t^x(s,v)=\langle\mathscr{D}\varphi_t^x(s,v),h\rangle$.

Based on (2.1) and the boundedness of b', one could verify that for each $h\in E$, the map $\varepsilon\mapsto\varphi_t^x(s,v+\varepsilon h)$ is a.s. continuous and satisfies

$$\sup_{x\in[0,1]}\left|\frac{\varphi_t^x(s,v+\varepsilon h)-\varphi_t^x(s,v)}{\varepsilon}-\mathscr{D}^h\varphi_t^x(s,v)\right|\leqslant\varepsilon|b|_2\frac{e^{2(t-s)|b|_1}}{4|b|_1}|h|_E^2,\tag{3.11}$$

where the Gateaux derivative $\mathscr{D}^h\varphi_t^x(s,v)$ satisfies that for $0\leqslant s<t\leqslant T$,

$$\mathscr{D}^h\varphi_t^x(s,v)=(G_{t-s}*h)(x)+\int_s^t\int_0^1 G_{t-r}(x,z)b'(\varphi_r^z(s,v))\mathscr{D}^h\varphi_r^z(s,v)\mathrm{d}z\mathrm{d}r,\quad a.s.$$

The estimate (3.11) implies that for each bounded set $B\subset E$, one has

$$\lim_{\varepsilon\to0}\frac{\varphi_t^x(s,v+\varepsilon h)-\varphi_t^x(s,v)}{\varepsilon}-\mathscr{D}^h\varphi_t^x(s,v)=0,$$

uniformly with respect to $h\in B$, which, along with (Marinelli & Scarpa, 2020, Lemma 2.1), indicates that $v\mapsto\varphi_t^x(s,v)$ is Fréchet differentiable. Hence, we also have that for $0\leqslant s<t\leqslant T$,

$$\langle\mathscr{D}\varphi_t^x(s,v),h\rangle=(G_{t-s}*h)(x)+\int_s^t\int_0^1 G_{t-r}(x,z)b'(\varphi_r^z(s,v))\langle\mathscr{D}\varphi_r^z(s,v),h\rangle\mathrm{d}z\mathrm{d}r,\quad a.s.\tag{3.12}$$

By the mean value theorem for Gateaux derivative and the chain rule for Fréchet derivative (Behmardi & Nayeri, 2008, Lemma 4.5), we have that for $\eta_i,\zeta_i\in E$,

$$\mathbb{E}\left[f(\varphi_T^x(t_i,\eta_i))-f(\varphi_T^x(t_i,\zeta_i))\right]=\int_0^1\mathbb{E}\left[f'(\varphi_T^x(t_i,\tau\eta_i+(1-\tau)\zeta_i))\langle\mathscr{D}\varphi_T^x(t_i,\tau\eta_i+(1-\tau)\zeta_i),\eta_i-\zeta_i\rangle\right]\mathrm{d}\tau.$$

Hence, it follows from (3.10) and the definition of Y_i^τ, and (3.8) that

$$\begin{aligned}
&\mathbb{E}\left[f(u(T,x))\right]-\mathbb{E}[f(U^{\delta,N}(x))]\\
=&\sum_{i=1}^N\int_0^1\mathbb{E}\left[\mathbb{E}\left[f'(\varphi_T^x(t_i,Y_i^\tau))\left\langle\mathscr{D}\varphi_T^x(t_i,Y_i^\tau),\varphi_{t_i}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))-\Phi_{t_i}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\right\rangle\right]|\mathscr{F}_{t_i}\right]\mathrm{d}\tau\\
=&\sum_{i=1}^N\int_0^1\mathbb{E}\left[f'(\varphi_T^x(t_i,Y_i^\tau))\left\langle\mathscr{D}\varphi_T^x(t_i,Y_i^\tau),\varphi_{t_i}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))-\Phi_{t_i}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\right\rangle\right]\mathrm{d}\tau\\
=&\sum_{i=1}^N\int_0^1\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1\mathbb{E}\left[f'(\varphi_T^x(t_i,Y_i^\tau))\left\langle\mathscr{D}\varphi_T^x(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\right\rangle b'(Z_i^\beta(r,y))\mathscr{E}^i(r,y)\right]\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\mathrm{d}\tau,
\end{aligned}\tag{3.13}$$

where $\mathscr{E}^i(r,y):=E^i_{initial\_u}(r,y)+E^i_{initial\_b}(r,y)+E^i_{initial\_\sigma}(r,y)+E^i_b(r,y)+E^i_\sigma(r,y)$. The error decomposition (3.13) is standard. However, due to the requirement that the constants C in Proposition 3.3 are independent of $f\in\Psi$, the classical estimates that $f'(\varphi_T^x(t_i,Y_i^\tau))$ is bounded by $|f|_1$ is not applicable to our case. Motivated by the work of Bally & Talay (1996), where the authors study the convergence rate of the density of the law of a small perturbation of the Euler–Maruyama method, the test function-independent estimates of terms on the right side of (3.13) will depend on the non-degeneracy of $\varphi_T^x(t_i,Y_i^\tau)$, since it is the prerequisite of the following Malliavin integration by parts formula.

Lemma 3.5 Let $\alpha\in\mathbb{N}$, $b\in\mathbf{C}_b^\infty$ and $\sigma>0$. If $G_1\in\mathbb{D}^\infty$ and $f\in\Psi$, then for any $i\in\{1,\ldots,N\}$, $x\in[0,1]$ and $\tau\in[0,1]$, there exists $H_{\alpha+1}(\varphi_T^x(t_i,Y_i^\tau),G_1)\in\mathbb{D}^\infty$ such that

$$\mathbb{E}[f^{(\alpha)}(\varphi_T^x(t_i,Y_i^\tau))G_1]=\mathbb{E}\left[F(\varphi_T^x(t_i,Y_i^\tau))H_{\alpha+1}(\varphi_T^x(t_i,Y_i^\tau),G_1)\right].\tag{3.14}$$

Moreover, there exists some constant $C=C(\alpha,T,\sigma,\|u_0\|_E)$ such that

$$|\mathbb{E}[f^{(\alpha)}(\varphi_T^x(t_i,Y_i^\tau))G_1]|\leqslant C\|G_1\|_{\alpha+1,2}.$$

Here, $f^{(\alpha)}(\varphi_T^x(t_i,Y_i^\tau))$ denotes the composition of the α-th derivative $f^{(\alpha)}$ of f and the random variable $\varphi_T^x(t_i,Y_i^\tau)$.

The proof of Lemma 3.5 relies on the non-degeneracy of $\varphi_T^x(t_i,Y_i^\tau)$, and is postponed to Section 4.

4. Technical estimates

In this section, we study the non-degeneracy property of $\{\varphi_T^x(t_i, Y_i^\tau)\}_{\{x\in[0,1],\, i\in\{1,\dots,N\},\, \tau\in[0,1]\}}$, which implies the non-degeneracy of $U^{\delta,N}(x)$ since $U^{\delta,N}(x) = \varphi_T^x(t_N, Y_N^\tau)$ with $\tau = 0$.

4.1 *Malliavin–Sobolev norm of* $\varphi_t^x(t_i, Y_i^\tau)$

In order to estimate the Malliavin–Sobolev norm $\|\cdot\|_{k,p}$ of $\varphi_t^x(t_i, Y_i^\tau)$, we need the following lemma.

Lemma 4.1 Let $\psi \in \mathbf{C}_b^k$ and $X \in \mathbb{D}^{k,\infty}$ for some $k \in \mathbb{N}_0$. Then $\psi(X) \in \mathbb{D}^{k,\infty}$. Moreover, for any integer $1 \leqslant \alpha \leqslant k$ and $p \geqslant 1$,

$$\|D^\alpha(\psi(X))\|_{L^p(\Omega,\mathbb{H}^{\otimes\alpha})} \leqslant C(\|D^\alpha X\|_{L^p(\Omega,\mathbb{H}^{\otimes\alpha})} + \|X\|_{\alpha-1,p\alpha}^\alpha + 1)$$

holds for some $C > 0$ depending on α, p and $|\psi|_i$, $1 \leqslant i \leqslant \alpha$. As a consequence,

$$\|\psi(X)\|_{\alpha,p} \leqslant C\|X\|_{\alpha,p} + C(\|X\|_{\alpha-1,p\alpha}^\alpha + \|\psi(X)\|_p + 1).$$

Proof. By (Sanz-Solé, 2005, (3.20)), for any integer $1 \leqslant \alpha \leqslant k$,

$$D^\alpha(\psi(X)) = \sum_{\ell=1}^{\alpha} \sum_{\mathscr{P}_\ell} c_\ell \psi^{(\ell)}(X) \prod_{i=1}^{\ell} D^{|I_i|} X,$$

where $\mathscr{P}_\ell$ is the set of partitions of $\{1,\dots,\alpha\}$ consisting of ℓ disjoint sets $I_1,\dots,I_\ell$, $\ell = 1,\dots,\alpha$, $|I_i|$ denotes the cardinal of the set I_i, and c_ℓ are positive constants. By $\psi \in \mathbf{C}_b^\infty$, we obtain

$$\|D^\alpha(\psi(X))\|_{L^p(\Omega,\mathbb{H}^{\otimes\alpha})} \leqslant C\|D^\alpha X\|_{L^p(\Omega,\mathbb{H}^{\otimes\alpha})} + C\sum_{\ell=2}^{\alpha}\sum_{\mathscr{P}_\ell} \|\prod_{i=1}^{\ell} D^{|I_i|}X\|_{L^p(\Omega,\mathbb{H}^{\otimes\alpha})}.$$

For the case of $2 \leqslant \ell \leqslant \alpha$, it holds that $|I_i| \leqslant \alpha - 1$ for $i \in \{1,\dots,\ell\}$, and

$$\|\prod_{i=1}^{\ell} D^{|I_i|}X\|_{L^p(\Omega,\mathbb{H}^{\otimes\alpha})} \leqslant \prod_{i=1}^{\ell} \|D^{|I_i|}X\|_{L^{p\ell}(\Omega,\mathbb{H}^{\otimes|I_i|})} \leqslant \prod_{i=1}^{\ell}\|X\|_{\alpha-1,p\ell} \leqslant \|X\|_{\alpha-1,p\alpha}^\alpha + 1,$$

which proves the first assertion. By further taking (2.10) into account, we obtain

$$\|\psi(X)\|_{\alpha,p} \leqslant C\|\psi(X)\|_p + C\sum_{j=1}^{\alpha}\|D^j(\psi(X))\|_{L^p(\Omega,\mathbb{H}^{\otimes j})} \leqslant C\|\psi(X)\|_p + C(\|X\|_{\alpha-1,p\alpha}^\alpha + 1 + \|X\|_{\alpha,p}).$$

The proof is completed. □

Lemma 4.2 Let $c_0, c_1 > 0$ and $g_s(t,x) \geqslant 0$ satisfy that for all $0 < s < t \leqslant T$ and $x \in [0,1]$,

$$g_s(t,x) \leqslant c_0 + c_1 \int_s^t \int_0^1 G_{t-r_1}(x,z_1) g_s(r_1,z_1)\mathrm{d}z_1\mathrm{d}r_1,$$

Then $g_s(t,x) \leqslant c_0 e^{c_1 T}$ for all $0 < s < t \leqslant T$ and $x \in [0,1]$.

Proof. Taking supremum over $x \in [0,1]$ and using (2.4), we obtain that for $0 < s < t \leqslant T$,

$$\sup_{x\in[0,1]} g_s(t,x) \leqslant c_0 + c_1 \sup_{x\in[0,1]} \int_s^t \int_0^1 G_{t-r_1}(x,z_1)\mathrm{d}z_1 \sup_{z_1\in[0,1]} g_s(r_1,z_1)\mathrm{d}r_1 \leqslant c_0 + c_1\int_s^t \sup_{z_1\in[0,1]} g_s(r_1,z_1)\mathrm{d}r_1.$$

Therefore, it follows from the Gronwall inequality that $\sup_{x\in[0,1]} g_s(t,x) \leqslant c_0 e^{c_1(t-s)} \leqslant c_0 e^{c_1 T}$. □

Based on Lemmas 4.1 and 4.2, we are ready to show that the Malliavin–Sobolev norms of $\varphi_t^x(t_i, Y_i^\tau)$ are uniformly bounded by some constant. In the sequel, the generic constant may depend on the supremum norm of derivatives of b.

Proposition 4.3 Assume that $b \in \mathbf{C}_b^\infty$. Then for any $k \in \mathbb{N}$, $p \geqslant 1$, there exists $C = C(k,p,T,\sigma,\|u_0\|_E)$ such that for any $\tau \in [0,1]$ and $i \in \{1,\dots,N\}$,

$$\|\Phi_{t_i}^y(0,u_0)\|_{k,p} + \|\varphi_t^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{k,p} \leqslant C, \quad \forall\, y \in [0,1],\ t \in (t_{i-1},T]. \tag{4.1}$$

$$\|\varphi_t^x(t_i,Y_i^\tau)\|_{k,p} \leqslant C, \quad \forall\, t \in [t_i,T],\ x \in [0,1]. \tag{4.2}$$

Proof. In this proof, we denote by $\mathscr{H}_M$ the property that (4.1) and (4.2) hold for all $p \in [1,\infty)$ and $k = M$. The proof is based on an induction argument on M.

We first prove $\mathscr{H}_0$. Let $i \in \{1,\dots,N\}$ be arbitrarily fixed. The Burkholder inequality and (2.8) give that

$$\left\|\int_0^{t_i}\int_0^1 G_{t_i-r}(y,z)\sigma W(\mathrm{d}r,\mathrm{d}z)\right\|_p \leqslant C(p,\sigma)\left(\int_0^{t_i}\int_0^1 G_{t_i-r}^2(y,z)\mathrm{d}z\mathrm{d}r\right)^{\frac{1}{2}} \leqslant C(p,T,\sigma), \quad \forall\, p \geqslant 1.$$

Therefore, by (3.6), (2.4), the assumption $u_0 \in E$ and the linear growth of b, for any $i = 1,\dots,N$,

$$\begin{aligned}\sup_{y\in[0,1]}\|\Phi_{t_i}^y(0,u_0)\|_p \leqslant& \|u_0\|_E + \sup_{y\in[0,1]}\left\|\int_0^{t_i}\int_0^1 G_{t_i-r}(y,z)\sigma W(\mathrm{d}r,\mathrm{d}z)\right\|_p \\ &+ C|b|_1\sum_{j=0}^{i-1}\int_{t_j}^{t_{j+1}}\sup_{y\in[0,1]}\int_0^1 G_{t_i-r}(y,z)\left(\sup_{z\in[0,1]}\|\Phi_{t_j}^z(0,u_0)\|_p + 1\right)\mathrm{d}z\mathrm{d}r \\ \leqslant& C(T,p,\|u_0\|_E,\sigma,|b|_1) + C|b|_1\sum_{j=0}^{i-1}\delta\sup_{z\in[0,1]}\|\Phi_{t_j}^z(0,u_0)\|_p.\end{aligned}$$

Utilizing the discrete Gronwall lemma produces

$$\sup_{y\in[0,1]}\|\Phi_{t_i}^y(0,u_0)\|_p \leqslant C, \quad \forall\, i = 1,\dots,N. \tag{4.3}$$

Similarly, it follows from the first identity in (3.7), the linear growth of b, the Burkholder and Minkowski inequalities that

$$\begin{aligned}\|\varphi_t^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_p \leqslant& \int_0^1 G_{t-t_{i-1}}(y,z)\|\Phi_{t_{i-1}}^z(0,u_0)\|_p\mathrm{d}z + C(p,\sigma)\left(\int_{t_{i-1}}^t\int_0^1 G_{t-r}^2(y,z)\mathrm{d}z\mathrm{d}r\right)^{\frac{1}{2}} \\ &+ C|b|_1\int_{t_{i-1}}^t\int_0^1 G_{t-r}(y,z)(1+\|\varphi_r^z(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_p)\mathrm{d}z\mathrm{d}r, \\ \leqslant& C(p,\sigma,|b|_1,T) + C|b|_1\int_{t_{i-1}}^t\int_0^1 G_{t-r}(y,z)\|\varphi_r^z(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_p\mathrm{d}z\mathrm{d}r,\end{aligned}$$

where the second inequality holds due to (4.3), (2.4), and (2.8). By Lemma 4.2, we obtain that for $t \in (t_{i-1},T]$,

$$\sup_{y\in[0,1]}\|\varphi_t^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_p \leqslant C, \quad \forall\, i = 1,\dots,N, \tag{4.4}$$

which together with (4.3) implies that (4.1) holds for $k = 0$. Similar to the proof of (4.4), it can be shown that (4.2) holds for $k = 0$ as well. Thus, we have proved $\mathscr{H}_0$.

Now we assume $\mathscr{H}_{M-1}$ and aim to show $\mathscr{H}_M$, $M \geqslant 1$. Notice that for $X \in \mathbb{D}^{M,p}$,

$$\|X\|_{M,p}^p = \|X\|_{M-1,p}^p + \|D^M X\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}^p. \tag{4.5}$$

By (3.6), the Minkowski inequality, Lemma 4.1 and $\mathscr{H}_{M-1}$, for $p \geqslant 1$,

$$\|D^M\Phi_{t_i}^y(0,u_0)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})} \leqslant \sum_{j=0}^{i-1}\int_{t_j}^{t_{j+1}}\int_0^1 G_{t_i-r}(y,z)\|D^M b(\Phi_{t_j}^z(0,u_0))\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z\mathrm{d}r$$

$$+\left\|D^M\int_0^{t_i}\int_0^1 G_{t_i-r}(y,z)\sigma W(\mathrm{d}r,\mathrm{d}z)\right\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\cdot \mathbf{1}_{\{M=1\}}$$
$$\leqslant C+C\sum_{j=0}^{i-1}\int_{t_j}^{t_{j+1}}\int_0^1 G_{t_i-r}(y,z)\|D^M\Phi^z_{t_j}(0,u_0)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z\mathrm{d}r, \tag{4.6}$$

since $D^M\int_0^{t_i}\int_0^1 G_{t_i-r}(y,z)\sigma W(\mathrm{d}r,\mathrm{d}z)$ vanishes for all $M\geqslant 2$, where $\mathbf{1}_{\{M=1\}}=1$ if $M=1$, otherwise $\mathbf{1}_{\{M=1\}}=0$. In the second step of (4.6), we have used the fact that $D_{\theta,\xi}\int_0^{t_i}\int_0^1 G_{t_i-r}(y,z)\sigma W(\mathrm{d}r,\mathrm{d}z)=G_{t_i-\theta}(y,\xi)\sigma$ for $\theta\in(0,t_i),\xi\in(0,1)$, and $D_{\theta,\xi}\int_0^{t_i}\int_0^1 G_{t_i-r}(y,z)\sigma W(\mathrm{d}r,\mathrm{d}z)=0$ for $\theta\in(t_i,T),\xi\in(0,1)$, and then by (2.8),

$$\left\|D\int_0^{t_i}\int_0^1 G_{t_i-r}(y,z)\sigma W(\mathrm{d}r,\mathrm{d}z)\right\|_{L^p(\Omega,\mathbb{H})}\leqslant \sigma\left(\int_0^{t_i}\int_0^1 G^2_{t_i-\theta}(y,\xi)\mathrm{d}\xi\mathrm{d}\theta\right)^{\frac{1}{2}}\leqslant C(p,\sigma,T). \tag{4.7}$$

By taking supremum over $y\in[0,1]$ and applying the discrete Gronwall lemma, we obtain

$$\sup_{y\in[0,1]}\|D^M\Phi^y_{t_i}(0,u_0)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\leqslant C,\quad \forall\, i\in\{1,\dots,N\}. \tag{4.8}$$

In the same way, by (4.8) and (2.1), we also have

$$\begin{aligned}\|D^M\varphi^y_t(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\leqslant{}& C+\int_0^1 G_{t-t_{i-1}}(y,z)\|D^M\Phi^z_{t_{i-1}}(0,u_0)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z\\ &+C\int_{t_{i-1}}^t\int_0^1 G_{t-r}(y,z)\|D^M\varphi^z_r(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z\mathrm{d}r\\ \leqslant{}& C+C\int_{t_{i-1}}^t\int_0^1 G_{t-r}(y,z)\|D^M\varphi^z_r(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z\mathrm{d}r,\end{aligned}$$

which together with Lemma 4.2 produces

$$\sup_{y\in[0,1]}\|D^M\varphi^y_t(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\leqslant C,\quad \forall\, t\in[t_{i-1},T],\ i\in\{1,\dots,N\}. \tag{4.9}$$

This along with (4.5), (4.8) and $\mathscr{H}_{M-1}$ completes the proof of (4.1) for $k=M$.

Moreover, (4.8), (4.9) and (3.9) also imply that for any $\tau\in[0,1]$,

$$\sup_{y\in[0,1]}\|D^M Y^\tau_i(y)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\leqslant\sup_{y\in[0,1]}\Big(\|D^M\Phi^y_{t_i}(0,u_0)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}+\|D^M\varphi^y_t(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\Big)\leqslant C.$$

Since $D^M\int_{t_i}^t\int_0^1 G_{t-s}(x,y)\sigma W(\mathrm{d}s,\mathrm{d}y)$ vanishes for all $M\geqslant 2$, by (2.1), it holds for $p\geqslant 1$ that

$$\begin{aligned}&\|D^M\varphi^x_t(t_i,Y^\tau_i)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\leqslant\int_0^1 G_{t-t_i}(x,y)\|D^M Y^\tau_i(y)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}y\\ &+\left\|D^M\int_{t_i}^t\int_0^1 G_{t-s}(x,y)\sigma W(\mathrm{d}s,\mathrm{d}y)\right\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\cdot\mathbf{1}_{\{M=1\}}+\int_{t_i}^t\int_0^1 G_{t-s}(x,y)\|D^M b(\varphi^y_s(t_i,Y^\tau_i))\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}y\mathrm{d}s,\end{aligned}$$

where the second term on the right side is estimated similarly as in (4.7), and is bounded by $C(p,\sigma)(t-t_i)^{\frac{1}{4}}\leqslant C(p,\sigma,T)$, due to (2.8). Using Lemma 4.1 and $\mathscr{H}_{M-1}$ yields

$$\|D^M b(\varphi^y_s(t_i,Y^\tau_i))\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\leqslant C\|D^M\varphi^y_s(t_i,Y^\tau_i)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}+C,$$

from which we get

$$\|D^M\varphi^x_t(t_i,Y^\tau_i)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\leqslant C+C\int_{t_i}^t\int_0^1 G_{t-s}(x,y)\|D^M\varphi^y_s(t_i,Y^\tau_i)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}y\mathrm{d}s.$$

This gives us that $\sup_{x\in[0,1]}\|D^M\varphi^x_t(t_i,Y^\tau_i)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\leqslant C$, thanks to Lemma 4.2. Then this estimate, (4.5) and $\mathscr{H}_{M-1}$ implies (4.2) for $k=M$. The proof is completed. □

The following corollary is a consequence of Proposition 4.3 and Lemma 4.1.

Corollary 4.4 Assume that $b \in \mathbf{C}_b^\infty$. Then for any $k \in \mathbb{N}$, $p \geqslant 1$, there exists $C = C(k,p,T,\sigma,\|u_0\|_E)$ such that for any τ, $\beta \in [0,1]$ and $i \in \{1,\dots,N\}$,

$$\text{(i)} \quad \sup_{\theta_1 \in (t_i,T],\, z\in[0,1]} \Big(\|b'(\varphi^z_{\theta_1}(t_i,Y_i^\tau))\|_{k,p} + \|b''(\varphi^z_{\theta_1}(t_i,Y_i^\tau))\|_{k,p} \Big) \leqslant C,$$

$$\text{(ii)} \quad \sup_{r \in (t_{i-1},t_i],\, y\in[0,1]} \Big(\|b'(Z_i^\beta(r,y))\|_{k,p} + \|b''(Z_i^\beta(r,y))\|_{k,p} \Big) \leqslant C,$$

$$\text{(iii)} \quad \sup_{\theta \in (0,t_{i-1}],\, \xi\in[0,1]} \|b(\Phi^\xi_{\lfloor\theta\rfloor}(0,u_0))\|_{k,p} + \sup_{\theta \in (t_{i-1},t_i],\, \xi\in[0,1]} \|b(\varphi^\xi_\theta(t_{i-1},\Phi_{t_{i-1}}(0,u_0)))\|_{k,p} \leqslant C.$$

Proof. Based on Proposition 4.3, we have that for any $k \in \mathbb{N}$, $p \geqslant 1$,

$$\|\varphi^z_{\theta_1}(t_i,Y_i^\tau)\|_{k,p} + \|\Phi^\xi_{\lfloor\theta\rfloor}(0,u_0)\|_{k,p} + \|\varphi^\xi_\theta(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{k,p} \leqslant C(k,p,T,\sigma,\|u_0\|_E).$$

Besides, by $Z_i^\beta(r,y) = \beta\varphi^y_r(t_{i-1},\Phi_{t_{i-1}}(0,u_0)) + (1-\beta)\Phi^y_{t_{i-1}}(0,u_0)$, we also have $\|Z_i^\beta(r,y)\|_{k,p} \leqslant C(k,p,T,\sigma,\|u_0\|_E)$. Notice that b', b'', as well as their derivatives of any order, are bounded. Hence, applying Lemma 4.1 with $\psi = b'$ (or $\psi = b''$), and $X = \varphi^z_{\theta_1}(t_i,Y_i^\tau)$ yields (i). Similarly, (ii) follows from Lemma 4.1 with $X = Z_i^\beta(r,y)$ and $\psi = b'$ (or $\psi = b''$). Finally, by Lemma 4.1 with $X = \Phi^\xi_{\lfloor\theta\rfloor}(0,u_0)$ and $\psi = b$, we arrive at

$$\begin{aligned}\|b(\Phi^\xi_{\lfloor\theta\rfloor}(0,u_0))\|_{k,p} &\leqslant C\|\Phi^\xi_{\lfloor\theta\rfloor}(0,u_0)\|_{k,p} + C(\|\Phi^\xi_{\lfloor\theta\rfloor}(0,u_0)\|^k_{k-1,pk} + \|b(\Phi^\xi_{\lfloor\theta\rfloor}(0,u_0))\|_p + 1)\\ &\leqslant C\|\Phi^\xi_{\lfloor\theta\rfloor}(0,u_0)\|_{k,p} + C(\|\Phi^\xi_{\lfloor\theta\rfloor}(0,u_0)\|^k_{k-1,pk} + \|\Phi^\xi_{\lfloor\theta\rfloor}(0,u_0)\|_p + 1) \leqslant C,\end{aligned}$$

since b is of linear growth. Analogous arguments also give $\|b(\varphi^\xi_\theta(t_{i-1},\Phi_{t_{i-1}}(0,u_0)))\|_{k,p} \leqslant C$. The proof is completed. □

4.2 *Negative moments estimates*

We first show the negative moments estimates for the numerical solution $U^{\delta,N}(x)$, which together with Proposition 4.3 proves Theorem 3.1.

Lemma 4.5 Assume that $b \in \mathbf{C}_b^1$ and $\sigma > 0$. Denote $M_i(\theta,y) := \int_0^1 D_{\theta,\xi}\Phi^y_{t_i}(0,u_0)\mathrm{d}\xi$. Then for each $i \in \{1,\dots,N\}$,

$$M_i(\theta,y) \geqslant \frac{1}{2}\sigma, \quad y \in [0,1],\ \theta \in (t_k,t_{k+1}) \text{ with } 0 \vee (i-1-\frac{\log\frac{3}{2}}{|b|_1\delta}) \leqslant k \leqslant i-1. \tag{4.10}$$

In particular, $(\Gamma_{U^{\delta,N}(x)})^{-1} \in L^{\infty-}(\Omega)$.

Proof. We first give an iteration formula for the Malliavin derivative for the numerical solution $\Phi^y_{t_i}(0,u_0)$. More precisely, by (3.6), for $\theta \in (t_k,t_{k+1})$ with $0 \leqslant k \leqslant i-1$, $y,\xi \in [0,1]$, and $i \in \{1,\dots,N\}$,

$$D_{\theta,\xi}\Phi^y_{t_i}(0,u_0) = \sum_{j=k+1}^{i-1}\int_{t_j}^{t_{j+1}}\int_0^1 G_{t_i-r}(y,z)b'(\Phi^z_{t_j}(0,u_0))D_{\theta,\xi}\Phi^z_{t_j}(0,u_0)\mathrm{d}z\mathrm{d}r + G_{t_i-\theta}(y,\xi)\sigma, \tag{4.11}$$

where we have used $D_{\theta,\xi}\Phi^y_{t_j}(0,u_0) = 0$ if $\theta > t_j$. Then by (2.4), for $\theta \in (t_k,t_{k+1})$ with $0 \leqslant k \leqslant i-1$, $y \in [0,1]$, and $i \in \{1,\dots,N\}$, $M_i(\theta,y)$ satisfies the following recursive relation

$$M_i(\theta,y) = \sum_{j=k+1}^{i-1}\int_{t_j}^{t_{j+1}}\int_0^1 G_{t_i-r}(y,z)b'(\Phi^z_{t_j}(0,u_0))M_j(\theta,z)\mathrm{d}z\mathrm{d}r + \sigma. \tag{4.12}$$

To get a lower bound for $M_i(\theta,y)$, we prove a discrete comparison principle. Define a two-parameter non-negative sequence $\{A_i^k\}_{0\leqslant k<i\leqslant N}$ by for any $i\in\{1,\ldots,N\}$, $A_i^{i-1}=\sigma$ and for any $0\leqslant k\leqslant i-2$,

$$A_i^k=\sum_{j=k+1}^{i-1}\int_{t_j}^{t_{j+1}}\int_0^1 G_{t_i-r}(y,z)|b|_1A_j^k\mathrm{d}z\mathrm{d}r+\sigma=\sum_{j=k+1}^{i-1}|b|_1\delta A_j^k+\sigma. \tag{4.13}$$

By (4.13), if $i_1-k_1=i_2-k_2$, then $A_{i_1}^{k_1}=A_{i_2}^{k_2}=:\mathscr{A}_{i_1-k_1}$. Rearranging (4.13), we derive that

$$\mathscr{A}_{i-k}=|b|_1\delta\mathscr{A}_{i-1-k}+\sum_{j=k+1}^{i-2}|b|_1\delta\mathscr{A}_{j-k}+\sigma=|b|_1\delta\mathscr{A}_{i-1-k}+\mathscr{A}_{i-1-k}=(1+|b|_1\delta)^{i-k-1}\sigma.$$

We claim that

$$|M_i(\theta,y)|\leqslant A_i^k,\quad\forall\,\theta\in(t_k,t_{k+1}),\;0\leqslant k\leqslant i-1,\;y\in[0,1]. \tag{4.14}$$

Indeed, we can prove (4.14) by an induction argument on $i-k$. For $i-k=1$, from (4.12), one has $M_i(\theta,y)=A_i^{i-1}=\sigma$, $\forall\,y\in[0,1]$, $\theta\in(t_{i-1},t_i)$ and $i\in\{1,\ldots,N\}$. Assume by induction that (4.14) holds for all integer i,k satisfying $1\leqslant i-k\leqslant i-k'-1$ $(k'\leqslant i-2)$. Now we show (4.14) for $i-k=i-k'$. Indeed, by the induction assumption $|M_j(\theta,z)|\leqslant A_j^{k'}$ for $k'+1\leqslant j\leqslant i-1$, (4.12), and (4.13),

$$\begin{aligned}|M_i(\theta,y)|&\leqslant\sum_{j=k'+1}^{i-1}\int_{t_j}^{t_{j+1}}\int_0^1 G_{t_i-r}(y,z)|b'(\Phi_{t_j}^z(0,u_0))||M_j(\theta,z)|\mathrm{d}z\mathrm{d}r+\sigma\\&\leqslant\sum_{j=k'+1}^{i-1}\int_{t_j}^{t_{j+1}}\int_0^1 G_{t_i-r}(y,z)|b|_1A_j^{k'}\mathrm{d}z\mathrm{d}r+\sigma=A_i^{k'}\end{aligned}$$

for $\theta\in(t_{k'},t_{k'+1})$, which implies (4.14) for $i-k=i-k'$. Thus we obtain (4.14).

By (4.12), (4.14) and (2.4), we deduce that if $|b|_1>0$, then

$$M_i(\theta,y)\geqslant\sigma-|b|_1\delta(A_{i-1}^k+A_{i-2}^k+\cdots+A_{k+1}^k)=\Big(2-(1+|b|_1\delta)^{i-k-1}\Big)\sigma\geqslant\frac12\sigma,$$

provided $0\leqslant i-k-1\leqslant(\log\frac32)/(\log(1+|b|_1\delta))$. By the relationship $0<\log(1+x)\leqslant x$, $\forall\,x>0$, we obtain (4.10). Obviously, if $|b|_1=0$, i.e., $b'\equiv0$, (4.10) is valid as well. Hence we have obtained the lower bound $\frac12\sigma$ of $M_i(\theta,y)$ when θ belongs to the interval $(t_{k_0^i},t_i)$, where $k_0^i=0\vee[i-\frac{K_b}{\delta}]$ with $K_b:=\frac{\log\frac32}{|b|_1}>0$ and $[\cdot]$ being the floor function. Since $t_i-t_{k_0^i}=\min\{t_i,t_i-\delta[i-\frac{K_b}{\delta}]\}\geqslant t_i\wedge K_b$, the length of the interval $(t_{k_0^i},t_i)$ is not smaller than $t_i\wedge K_b$, which is sufficient to derive a uniform lower bound independent of δ for $\Gamma_{U^{\delta,N}(x)}$. More precisely, by the Cauchy–Schwarz inequality and (4.10), we obtain

$$\begin{aligned}\Gamma_{U^{\delta,N}(x)}=&\Gamma_{\Phi_{t_N}^y(0,u_0)}:=\int_0^T\int_0^1|D_{\theta,\xi}\Phi_{t_N}^x(0,u_0)|^2\mathrm{d}\xi\mathrm{d}\theta\geqslant\int_0^T\left(\int_0^1 D_{\theta,\xi}\Phi_{t_N}^x(0,u_0)\mathrm{d}\xi\right)^2\mathrm{d}\theta\\&=\int_0^T|M_N(\theta,x)|^2\mathrm{d}\theta\geqslant\int_{t_{k_0^N}}^{t_N}|M_N(\theta,x)|^2\mathrm{d}\theta\geqslant\frac{\sigma^2}{4}(T\wedge K_b).\end{aligned}$$

This shows $(\Gamma_{U^{\delta,N}(x)})^{-1}\in L^{\infty-}(\Omega)$, and the proof is completed. □

We consider in Proposition 4.7 the uniform positive lower bound, independent of the sample ω and stepsize $\delta>0$, of the Malliavin covariance matrix $\{\Gamma_{\varphi_T^x(t_i,Y_i^\tau)}\}_{x\in[0,1],\,i\in\{1,\ldots,N\},\,\tau\in[0,1]}$. We need a comparison theorem for stochastic heat equation with Neumann boundary conditions, whose proof is omitted since it is similar to the case of Dirichlet boundary conditions (see (Mueller & Nualart, 2008, Lemma 4)).

Lemma 4.6 Let $u_i(t,x)$: $i=1,2$ be two solutions of

$$\frac{\partial u_i}{\partial t}=\frac{\partial^2 u_i}{\partial x^2}+B_iu_i+Hu_i\frac{\partial^2 W}{\partial t\partial x},\qquad u_i(0,x)=u_0^{(i)}(x)$$

with Neumann boundary conditions, where $B_i(t,x)$, $i=1,2$ and $H(t,x)$ are bounded and adapted processes, and $u_0^{(i)}(x)$, $i=1,2$, are nonnegative continuous functions not identically zero. Also assume that $B_1(t,x)\leqslant B_2(t,x)$, $u_0^{(1)}(x)\leqslant u_0^{(2)}(x)$ hold with probability one for all $t\geqslant 0$, $x\in[0,1]$. Then with probability 1,

$$u_1(t,x)\leqslant u_2(t,x),\quad \forall\, t\geqslant 0,\ x\in[0,1]. \tag{4.15}$$

If in Lemma 4.6, the assumptions of the initial conditions are replaced by $u_0^{(1)}(x)\leqslant u_0^{(2)}(x)$ and $u_0^{(i)}(x)$, $i=1,2$, are nonpositive continuous functions not identically zero, then by applying Lemma 4.6 to $-u_1$ and $-u_2$ yields that $-u_2(t,x)\leqslant -u_1(t,x)$, which also gives (4.15).

Proposition 4.7 Assume that $b\in\mathbf{C}_b^1$ and $\sigma>0$. Then the Malliavin covariance matrix $\{\Gamma_{\varphi_T^x(t_i,Y_i^\tau)}\}_{x\in[0,1],i\in\{1,\dots,N\},\tau\in[0,1]}$ satisfies $\Gamma_{\varphi_T^x(t_i,Y_i^\tau)}\geqslant c$, for some $c=c(T,|b|_1,\sigma)>0$ independent of δ, x, i and τ.

Proof. By the Cauchy–Schwarz inequality, we infer that

$$\Gamma_{\varphi_T^x(t_i,Y_i^\tau)}:=\int_0^T\int_0^1|D_{\theta,\xi}\varphi_T^x(t_i,Y_i^\tau)|^2\mathrm{d}\xi\,\mathrm{d}\theta\geqslant\int_0^T\left(\int_0^1 D_{\theta,\xi}\varphi_T^x(t_i,Y_i^\tau)\mathrm{d}\xi\right)^2\mathrm{d}\theta, \tag{4.16}$$

for $i\in\{1,\dots,N\}$. Denote $X_\theta^i(t,x):=\int_0^1 D_{\theta,\xi}\varphi_t^x(t_i,Y_i^\tau)\mathrm{d}\xi$, where we drop its explicit dependence upon τ for simplicity. By (2.1) and the chain rule (see e.g. (Nualart, 2006, Proposition 1.5.1)), we have

$$X_\theta^i(t,x)=\sigma\mathbf{1}_{(t_i,t]}(\theta)+\int_0^1 G_{t-t_i}(x,y)\int_0^1 D_{\theta,\xi}Y_i^\tau(y)\mathrm{d}\xi\,\mathrm{d}y+\int_{t_i}^t\int_0^1 G_{t-r}(x,y)b'(\varphi_r^y(t_i,Y_i^\tau))X_\theta^i(r,y)\mathrm{d}y\mathrm{d}r. \tag{4.17}$$

Next we estimate $X_\theta^i(t,x)$ in the following two cases, i.e., $\theta>t_i$ and $\theta<t_i$.

Case 1: Let $\theta\in(t_i,T]$. Notice that $\theta>t_i$ implies $D_{\theta,\xi}Y_i^\tau(y)=0$ since Y_i^τ defined in (3.9) is $\mathscr{F}_{t_i}$-measurable. Then it follows from (4.17) that

$$\partial_t X_\theta^i(t,x)=\partial_{xx}X_\theta^i(t,x)+b'(\varphi_t^x(t_i,Y_i^\tau))X_\theta^i(t,x),\quad \theta<t\leqslant T$$

with the initial condition $X_\theta^i(\theta,x)=\sigma$, $\forall\,x\in[0,1]$. By Lemma 4.6 (with $B_1\equiv-|b|_1$, $B_2(t,x)=b'(\varphi_t^x(t_i,Y_i^\tau))$, $H\equiv 0$), we obtain

$$X_\theta^i(T,x)\geqslant e^{-|b|_1(T-\theta)}\sigma. \tag{4.18}$$

Case 2: Let $\theta\in(0,t_i)$. Notice that $\theta<t_i$ implies $\mathbf{1}_{(t_i,t]}(\theta)=0$. By (3.9) and (4.17), it is sufficient to estimate $\int_0^1 D_{\theta,\xi}Y_i^\tau(y)\mathrm{d}\xi$ in terms of the estimates of $\int_0^1 D_{\theta,\xi}\Phi_{t_i}^y(0,u_0)\mathrm{d}\xi$ and $\int_0^1 D_{\theta,\xi}\varphi_{t_i}^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\mathrm{d}\xi$. By (2.1), for $\theta<t_i$,

$$\begin{aligned}\int_0^1 D_{\theta,\xi}\varphi_{t_i}^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\mathrm{d}\xi&=\sigma\mathbf{1}_{(t_{i-1},t_i)}(\theta)+\int_0^1 G_\delta(y,z)\int_0^1 D_{\theta,\xi}\Phi_{t_{i-1}}^z(0,u_0)\mathrm{d}\xi\,\mathrm{d}z\\&+\int_{t_{i-1}}^{t_i}\int_0^1 G_{t_i-r}(y,z)b'(\varphi_r^z(t_{i-1},\Phi_{t_{i-1}}(0,u_0)))\int_0^1 D_{\theta,\xi}\varphi_r^z(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\mathrm{d}\xi\,\mathrm{d}z\mathrm{d}r.\end{aligned}$$

If $\theta\in(t_{i-1},t_i)$, then $D_{\theta,\xi}\Phi_{t_{i-1}}^z(0,u_0)=0$ and thus similar to (4.18), applying Lemma 4.6 (with $B_1\equiv-|b|_1$, $B_2(t,x)=b'(\varphi_t^x(t_{i-1},\Phi_{t_{i-1}}(0,u_0)))$, $H\equiv 0$) will lead to

$$\int_0^1 D_{\theta,\xi}\varphi_{t_i}^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\mathrm{d}\xi\geqslant e^{-|b|_1(t_i-\theta)}\sigma\geqslant e^{-|b|_1\delta}\sigma,$$

which along with (3.9), $D_{\theta,\xi}\Phi_{t_i}^y(0,u_0)=\sigma G_{t_i-\theta}(y,\xi)$, $\theta\in(t_{i-1},t_i)$, and (2.4), implies for all $\tau,y\in[0,1]$,

$$\int_0^1 D_{\theta,\xi}Y_i^\tau(y)\mathrm{d}\xi\geqslant(1-\tau)\sigma+\tau e^{-|b|_1\delta}\sigma\geqslant e^{-|b|_1\delta}\sigma,\quad\forall\,\theta\in(t_{i-1},t_i). \tag{4.19}$$

Denote $K_b := \frac{\log\frac{3}{2}}{|b|_1}$. Due to (4.10), we have

$$\int_0^1 D_{\theta,\xi}\Phi^z_{t_{i-1}}(0,u_0)\mathrm{d}\xi \geqslant \frac{1}{2}\sigma, \quad \forall\, z\in[0,1],\ \theta\in(t_k,t_{k+1}),\ 0\vee(i-2-\frac{K_b}{\delta})\leqslant k\leqslant i-2,$$

which, together with (2.4) and Lemma 4.6 (with $B_2(t,x)=b'(\varphi^x_t(t_{i-1},\Phi_{t_{i-1}}(0,u_0)))$, $B_1\equiv -|b|_1$, $H\equiv 0$) indicates

$$\int_0^1 D_{\theta,\xi}\varphi^y_{t_i}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\mathrm{d}\xi \geqslant \frac{1}{2}e^{-|b|_1\delta}\sigma, \quad \forall\, y\in[0,1],\ \theta\in(t_k,t_{k+1}),\ 0\vee(i-2-\frac{K_b}{\delta})\leqslant k\leqslant i-2.$$

Therefore, (4.10) and (3.9) indicate for all $\tau,y\in[0,1]$,

$$\int_0^1 D_{\theta,\xi}Y^\tau_i(y)\mathrm{d}\xi \geqslant (1-\tau)\frac{1}{2}\sigma+\frac{\tau}{2}e^{-|b|_1\delta}\sigma \geqslant \frac{1}{2}e^{-|b|_1\delta}\sigma, \quad \theta\in(t_k,t_{k+1}),\ 0\vee(i-1-\frac{K_b}{\delta})\leqslant k\leqslant i-2. \tag{4.20}$$

Combining (4.19) and (4.20), we conclude

$$\int_0^1 D_{\theta,\xi}Y^\tau_i(y)\mathrm{d}\xi \geqslant \frac{1}{2}e^{-|b|_1\delta}\sigma, \quad \theta\in(t_k,t_{k+1}),\ 0\vee(i-1-\frac{K_b}{\delta})\leqslant k\leqslant i-1. \tag{4.21}$$

Now we turn to (4.17) and estimate $X^i_\theta(T,x)$. Taking into account (4.21) and applying Lemma 4.6 (with $B_2(t,x)=b'(\varphi^x_t(t_i,Y^\tau_i))$, $B_1\equiv -|b|_1$ and $H\equiv 0$) yield

$$X^i_\theta(T,x) \geqslant \int_0^1 G_{T-t_i}(x,y)e^{-|b|_1(T-t_i)}\int_0^1 D_{\theta,\xi}Y^\tau_i(y)\mathrm{d}\xi\mathrm{d}y \geqslant \frac{1}{2}e^{-|b|_1(T-t_{i-1})}\sigma, \tag{4.22}$$

for any $\tau,x\in[0,1]$ and $\theta\in(t_k,t_{k+1})$ with $0\vee(i-1-K_b/\delta)\leqslant k\leqslant i-1$.

So far, we have dominated $X^i_\theta(T,x)$ from below when $\theta>t_i$ in **Case 1** and when $\theta\in(t_k,t_{k+1})$ with $0\vee(i-1-K_b/\delta)\leqslant k\leqslant i-1$ in **Case 2**, based on which, we now give a lower bound of $\Gamma_{\varphi^x_T(t_i,Y^\tau_i)}$ as follows. By (4.16) and (4.18),

$$\Gamma_{\varphi^x_T(t_i,Y^\tau_i)} \geqslant \int_0^T \left|X^i_\theta(T,x)\right|^2\mathrm{d}\theta \geqslant \sum_{k=0}^{i-1}\int_{t_k}^{t_{k+1}}\left|X^i_\theta(T,x)\right|^2\mathrm{d}\theta+\int_{t_i}^T e^{-2|b|_1(T-\theta)}\sigma^2\mathrm{d}\theta.$$

For $0\leqslant i\leqslant \frac{N}{2}$,

$$\Gamma_{\varphi^x_T(t_i,Y^\tau_i)} \geqslant \int_{\frac{T}{2}}^T e^{-2|b|_1(T-\theta)}\sigma^2\mathrm{d}\theta = \frac{1-e^{-|b|_1T}}{2|b|_1}\sigma^2 =: c_1.$$

For $\frac{N}{2}<i\leqslant N$, utilizing (4.22) yields

$$\begin{aligned}\Gamma_{\varphi^x_T(t_i,Y^\tau_i)} &\geqslant \sum_{0\vee(i-1-K_b/\delta)\leqslant k\leqslant i-1}\int_{t_k}^{t_{k+1}}\left|X^i_\theta(T,x)\right|^2\mathrm{d}\theta \geqslant \sum_{k=0\vee[i-\frac{K_b}{\delta}]}^{i-1}\int_{t_k}^{t_{k+1}}\left|X^i_\theta(T,x)\right|^2\mathrm{d}\theta\\ &\geqslant \frac{\sigma^2}{4}e^{-|b|_1T}\delta\min\left\{i,\frac{K_b}{\delta}\right\} \geqslant \frac{\sigma^2}{8}e^{-|b|_1T}(T\wedge K_b) =: c_2,\end{aligned}$$

where we have used $i-1-\frac{K_b}{\delta}<[i-1-\frac{K_b}{\delta}]+1=[i-\frac{K_b}{\delta}]\leqslant i-\frac{K_b}{\delta}$. Finally, we finish the proof by choosing $c=c_1\wedge c_2$. □

Based on Propositions 4.3 and 4.7, we now give the proof of Lemma 3.5.

Proof of Lemma 3.5: Since $f\in\Psi$, there exists F such that $F'=f$ and $0\leqslant F\leqslant 1$. Invoking Propositions 4.3 and 4.7, it follows from (Nualart, 2006, formula (2.32) or Proposition 2.1.4) that for any $\alpha\in\mathbb{N}$, $k\geqslant 1$, there exists $H_{\alpha+1}(\varphi^x_T(t_i,Y^\tau_i),G_1)\in\mathbb{D}^\infty$ satisfying (3.14), and furthermore, for $p>p_1\geqslant 1$, there exist constants η, γ and integers n,m such that

$$\|H_{\alpha+1}(\varphi^x_T(t_i,Y^\tau_i),G_1)\|_{p_1}\leqslant C(p_1,p)\|\Gamma^{-1}_{\varphi^x_T(t_i,Y^\tau_i)}\|^m_\eta\|\varphi^x_T(t_i,Y^\tau_i)\|^n_{\alpha+1,\gamma}\|G_1\|_{\alpha+1,p}.$$

Hence, by taking $p=2$, $p_1=1$ and using Propositions 4.3 and 4.7, we complete the proof. □

5. Regularity estimates for derivatives

In this section, we present the estimates of the moments of the Gateaux derivative and the Malliavin derivative of $\varphi_t^x(t_i,Y_i^\tau)$ in Lemmas 5.3 and 5.4, respectively, which will be used in the proof of Proposition 3.3. As we will see, the p-th moments of the derivatives of $\varphi_t^x(t_i,Y_i^\tau)$ are dominated by the Green function, instead of being bounded by a constant in the case of stochastic ordinary differential equations. This is one main difference in the weak convergence analysis between SPDEs and stochastic ordinary differential equations.

Lemma 5.1 Let $M \geqslant 1$, $X,H \in \mathbb{D}^{M,\infty}$ and $g \in C_b^\infty$. Then for any $q \geqslant 1$,

$$\|D^M\{g'(X)H\}\|_{L^q(\Omega;\mathbb{H}^{\otimes M})} \leqslant |g|_1 \|D^M H\|_{L^q(\Omega;\mathbb{H}^{\otimes M})} + C\|g'(X)\|_{M,2q}\|H\|_{M-1,2q}.$$

Proof. By Leibnitz's rule (Nualart, 2006, Proposition 1.5.6 or Exercise 1.2.13) and Hölder's inequality,

$$\|D^M\{g'(X)H\}\|_{L^q(\Omega,\mathbb{H}^{\otimes M})} \leqslant |g|_1 \|D^M H\|_{L^q(\Omega,\mathbb{H}^{\otimes M})} + \sum_{j=0}^{M-1} \binom{M}{j} \|D^{M-j}g'(X)\|_{L^{2q}(\Omega,\mathbb{H}^{\otimes(M-j)})} \|D^j H\|_{L^{2q}(\Omega,\mathbb{H}^{\otimes j})}.$$

Noticing that for $j \in \{0,1,\dots,M-1\}$,

$$\|D^{M-j}g'(X)\|_{L^{2q}(\Omega,\mathbb{H}^{\otimes(M-j)})} \|D^j H\|_{L^{2q}(\Omega,\mathbb{H}^{\otimes j})} \leqslant \|g'(X)\|_{M-j,2q}\|H\|_{j,2q} \leqslant \|g'(X)\|_{M,2q}\|H\|_{M-1,2q}.$$

Putting these two estimates together finishes the proof. □

The following two-parameter Gronwall lemma is essential in the moment estimates for derivatives of $\varphi_t^x(t_i,Y_i^\tau)$.

Lemma 5.2 Let $g_{s,y}(t,x) \geqslant 0$ satisfy that for all $0 < s < t \leqslant T$ and $x,y \in [0,1]$,

$$g_{s,y}(t,x) \leqslant cG_{t-s}(x,y) + c\int_s^t \int_0^1 G_{t-r_1}(x,z_1) g_{s,y}(r_1,z_1)\mathrm{d}z_1\mathrm{d}r_1,$$

where $c > 0$ is a constant. Then for some $C = C(T,c)$, it holds that $g_{s,y}(t,x) \leqslant CG_{t-s}(x,y)$.

Proof. Taking the supremum over $x \in [0,1]$, and using (2.4) and $G_{t-s}(x,y) \leqslant KP_{t-s}(x,y) \leqslant \frac{C}{\sqrt{t-s}}$, we have

$$\sup_{x\in[0,1]} g_{s,y}(t,x) \leqslant \frac{C}{\sqrt{t-s}} + c\int_s^t \sup_{z_1\in[0,1]} g_{s,y}(r_1,z_1)\mathrm{d}r_1, \quad \forall\, y \in [0,1],$$

which, together with the Gronwall inequality, implies that for some $C = C(T)$,

$$\begin{aligned}\sup_{x\in[0,1]} g_{s,y}(t,x) &\leqslant \frac{C}{\sqrt{t-s}} + \int_s^t \frac{C}{\sqrt{r-s}} c \exp\left(\int_r^t c\mathrm{d}u\right)\mathrm{d}r \\ &\leqslant \frac{C}{\sqrt{t-s}} + 2ce^{cT}C\sqrt{T}, \quad \forall\, y \in [0,1]. \end{aligned} \tag{5.1}$$

By an iteration procedure and (2.5), we have

$$\begin{aligned} g_{s,y}(t,x) &\leqslant cG_{t-s}(x,y) + c^2\int_s^t\int_0^1 G_{t-r_1}(x,z_1)G_{r_1-s}(z_1,y)\mathrm{d}z_1\mathrm{d}r_1 + \cdots \\ &+ c^{n+1}\int_s^t\int_0^1\int_s^{r_1}\int_0^1\cdots\int_s^{r_{n-1}}\int_0^1 G_{t-r_1}(x,z_1)G_{r_1-r_2}(z_1,z_2)\cdots G_{r_{n-1}-r_n}(z_{n-1},z_n)G_{r_n-s}(z_n,y)\mathrm{d}z_n\mathrm{d}r_n\cdots\mathrm{d}z_2\mathrm{d}r_2\mathrm{d}z_1\mathrm{d}r_1 \\ &+ c^{n+1}\int_s^t\int_0^1\int_s^{r_1}\int_0^1\cdots\int_s^{r_n}\int_0^1 G_{t-r_1}(x,z_1)G_{r_1-r_2}(z_1,z_2)\cdots G_{r_n-r_{n+1}}(z_n,z_{n+1})g_{s,y}(r_{n+1},z_{n+1})\mathrm{d}z_{n+1}\mathrm{d}r_{n+1}\cdots\mathrm{d}z_2\mathrm{d}r_2\mathrm{d}z_1\mathrm{d}r_1 \\ &\leqslant \left(c + c^2(t-s) + \cdots + c^{n+1}\frac{(t-s)^n}{n!}\right)G_{t-s}(x,y) + c^{n+1}\frac{(t-s)^n}{n!}\int_s^t\int_0^1 G_{t-r_{n+1}}(x,z_{n+1})g_{s,y}(r_{n+1},z_{n+1})\mathrm{d}z_{n+1}\mathrm{d}r_{n+1}, \end{aligned}$$

where the first term on the right-hand side is bounded by $ce^{cT}G_{t-s}(x,y)$, and due to (5.1) and (2.4), the second term is dominated by

$$c^{n+1}\frac{(t-s)^n}{n!}C(T)\left(\int_s^t \frac{1}{\sqrt{r_{n+1}-s}}\mathrm{d}r_{n+1} + 1\right),$$

which tends to 0 as $n \to \infty$. The proof is completed. □

Lemma 5.3 Assume that $b \in \mathbf{C}_b^\infty$. Then for any $k \in \mathbb{N}$ and $p \geqslant 1$, there exists $C = C(k,p,T,\sigma)$ such that

$$\|\langle \mathscr{D}\varphi_t^x(t_i,Y_i^\tau), G_{t_i-r}(\cdot,y)\rangle\|_{k,p} \leqslant CG_{t-r}(x,y) \tag{5.2}$$

holds for every $r \in [t_{i-1},t_i)$, $t_i \leqslant t \leqslant T$, $i \in \{1,\ldots,N\}$ and $\tau,x,y \in [0,1]$.

Proof. In this proof, we denote by $\mathscr{K}_M$ the property that (5.2) holds for $k = M$ and all $p \geqslant 1$. The proof is completed by an induction argument on M. First, by (2.5) and (3.12), we obtain

$$\begin{aligned}&\langle \mathscr{D}\varphi_t^x(t_i,Y_i^\tau), G_{t_i-r}(\cdot,y)\rangle = G_{t-r}(x,y)\\&\quad+\int_{t_i}^t\int_0^1 G_{t-\theta_1}(x,z)b'(\varphi_{\theta_1}^z(t_i,Y_i^\tau))\langle \mathscr{D}\varphi_{\theta_1}^z(t_i,Y_i^\tau), G_{t_i-r}(\cdot,y)\rangle \mathrm{d}z\mathrm{d}\theta_1.\end{aligned} \tag{5.3}$$

Utilizing (5.3), the Minkowski inequality and the boundedness of b' gives that for $p \geqslant 1$,

$$\begin{aligned}\|\langle \mathscr{D}\varphi_t^x(t_i,Y_i^\tau), G_{t_i-r}(\cdot,y)\rangle\|_p &\leqslant G_{t-r}(x,y) + |b|_1\int_{t_i}^t\int_0^1 G_{t-\theta_1}(x,z)\|\langle \mathscr{D}\varphi_{\theta_1}^z(t_i,Y_i^\tau), G_{t_i-r}(\cdot,y)\rangle\|_p \mathrm{d}z\mathrm{d}\theta_1\\&\leqslant G_{t-r}(x,y) + |b|_1\int_{r}^t\int_0^1 G_{t-\theta_1}(x,z)\|\langle \mathscr{D}\varphi_{\theta_1}^z(t_i,Y_i^\tau), G_{t_i-r}(\cdot,y)\rangle\|_p \mathrm{d}z\mathrm{d}\theta_1.\end{aligned}$$

A direct application of Lemma 5.2 completes the proof of $\mathscr{K}_0$.

Assume $\mathscr{K}_{M-1}$, and we proceed to prove $\mathscr{K}_M$. By Lemma 5.1 with $q = p$, $g = b$, $X = \varphi_{\theta_1}^z(t_i,Y_i^\tau)$ and $H = \langle \mathscr{D}\varphi_{\theta_1}^z(t_i,Y_i^\tau), G_{t_i-r}(\cdot,y)\rangle$, we get that for $t_i < \theta_1 < t$,

$$\begin{aligned}&\|D^M\{b'(\varphi_{\theta_1}^z(t_i,Y_i^\tau))\langle \mathscr{D}\varphi_{\theta_1}^z(t_i,Y_i^\tau), G_{t_i-r}(\cdot,y)\rangle\}\|_{L^p(\Omega;\mathbb{H}^{\otimes M})}\\&\leqslant |b|_1\|D^M\langle \mathscr{D}\varphi_{\theta_1}^z(t_i,Y_i^\tau), G_{t_i-r}(\cdot,y)\rangle\|_{L^p(\Omega;\mathbb{H}^{\otimes M})} + C\|b'(\varphi_{\theta_1}^z(t_i,Y_i^\tau))\|_{M,2p}\|\langle \mathscr{D}\varphi_{\theta_1}^z(t_i,Y_i^\tau), G_{t_i-r}(\cdot,y)\rangle\|_{M-1,2p}\\&\leqslant |b|_1\|D^M\langle \mathscr{D}\varphi_{\theta_1}^z(t_i,Y_i^\tau), G_{t_i-r}(\cdot,y)\rangle\|_{L^p(\Omega;\mathbb{H}^{\otimes M})} + CG_{\theta_1-r}(z,y),\end{aligned} \tag{5.4}$$

thanks to Corollary 4.4 and $\mathscr{K}_{M-1}$. By taking M-th ($M \geqslant 1$) Malliavin derivatives on both sides of (5.3), and using (2.5) and (5.4), we have

$$\begin{aligned}&\|D^M\langle \mathscr{D}\varphi_t^x(t_i,Y_i^\tau), G_{t_i-r}(\cdot,y)\rangle\|_{L^p(\Omega,\mathbb{H}^{\otimes M})} \leqslant C\int_{t_i}^t\int_0^1 G_{t-\theta_1}(x,z)G_{\theta_1-r}(z,y)\mathrm{d}z\mathrm{d}\theta_1\\&+|b|_1\int_{t_i}^t\int_0^1 G_{t-\theta_1}(x,z)\|D^M\langle \mathscr{D}\varphi_{\theta_1}^z(t_i,Y_i^\tau), G_{t_i-r}(\cdot,y)\rangle\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z\mathrm{d}\theta_1\\&\leqslant C(t-t_i)G_{t-r}(x,y) + |b|_1\int_r^t\int_0^1 G_{t-\theta_1}(x,z)\|D^M\langle \mathscr{D}\varphi_{\theta_1}^z(t_i,Y_i^\tau), G_{t_i-r}(\cdot,y)\rangle\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z\mathrm{d}\theta_1.\end{aligned}$$

Consequently, it follows from Lemma 5.2 that $\|D^M\langle \mathscr{D}\varphi_t^x(t_i,Y_i^\tau), G_{t_i-r}(\cdot,y)\rangle\|_{L^p(\Omega,\mathbb{H}^{\otimes M})} \leqslant CG_{t-r}(x,y)$, which, together with $\mathscr{K}_{M-1}$ and (4.5), yields $\mathscr{K}_M$. The proof is completed. □

Now we are in a position to estimate moments of the Malliavin derivative $D\Phi_{t_i}^y(0,u_0) = \{D_{\theta,\xi}\Phi_{t_i}^y(0,u_0), \theta \in [0,t_i)\times[0,1]\}$ of the numerical solution $U^{\delta,i}(x) = \Phi_{t_i}^x(0,u_0)$. Compared with the result in Proposition 4.3 where the Malliavin–Sobolev norms of $\Phi_{t_i}^y(0,u_0)$ are uniformly bounded by constants, Lemma 5.4 states that the Malliavin–Sobolev norms of $D_{\theta,\xi}\Phi_{t_i}^y(0,u_0)$ are bounded by multiples of the Green function $G_{t_i-\theta}(y,\xi)$.

Lemma 5.4 Assume that $b \in \mathbf{C}_b^\infty$. Then for any $k \in \mathbb{N}$, $p \geqslant 1$, there exists $C = C(k,p,T,\sigma)$ such that for every $\theta \in (0,t_i), y, \xi \in [0,1]$, and $i \in \{1,\ldots,N\}$,

$$\|D_{\theta,\xi}\Phi_{t_i}^y(0,u_0)\|_{k,p} \leqslant CG_{t_i-\theta}(y,\xi).$$

Proof. In this proof, we denote by $\mathscr{B}_{M,j}$ the property that $\|D_{\theta,\xi}\Phi_{t_j}^y(0,u_0)\|_{k,p} \leqslant CG_{t_j-\theta}(y,\xi)$ holds for $k = M$ and all $p \geqslant 1$, $\xi, y \in [0,1]$, $\theta \in (0,t_{j-1})$, and by $\mathscr{B}_M$ the property that $\mathscr{B}_{M,j}$ holds for all $j \in \{2,\ldots,N\}$. Since for $\theta \in (t_{i-1},t_i)$, $i \in \{1,\ldots,N\}$, $D_{\theta,\xi}\Phi_{t_i}^y(0,u_0) = \sigma G_{t_i-\theta}(y,\xi)$, it suffices to prove $\mathscr{B}_M$ for $M \in \mathbb{N}$.

Step 1: We show $\mathscr{B}_0$, i.e., $\mathscr{B}_{0,i}$ holds for $i \in \{2,\ldots,N\}$.

In fact, if $i=2$, then for any $\theta \in (0,t_1)$, $\xi \in [0,1]$,

$$D_{\theta,\xi}\Phi^y_{t_2}(0,u_0) = \sigma G_{t_2-\theta}(y,\xi) + \int_{t_1}^{t_2}\int_0^1 G_{t_2-r}(y,z)b'(\Phi^z_{t_1}(0,u_0))D_{\theta,\xi}\Phi^z_{t_1}(0,u_0)\mathrm{d}z\mathrm{d}r, \tag{5.5}$$

which along with the fact that for any $\theta \in (0,t_1)$, $\xi \in [0,1]$, $D_{\theta,\xi}\Phi^y_{t_1}(0,u_0) = \sigma G_{t_1-\theta}(y,\xi)$ implies

$$\|D_{\theta,\xi}\Phi^y_{t_2}(0,u_0)\|_p \leqslant CG_{t_2-\theta}(y,\xi) + |b|_1\sigma\int_{t_1}^{t_2}\int_0^1 G_{t_2-r}(y,z)G_{t_1-\theta}(z,\xi)\mathrm{d}z\mathrm{d}r.$$

By (2.5) and (2.7),

$$\begin{aligned}\int_{t_1}^{t_2}\int_0^1 G_{t_2-r}(y,z)G_{t_1-\theta}(z,\xi)\mathrm{d}z\mathrm{d}r &= \int_{t_1}^{t_2} G_{t_2-r+t_1-\theta}(y,\xi)\mathrm{d}r \leqslant C(T)\int_{t_1}^{t_2}\sqrt{\frac{t_2-\theta}{t_2-r+t_1-\theta}}G_{t_2-\theta}(y,\xi)\mathrm{d}r \\ &\leqslant C(T)\left(\sqrt{t_2-\theta}-\sqrt{t_1-\theta}\right)\sqrt{t_2-\theta}G_{t_2-\theta}(y,\xi) \\ &\leqslant C(T)G_{t_2-\theta}(y,\xi),\end{aligned} \tag{5.6}$$

which implies $\mathscr{B}_{0,2}$.

To show $\mathscr{B}_{0,i}$ for general $i \in \{3,\ldots,N\}$, we assume by induction that $\mathscr{B}_{0,j}$ holds for all $2 \leqslant j \leqslant i-1$, and aim to prove $\mathscr{B}_{0,i}$. For $\theta \in (0,t_{i-1})$, we have $\theta \in (t_{k-1},t_k]$ for some $k \in \{1,\ldots,i-1\}$. Then by (4.11), the induction assumption $\mathscr{B}_{0,j}$ with $2 \leqslant j \leqslant i-1$, as well as the semigroup property (2.5),

$$\begin{aligned}\|D_{\theta,\xi}\Phi^y_{t_i}(0,u_0)\|_p \leqslant& |b|_1\sum_{j=k}^{i-1}\int_{t_j}^{t_{j+1}}\int_0^1 G_{t_i-r}(y,z)\|D_{\theta,\xi}\Phi^z_{t_j}(0,u_0)\|_p\mathrm{d}z\mathrm{d}r + G_{t_i-\theta}(y,\xi)\sigma \\ \leqslant& C\sum_{j=k}^{i-1}\int_{t_j}^{t_{j+1}}\int_0^1 G_{t_i-r}(y,z)G_{t_j-\theta}(z,\xi)\mathrm{d}z\mathrm{d}r + G_{t_i-\theta}(y,\xi)\sigma \\ \leqslant& C\sum_{j=k}^{i-1}\int_{t_j}^{t_{j+1}} G_{t_i-r+t_j-\theta}(y,\xi)\mathrm{d}r + G_{t_i-\theta}(y,\xi)\sigma.\end{aligned}$$

Since $\sum_{j=k}^{i-1}\int_{t_j}^{t_{j+1}}\sqrt{\frac{t_i-\theta}{t_i-r+t_j-\theta}}\mathrm{d}r = 2\delta\sum_{j=k}^{i-1}\frac{\sqrt{t_i-\theta}}{\sqrt{t_i-\theta}+\sqrt{t_{i-1}-\theta}} \leqslant 2T$, it follows from (2.7) that

$$\sum_{j=k}^{i-1}\int_{t_j}^{t_{j+1}} G_{t_i-r+t_j-\theta}(y,\xi)\mathrm{d}r \leqslant C\sum_{j=k}^{i-1}\int_{t_j}^{t_{j+1}}\sqrt{\frac{t_i-\theta}{t_i-r+t_j-\theta}}\mathrm{d}rG_{t_i-\theta}(y,\xi) \leqslant CG_{t_i-\theta}(y,\xi), \tag{5.7}$$

which completes the proof of $\mathscr{B}_{0,i}$.

Step 2: We assume by induction $\mathscr{B}_{M-1}$ ($M \geqslant 1$), and proceed to show $\mathscr{B}_M$.

It suffices to show $\mathscr{B}_{M,i}$ for all $i \in \{2,\ldots,N\}$. Actually, for $i=2$ and $\theta \in (0,t_1)$, by taking the M-th Malliavin derivative on both sides of (5.5),

$$\|D^M D_{\theta,\xi}\Phi^y_{t_2}(0,u_0)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})} \leqslant \int_{t_1}^{t_2}\int_0^1 G_{t_2-r}(y,z)\|D^M\{b'(\Phi^z_{t_1}(0,u_0))D_{\theta,\xi}\Phi^z_{t_1}(0,u_0)\}\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z\mathrm{d}r.$$

By Lemma 5.1 with $q=p$, $g=b$, $X=\Phi^z_{t_l}(0,u_0)$, and $H=D_{\theta,\xi}\Phi^z_{t_l}(0,u_0)$, we also have for $l \in \{1,\ldots,N\}$,

$$\begin{aligned}&\|D^M\{b'(\Phi^z_{t_l}(0,u_0))D_{\theta,\xi}\Phi^z_{t_l}(0,u_0)\}\|_{L^p(\Omega,\mathbb{H}^{\otimes M})} \\ \leqslant& |b|_1\|D^M D_{\theta,\xi}\Phi^z_{t_l}(0,u_0)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})} + C\|b'(\Phi^z_{t_l}(0,u_0))\|_{M,2p}\|D_{\theta,\xi}\Phi^z_{t_l}(0,u_0)\|_{M-1,2p}.\end{aligned}$$

By (4.1) and Lemma 4.1, it holds that $\|b'(\Phi^z_{t_l}(0,u_0))\|_{M,2p} \leqslant C$, which in combination with $\mathscr{B}_{M-1}$ gives

$$\|D^M\{b'(\Phi^z_{t_l}(0,u_0))D_{\theta,\xi}\Phi^z_{t_l}(0,u_0)\}\|_{L^p(\Omega,\mathbb{H}^{\otimes M})} \leqslant |b|_1\|D^M D_{\theta,\xi}\Phi^z_{t_l}(0,u_0)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})} + CG_{t_l-\theta}(z,\xi). \tag{5.8}$$

Using (5.8) with $l=1$, as well as the fact that $DD_{\theta,\xi}\Phi^z_{t_1}(0,u_0)$ vanishes, leads to

$$\|D^M\{b'(\Phi^z_{t_1}(0,u_0))D_{\theta,\xi}\Phi^z_{t_1}(0,u_0)\}\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\leqslant CG_{t_1-\theta}(z,\xi).$$

Combining the above arguments together, we conclude

$$\|D^M D_{\theta,\xi}\Phi^y_{t_2}(0,u_0)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\leqslant C\int_{t_1}^{t_2}\int_0^1 G_{t_2-r}(y,z)G_{t_1-\theta}(z,\xi)\mathrm{d}z\mathrm{d}r\leqslant C(T)G_{t_2-\theta}(y,\xi),$$

where the last inequality follows from (5.6). This along with $\mathscr{B}_{M-1}$ proves $\mathscr{B}_{M,2}$.

To show $\mathscr{B}_{M,i}$ for general $i\in\{3,\dots,N\}$, we assume by induction that $\mathscr{B}_{M,j}$ holds for $j\in\{2,\dots,i-1\}$. For $\theta\in(0,t_{i-1})$, we have $\theta\in(t_{k-1},t_k]$ for some $k\in\{1,\dots,i-1\}$. Taking the M-th Malliavin derivative on both sides of (4.11) and using (5.8), we obtain

$$\|D^M D_{\theta,\xi}\Phi^y_{t_i}(0,u_0)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\leqslant\sum_{j=k}^{i-1}\int_{t_j}^{t_{j+1}}\int_0^1 G_{t_i-r}(y,z)\|D^M\{b'(\Phi^z_{t_j}(0,u_0))D_{\theta,\xi}\Phi^z_{t_j}(0,u_0)\}\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z\mathrm{d}r$$

$$\leqslant\sum_{j=k}^{i-1}\int_{t_j}^{t_{j+1}}\int_0^1 G_{t_i-r}(y,z)|b|_1\|D^M D_{\theta,\xi}\Phi^z_{t_j}(0,u_0)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z\mathrm{d}r+C\sum_{j=k}^{i-1}\int_{t_j}^{t_{j+1}}\int_0^1 G_{t_i-r}(y,z)G_{t_j-\theta}(z,\xi)\mathrm{d}z\mathrm{d}r.$$

Using the assumption that $\mathscr{B}_{M,j}$ holds for $j\in\{2,\dots,i-1\}$ and (5.7), we conclude

$$\|D^M D_{\theta,\xi}\Phi^y_{t_i}(0,u_0)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\leqslant C\sum_{j=k}^{i-1}\int_{t_j}^{t_{j+1}}\int_0^1 G_{t_i-r}(y,z)G_{t_j-\theta}(z,\xi)\mathrm{d}z\mathrm{d}r\leqslant CG_{t_i-\theta}(y,\xi),$$

which together with $\mathscr{B}_{M-1}$ indicates $\mathscr{B}_{M,i}$.

Combining *Step 1* and *Step 2*, we complete the proof.

□

The following Lemma considers the bounds for the Malliavin derivative of the exact flow of Eq. (1.1).

Lemma 5.5 Assume that $b\in\mathbf{C}_b^\infty$. Given $i\in\{0,1,\dots,N-1\}$, let $X_i:\Omega\times[0,1]\to\mathbb{R}$ satisfy that $X_i(z)$ is $\mathscr{F}_{t_i}$ measurable for any $z\in[0,1]$. If for all $k\in\mathbb{N}$ and $p\geqslant 1$,

$$\|\varphi^y_t(t_i,X_i)\|_{k,p}\leqslant C(k,p,T),\quad\forall\, t\in[t_i,T],\ y\in[0,1],\tag{5.9}$$

and

$$\|D_{\theta,\xi}X_i(z)\|_{k,p}\leqslant C(k,p,T)G_{t_i-\theta}(z,\xi),\quad\forall\,\theta\in(0,t_i),\ z,\xi\in[0,1],\tag{5.10}$$

then there is some constant $C=C(k,p,T)$ such that for any $k\in\mathbb{N}$ and $p\geqslant 1$,

$$\|D_{\theta,\xi}\varphi^y_t(t_i,X_i)\|_{k,p}\leqslant C(k,p,T)G_{t-\theta}(y,\xi),\quad\forall\, t\in(t_i,T],\ \theta\in(0,t_i),\ y,\xi\in[0,1].\tag{5.11}$$

Proof. In this proof, we denote by $\mathscr{G}_M$ the property that (5.11) holds for all $p\geqslant 1$, $t\in(t_i,T]$, $\theta\in(0,t_i)$, $y,\xi\in[0,1]$ and $k=M$. We first prove $\mathscr{G}_0$. Notice that for $\theta\in(0,t_i)$,

$$D_{\theta,\xi}\varphi^y_t(t_i,X_i)=\int_0^1 G_{t-t_i}(y,z)D_{\theta,\xi}X_i(z)\mathrm{d}z+\int_{t_i}^t\int_0^1 G_{t-r}(y,z)b'(\varphi^z_r(t_i,X_i))D_{\theta,\xi}\varphi^z_r(t_i,X_i)\mathrm{d}z\mathrm{d}r.\tag{5.12}$$

Using the Minkowski inequality, (5.10), (2.5) and the boundedness of b', we have

$$\begin{aligned}\|D_{\theta,\xi}\varphi^y_t(t_i,X_i)\|_p&\leqslant CG_{t-\theta}(y,\xi)+|b|_1\int_{t_i}^t\int_0^1 G_{t-r}(y,z)\|D_{\theta,\xi}\varphi^z_r(t_i,X_i)\|_p\mathrm{d}z\mathrm{d}r\\&\leqslant CG_{t-\theta}(y,\xi)+|b|_1\int_{\theta}^t\int_0^1 G_{t-r}(y,z)\|D_{\theta,\xi}\varphi^z_r(t_i,X_i)\|_p\mathrm{d}z\mathrm{d}r.\end{aligned}$$

Then $\mathscr{G}_0$ follows from Lemma 5.2.

Now we assume $\mathscr{G}_{M-1}$ and proceed to show $\mathscr{G}_M$. Taking the M-th Malliavin derivative on both sides of (5.12), we have

$$\|D^M D_{\theta,\xi}\varphi_t^y(t_i,X_i)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})} \leqslant \int_0^1 G_{t-t_i}(y,z)\|D^M D_{\theta,\xi}X_i(z)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z + \int_{t_i}^t\int_0^1 G_{t-r}(y,z)\|D^M\{b'(\varphi_r^z(t_i,X_i))D_{\theta,\xi}\varphi_r^z(t_i,X_i)\}\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z\mathrm{d}r. \quad (5.13)$$

By Lemma 4.1 and (5.9), $\|b'(\varphi_r^z(t_i,X_i))\|_{M,2p} \leqslant C(M,p,T)$ for $r \in (t_i,t]$ and $z \in [0,1]$. By Lemma 5.1 with $g = b$, $X = \varphi_r^z(t_i,X_i)$, and $H = D_{\theta,\xi}\varphi_r^z(t_i,X_i)$,

$$\begin{aligned}
&\|D^M\{b'(\varphi_r^z(t_i,X_i))D_{\theta,\xi}\varphi_r^z(t_i,X_i)\}\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\\
&\leqslant |b|_1\|D^M D_{\theta,\xi}\varphi_r^z(t_i,X_i)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})} + C\|b'(\varphi_r^z(t_i,X_i))\|_{M,2p}\|D_{\theta,\xi}\varphi_r^z(t_i,X_i)\|_{M-1,2p}\\
&\leqslant |b|_1\|D^M D_{\theta,\xi}\varphi_r^z(t_i,X_i)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})} + CG_{r-\theta}(z,\xi), \qquad (5.14)
\end{aligned}$$

where we have used $\mathscr{G}_{M-1}$ in the last line. Inserting (5.10) and (5.14) into (5.13), and using (2.5), we arrive at

$$\begin{aligned}
\|D^M D_{\theta,\xi}\varphi_t^y(t_i,X_i)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})} &\leqslant \int_0^1 G_{t-t_i}(y,z)G_{t_i-\theta}(z,\xi)\mathrm{d}z + C\int_{t_i}^t\int_0^1 G_{t-r}(y,z)G_{r-\theta}(z,\xi)\mathrm{d}z\mathrm{d}r\\
&\quad + \int_{t_i}^t\int_0^1 G_{t-r}(y,z)|b|_1\|D^M D_{\theta,\xi}\varphi_r^z(t_i,X_i)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z\mathrm{d}r\\
&\leqslant CG_{t-\theta}(y,\xi) + C\int_\theta^t\int_0^1 G_{t-r}(y,z)\|D^M D_{\theta,\xi}\varphi_r^z(t_i,X_i)\|_{L^p(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z\mathrm{d}r.
\end{aligned}$$

Finally, by Lemma 5.2, we obtain $\mathscr{G}_M$, which completes the proof. □

Lemma 5.6 Assume that $b \in \mathbf{C}_b^\infty$. Then for any $k \in \mathbb{N}$ and $p \geqslant 1$, there exists $C = C(k,p,T,\sigma)$ such that for every $\theta_1 \in (t_{i-1},r)$, $r \in (t_{i-1},t_i]$, $\beta,y,\xi \in [0,1]$, and $i \in \{2,\ldots,N\}$,

$$\|D_{\theta_1,\xi}Z_i^\beta(r,y)\|_{k,p} \leqslant CG_{r-\theta_1}(y,\xi). \quad (5.15)$$

Proof. By the definition of $Z_i^\beta(r,y)$, we have that for $\theta_1 \in (t_{i-1},r) \subset (t_{i-1},t_i]$,

$$D_{\theta_1,\xi}Z_i^\beta(r,y) = \beta D_{\theta_1,\xi}\varphi_r^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0)) + (1-\beta)D_{\theta_1,\xi}\Phi_{t_{i-1}}^y(0,u_0) = \beta D_{\theta_1,\xi}\varphi_r^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0)),$$

since $\theta_1 > t_{i-1}$ implies $D_{\theta_1,\xi}\Phi_{t_{i-1}}^y(0,u_0) = 0$. Therefore, (5.15) is equivalent to

$$\|D_{\theta_1,\xi}\varphi_r^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{k,p} \leqslant CG_{r-\theta_1}(y,\xi). \quad (5.16)$$

Note that for $\theta_1 \in (t_{i-1},r)$

$$\begin{aligned}
&D_{\theta_1,\xi}\varphi_r^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0)) = \sigma G_{r-\theta_1}(y,\xi)\\
&+\int_{\theta_1}^r\int_0^1 G_{r-r_1}(y,z_1)b'(\varphi_{r_1}^{z_1}(t_{i-1},\Phi_{t_{i-1}}(0,u_0)))D_{\theta_1,\xi}\varphi_{r_1}^{z_1}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\mathrm{d}z_1\mathrm{d}r_1. \qquad (5.17)
\end{aligned}$$

Taking $\|\cdot\|_p$-norm on both sides of (5.17), we obtain for $p \geqslant 1$,

$$\|D_{\theta_1,\xi}\varphi_r^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_p \leqslant CG_{r-\theta_1}(y,\xi) + |b|_1\int_{\theta_1}^r\int_0^1 G_{r-r_1}(y,z_1)\|D_{\theta_1,\xi}\varphi_{r_1}^{z_1}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_p\mathrm{d}z_1\mathrm{d}r_1,$$

which implies $\|D_{\theta_1,\xi}\varphi_r^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_p \leqslant CG_{r-\theta_1}(y,\xi)$, thanks to Lemma 5.2.

Assume by induction that (5.16) holds for any $\theta_1 \in (t_{i-1},r)$, $p \geqslant 1$ and $k = M-1$ ($M \geqslant 1$), and we aim to prove (5.16) with $k = M$. Taking the M-th Malliavin derivative on both sides of (5.17), we obtain

$$\|D^M D_{\theta_1,\xi}\varphi_r^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{L^p(\Omega;\mathbb{H}^{\otimes M})}$$

$$\leqslant \int_{\theta_1}^{r}\int_0^1 G_{r-r_1}(y,z_1)\|D^M\{b'(\varphi_{r_1}^{z_1}(t_{i-1},\Phi_{t_{i-1}}(0,u_0)))D_{\theta_1,\xi}\varphi_{r_1}^{z_1}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\}\|_{L^p(\Omega;\mathbb{H}^{\otimes M})}\mathrm{d}z_1\mathrm{d}r_1$$
$$\leqslant |b|_1\int_{\theta_1}^{r}\int_0^1 G_{r-r_1}(y,z_1)\|D^M D_{\theta_1,\xi}\varphi_{r_1}^{z_1}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{L^p(\Omega;\mathbb{H}^{\otimes M})}\mathrm{d}z_1\mathrm{d}r_1$$
$$+C\int_{\theta_1}^{r}\int_0^1 G_{r-r_1}(y,z_1)\|b'(\varphi_{r_1}^{z_1}(t_{i-1},\Phi_{t_{i-1}}(0,u_0)))\|_{M,2p}\|D_{\theta_1,\xi}\varphi_{r_1}^{z_1}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{M-1,2p}\mathrm{d}z_1\mathrm{d}r_1,$$

where in the second inequality, we have used Lemma 5.1 with $g=b$, $X=\varphi_{r_1}^{z_1}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))$ and $H=D_{\theta_1,\xi}\varphi_{r_1}^{z_1}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))$. Using the induction assumption, we have $\|D_{\theta_1,\xi}\varphi_{r_1}^{z_1}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{M-1,2p}\leqslant CG_{r_1-\theta_1}(z,\xi)$. Using Lemma 4.1 and Proposition 4.3, we also have $\|b'(\varphi_{r_1}^{z_1}(t_{i-1},\Phi_{t_{i-1}}(0,u_0)))\|_{M,2p}\leqslant C$. Therefore,

$$\|D^M D_{\theta_1,\xi}\varphi_{r}^{y}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{L^p(\Omega;\mathbb{H}^{\otimes M})}$$
$$\leqslant |b|_1\int_{\theta_1}^{r}\int_0^1 G_{r-r_1}(y,z_1)\|D^M D_{\theta_1,\xi}\varphi_{r_1}^{z_1}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{L^p(\Omega;\mathbb{H}^{\otimes M})}\mathrm{d}z_1\mathrm{d}r_1+C\int_{\theta_1}^{r}\int_0^1 G_{r-r_1}(y,z_1)G_{r_1-\theta_1}(z_1,\xi)\mathrm{d}z_1\mathrm{d}r_1$$
$$\leqslant |b|_1\int_{\theta_1}^{r}\int_0^1 G_{r-r_1}(y,z_1)\|D^M D_{\theta_1,\xi}\varphi_{r_1}^{z_1}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{L^p(\Omega;\mathbb{H}^{\otimes M})}\mathrm{d}z_1\mathrm{d}r_1+CG_{r-\theta_1}(y,\xi),$$

thanks to (2.5). Applying Lemma 5.2 leads to $\|D^M D_{\theta_1,\xi}\varphi_{r}^{y}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{L^p(\Omega;\mathbb{H}^{\otimes M})}\leqslant CG_{r-\theta_1}(y,\xi)$. This together with the induction assumption finishes the proof of (5.16).

□

Corollary 5.7 Assume that $b\in\mathbf{C}_b^\infty$. Then for any $k\in\mathbb{N}$ and $p\geqslant 1$, there exists $C=C(k,p,T,\sigma)$ such that for every $\theta\in(0,t_{i-1})$, $\theta_1\in(0,r)$, $r\in(t_{i-1},t_i]$, $\beta,\tau,y,\xi\in[0,1]$, $s\in(t_{i-1},T]$, $t\in(t_i,T]$ and $i\in\{2,\ldots,N\}$,

$$\|D_{\theta,\xi}\varphi_s^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\|_{k,p}\leqslant CG_{s-\theta}(y,\xi), \tag{5.18}$$
$$\|D_{\theta,\xi}\varphi_t^y(t_i,Y_i^\tau)\|_{k,p}\leqslant CG_{t-\theta}(y,\xi), \tag{5.19}$$
$$\|D_{\theta_1,\xi}Z_i^\beta(r,y)\|_{k,p}\leqslant CG_{r-\theta_1}(y,\xi). \tag{5.20}$$

Proof. (i) Let $X_{i-1}:=\Phi_{t_{i-1}}(0,u_0)$ for $i\in\{2,\ldots,N\}$. Then (5.9) and (5.10) follow from (4.1) and Lemma 5.4, respectively. This allows us to apply Lemma 5.5 to get (5.18).
(ii) Let $X_i:=Y_i^\tau$ for $i\in\{1,\ldots,N-1\}$. Then (4.2) implies (5.9). Recall that by (3.9), $Y_i^\tau(y)=\tau\Phi_{t_i}^y(0,u_0)+(1-\tau)\varphi_{t_i}^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))$. Hence, (5.18) and Lemma 5.4 imply (5.10). Using Lemma 5.5 yields (5.19).
(iii) By the definition of $Z_i^\beta(r,y)$, $D_{\theta_1,\xi}Z_i^\beta(r,y)=\beta D_{\theta_1,\xi}\varphi_r^y(t_{i-1},\Phi_{t_{i-1}}(0,u_0))+(1-\beta)D_{\theta_1,\xi}\Phi_{t_{i-1}}^y(0,u_0)$. If $\theta_1\in(0,t_{i-1})$, then (5.20) follows from (5.18) and Lemma 5.4. If $\theta_1\in(t_{i-1},r)$, then (5.20) follows from Lemma 5.6. The proof is completed. □

Lemma 5.8 Assume that $b\in\mathbf{C}_b^\infty$. Then for any $k\in\mathbb{N}$ and $p\geqslant 1$, there exists $C=C(k,p,T,\sigma)$ such that

$$\|D_{\theta,\xi}\langle\mathscr{D}\varphi_t^x(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\rangle\|_{k,p}\leqslant CG_{t-\theta}(x,\xi) \tag{5.21}$$

holds for every $\theta\in(0,t_{i-1})$, $r\in[t_{i-1},t_i)$, $t_i<t\leqslant T$, $i\in\{1,\ldots,N\}$ and $\tau,x,y,\xi\in[0,1]$.

Proof. Denote by $\mathscr{H}_M$ the property that (5.21) holds for $k=M$ and all $p\geqslant 1$, $\theta\in(0,t_{i-1})$, $r\in[t_{i-1},t_i)$, $t_i<t\leqslant T$, $i\in\{1,\ldots,N\}$ and $\tau,x,y,\xi\in[0,1]$.

We first prove $\mathscr{H}_0$. Taking the Malliavin derivative $D_{\theta,\xi}$ on both sides of (5.3) gives

$$D_{\theta,\xi}\langle\mathscr{D}\varphi_t^x(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\rangle$$
$$=\int_{t_i}^{t}\int_0^1 G_{t-\theta_1}(x,z)b''(\varphi_{\theta_1}^z(t_i,Y_i^\tau))D_{\theta,\xi}\varphi_{\theta_1}^z(t_i,Y_i^\tau)\langle\mathscr{D}\varphi_{\theta_1}^z(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\rangle\mathrm{d}z\mathrm{d}\theta_1$$
$$+\int_{t_i}^{t}\int_0^1 G_{t-\theta_1}(x,z)b'(\varphi_{\theta_1}^z(t_i,Y_i^\tau))D_{\theta,\xi}\langle\mathscr{D}\varphi_{\theta_1}^z(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\rangle\mathrm{d}z\mathrm{d}\theta_1. \tag{5.22}$$

By Proposition 2.2, Corollary 4.4, Lemma 5.3, and (5.19), we have for $q \geqslant 1$ and $k \in \mathbb{N}$,

$$\begin{aligned}&\|b''(\varphi^z_{\theta_1}(t_i,Y^\tau_i))D_{\theta,\xi}\varphi^z_{\theta_1}(t_i,Y^\tau_i)\langle \mathscr{D}\varphi^z_{\theta_1}(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\|_{k,q}\\ \leqslant{}& C\|b''(\varphi^z_{\theta_1}(t_i,Y^\tau_i))\|_{k,3q}\|D_{\theta,\xi}\varphi^z_{\theta_1}(t_i,Y^\tau_i)\|_{k,3q}\|\langle \mathscr{D}\varphi^z_{\theta_1}(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\|_{k,3q}\\ \leqslant{}& C(k,q)G_{\theta_1-\theta}(z,\xi)G_{\theta_1-r}(z,y).\end{aligned} \tag{5.23}$$

Notice that by (2.6), $G_{\theta_1-r}(z,y) \leqslant C(T)\frac{1}{\sqrt{\theta_1-r}}$ for $\theta_1 > r$, which along with (2.5) implies for $t_{i-1} < r < t_i < t$,

$$\int_{t_i}^t\int_0^1 G_{t-\theta_1}(x,z)G_{\theta_1-\theta}(z,\xi)G_{\theta_1-r}(z,y)\mathrm{d}z\mathrm{d}\theta_1 \leqslant C\int_{t_i}^t G_{t-\theta}(x,\xi)\frac{1}{\sqrt{\theta_1-r}}\mathrm{d}\theta_1 \leqslant 2C\sqrt{t-r}G_{t-\theta}(x,\xi). \tag{5.24}$$

Therefore, we have

$$\begin{aligned}\|D_{\theta,\xi}\langle \mathscr{D}\varphi^x_t(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\|_p \leqslant{}& CG_{t-\theta}(x,\xi)+C\int_{t_i}^t\int_0^1 G_{t-\theta_1}(x,z)\|D_{\theta,\xi}\langle \mathscr{D}\varphi^z_{\theta_1}(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\|_p\mathrm{d}z\mathrm{d}\theta_1\\ \leqslant{}& CG_{t-\theta}(x,\xi)+C\int_{\theta}^t\int_0^1 G_{t-\theta_1}(x,z)\|D_{\theta,\xi}\langle \mathscr{D}\varphi^z_{\theta_1}(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\|_p\mathrm{d}z\mathrm{d}\theta_1,\end{aligned}$$

which together with Lemma 5.2 implies $\mathscr{H}_0$.

Now we assume by induction $\mathscr{H}_{M-1}$ and aim to prove $\mathscr{H}_M$. Taking the Malliavin derivative D^M on both sides of (5.22) gives

$$\begin{aligned}&D^M D_{\theta,\xi}\langle \mathscr{D}\varphi^x_t(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\\ ={}&\int_{t_i}^t\int_0^1 G_{t-\theta_1}(x,z)D^M\{b''(\varphi^z_{\theta_1}(t_i,Y^\tau_i))D_{\theta,\xi}\varphi^z_{\theta_1}(t_i,Y^\tau_i)\langle \mathscr{D}\varphi^z_{\theta_1}(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\}\mathrm{d}z\mathrm{d}\theta_1\\ &+\int_{t_i}^t\int_0^1 G_{t-\theta_1}(x,z)D^M\{b'(\varphi^z_{\theta_1}(t_i,Y^\tau_i))D_{\theta,\xi}\langle \mathscr{D}\varphi^z_{\theta_1}(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\}\mathrm{d}z\mathrm{d}\theta_1.\end{aligned}$$

It follows from Lemma 5.1 with $g=b$, $X=\varphi^z_{\theta_1}(t_i,Y^\tau_i)$, and $H=D_{\theta,\xi}\langle \mathscr{D}\varphi^z_{\theta_1}(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle$ and the first inequality of Corollary 4.4, and the induction assumption $\mathscr{H}_{M-1}$ that

$$\begin{aligned}&\|D^M\{b'(\varphi^z_{\theta_1}(t_i,Y^\tau_i))D_{\theta,\xi}\langle \mathscr{D}\varphi^z_{\theta_1}(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\}\|_{L^q(\Omega,\mathbb{H}^{\otimes M})}\\ \leqslant{}&|b|_1\|D^M D_{\theta,\xi}\langle \mathscr{D}\varphi^z_{\theta_1}(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\|_{L^q(\Omega,\mathbb{H}^{\otimes M})}+C\|D_{\theta,\xi}\langle \mathscr{D}\varphi^z_{\theta_1}(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\|_{M-1,2q}\\ \leqslant{}&|b|_1\|D^M D_{\theta,\xi}\langle \mathscr{D}\varphi^z_{\theta_1}(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\|_{L^q(\Omega,\mathbb{H}^{\otimes M})}+CG_{\theta_1-\theta}(z,\xi).\end{aligned}$$

This, in combination with (5.23) and (5.24), indicates

$$\begin{aligned}\|D^M D_{\theta,\xi}\langle \mathscr{D}\varphi^x_t(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\|_{L^q(\Omega,\mathbb{H}^{\otimes M})} \leqslant{}& C\int_{t_i}^t\int_0^1 G_{t-\theta_1}(x,z)G_{\theta_1-\theta}(z,\xi)\mathrm{d}z\mathrm{d}\theta_1+CG_{t-\theta}(x,\xi)\\ &+C\int_{t_i}^t\int_0^1 G_{t-\theta_1}(x,z)\|D^M D_{\theta,\xi}\langle \mathscr{D}\varphi^z_{\theta_1}(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\|_{L^q(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z\mathrm{d}\theta_1.\end{aligned}$$

Further, taking into account (2.5) and $\theta < t_i$, we arrive at

$$\begin{aligned}&\|D^M D_{\theta,\xi}\langle \mathscr{D}\varphi^x_t(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\|_{L^q(\Omega,\mathbb{H}^{\otimes M})}\\ \leqslant{}& CG_{t-\theta}(x,\xi)+C\int_{\theta}^t\int_0^1 G_{t-\theta_1}(x,z)\|D^M D_{\theta,\xi}\langle \mathscr{D}\varphi^z_{\theta_1}(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\|_{L^q(\Omega,\mathbb{H}^{\otimes M})}\mathrm{d}z\mathrm{d}\theta_1.\end{aligned}$$

Hence, it follows from Lemma 5.2 that

$$\|D^M D_{\theta,\xi}\langle \mathscr{D}\varphi^x_t(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle\|_{L^q(\Omega,\mathbb{H}^{\otimes M})} \leqslant CG_{t-\theta}(x,\xi),$$

which together with (4.5) and the induction assumption $\mathscr{H}_{M-1}$ completes the proof of $\mathscr{H}_M$. □

6. Proof of Proposition 3.3

In this section, we give the proof of Proposition 3.3. We begin with (3.13) and proceed to estimate

$$
\begin{aligned}
&\mathbb{E}[f(u(T,x))]-\mathbb{E}[f(U^{\delta,N}(x))]\\
=&\sum_{i=1}^{N}\int_0^1\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1\mathbb{E}\left[f'(\varphi_T^x(t_i,Y_i^\tau))\left\langle\mathscr{D}\varphi_T^x(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\right\rangle b'(Z_i^\beta(r,y))\mathscr{E}^i(r,y)\right]\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\mathrm{d}\tau,
\end{aligned}
$$

where $\mathscr{E}^i:=E^i_{initial\_u}+E^i_{initial\_b}+E^i_{initial\_\sigma}+E^i_b+E^i_\sigma$ is given in (3.8).

For $i\in\{1,\dots,N\}$ and $\star\in\{initial\_u,initial\_b,initial\_\sigma,b,\sigma\}$, denote

$$
\mathscr{R}^i_\star(r,y):=\mathbb{E}\left[f'(\varphi_T^x(t_i,Y_i^\tau))\left\langle\mathscr{D}\varphi_T^x(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\right\rangle b'(Z_i^\beta(r,y))E^i_\star(r,y)\right]. \tag{6.1}
$$

Hence, it follows that

$$
\begin{aligned}
&\mathbb{E}[f(u(T,x))]-\mathbb{E}[f(U^{\delta,N}(x))]\\
=&\sum_{i=1}^{N}\int_0^1\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1[\mathscr{R}^i_{initial\_u}(r,y)+\mathscr{R}^i_{initial\_b}(r,y)+\mathscr{R}^i_{initial\_\sigma}(r,y)+\mathscr{R}^i_b(r,y)+\mathscr{R}^i_\sigma(r,y)]\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\mathrm{d}\tau\\
=:&\,\mathscr{I}_{initial\_u}+\mathscr{I}_{initial\_b}+\mathscr{I}_{initial\_\sigma}+\mathscr{I}_b+\mathscr{I}_\sigma,
\end{aligned}
$$

where

$$
\mathscr{I}_\star=\sum_{i=1}^{N}\int_0^1\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1\mathscr{R}^i_\star(r,y)\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\mathrm{d}\tau, \tag{6.2}
$$

for each $\star\in\{initial\_u,initial\_b,initial\_\sigma,b,\sigma\}$. Here, we drop the explicit dependence of $\mathscr{I}_j$ upon τ,β,x, and note that the constants C throughout this proof are independent of τ, $\beta\in[0,1]$ and $x\in[0,1]$.

For fixed $0\leqslant r<t_i\leqslant T$ and $y\in[0,1]$, we have $G_{t_i-r}(\cdot,y)\in E$. For each $i\in\{1,\dots,N\}$, if $Q_i(r,y)\in\mathbb{D}^\infty$ for every $(r,y)\in[t_{i-1},t_i]\times[0,1]$, then it follows from Lemma 3.5 that for $\alpha\in\{1,2\}$ and $g=b$ (or $g=b'$),

$$
\begin{aligned}
&|\mathbb{E}[f^{(\alpha)}(\varphi_T^x(t_i,Y_i^\tau))\left\langle\mathscr{D}\varphi_T^x(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\right\rangle g'(Z_i^\beta(r,y))Q_i(r,y)]|\\
\leqslant&\, C\|\left\langle\mathscr{D}\varphi_T^x(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\right\rangle g'(Z_i^\beta(r,y))Q_i(r,y)\|_{\alpha+1,2}\\
\leqslant&\, C\|\left\langle\mathscr{D}\varphi_T^x(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\right\rangle\|_{\alpha+1,4}\|g'(Z_i^\beta(r,y))Q_i(r,y)\|_{\alpha+1,4}\\
\leqslant&\, CG_{T-r}(x,y)\|g'(Z_i^\beta(r,y))Q_i(r,y)\|_{\alpha+1,4}\leqslant CG_{T-r}(x,y)\|Q_i(r,y)\|_{\alpha+1,8},
\end{aligned}
\tag{6.3}
$$

thanks to Proposition 2.2, Lemma 5.3 and Corollary 4.4.

(a) Estimate of $\mathscr{I}_{initial\_u}$. By (6.3) (with $Q_i=E^i_{initial\_u}$, $\alpha=1$, and $g=b$) and (6.1), for any $\nu\in(\frac{1}{3},1)$,

$$
\begin{aligned}
|\mathscr{R}^i_{initial\_u}(r,y)|&\leqslant CG_{T-r}(x,y)\int_0^1|G_r(y,\xi)-G_{t_{i-1}}(y,\xi)||u_0(\xi)|\mathrm{d}\xi\\
&\leqslant C\|u_0\|_E G_{T-r}(x,y)(r-t_{i-1})^\nu(t_{i-1})^{-\nu},\quad\forall\, i\in\{2,\dots,N\},
\end{aligned}
$$

due to Lemma 2.1. For $i=1$, by (2.4),

$$
|\mathscr{R}^1_{initial\_u}(r,y)|\leqslant CG_{T-r}(x,y)\left|\int_0^1 G_r(y,\xi)u_0(\xi)\mathrm{d}\xi-u_0(y)\right|\mathrm{d}\xi\leqslant C\|u_0\|_E G_{T-r}(x,y).
$$

Therefore, by (6.2), it holds for $\nu\in(\frac{1}{3},1)$ that

$$
|\mathscr{I}_{initial\_u}|\leqslant\sum_{i=1}^{N}\int_0^1\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1|\mathscr{R}^i_{initial\_u}(r,y)|\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\mathrm{d}\tau
$$

$$\leqslant C\int_{t_0}^{t_1}\int_0^1 G_{T-r}(x,y)\mathrm{d}y\mathrm{d}r+\sum_{i=2}^{N}\int_0^1\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1|\mathscr{R}^i_{initial\_u}(r,y)|\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\mathrm{d}\tau$$
$$\leqslant C\delta+C\sum_{i=2}^{N}\int_0^1 G_{T-r}(x,y)\mathrm{d}y\int_{t_{i-1}}^{t_i}(r-t_{i-1})^{\nu}(t_{i-1})^{-\nu}\mathrm{d}r\leqslant C\delta+C\delta^{\nu}\int_0^T\frac{1}{r^{\nu}}\mathrm{d}r\leqslant C(\nu,T)\delta^{\nu}.$$

(b) Estimate of $\mathscr{I}_{initial\_b}$. By (6.3) (with $Q_i=E^i_{initial\_b}$, $\alpha=1$ and $g=b$), (6.1) and Proposition 2.2,

$$|\mathscr{R}^i_{initial\_b}(r,y)|\leqslant CG_{T-r}(x,y)\left\|\int_0^{t_{i-1}}\int_0^1\{G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)\}b\left(\Phi^{\xi}_{\lfloor\theta\rfloor}(0,u_0)\right)\mathrm{d}\xi\mathrm{d}\theta\right\|_{2,8}$$
$$\leqslant CG_{T-r}(x,y)\int_0^{t_{i-1}}\int_0^1|G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)|\|b(\Phi^{\xi}_{\lfloor\theta\rfloor}(0,u_0))\|_{2,8}\mathrm{d}\xi\mathrm{d}\theta.$$

By further taking into account Corollary 4.4 and Lemma 2.1, we arrive at

$$|\mathscr{R}^i_{initial\_b}(r,y)|\leqslant CG_{T-r}(x,y)\int_0^{t_{i-1}}\int_0^1|G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)|\mathrm{d}\xi\mathrm{d}\theta$$
$$\leqslant C(\nu)G_{T-r}(x,y)\int_0^{t_{i-1}}(r-t_{i-1})^{\nu}(t_{i-1}-\theta)^{-\nu}\mathrm{d}\theta$$

for $\nu\in(1/3,1)$. Hence, in view of (6.1), it holds that

$$|\mathscr{I}_{initial\_b}|\leqslant C(\nu)\sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}\int_0^1 G_{T-r}(x,y)\mathrm{d}y\int_0^{t_{i-1}}(r-t_{i-1})^{\nu}(t_{i-1}-\theta)^{-\nu}\mathrm{d}\theta\mathrm{d}r$$
$$\leqslant C(\nu)\delta^{\nu}\sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}\int_0^{t_{i-1}}(t_{i-1}-\theta)^{-\nu}\mathrm{d}\theta\mathrm{d}r\leqslant C(\nu,T)\delta^{\nu}.$$

(c) Estimate of $\mathscr{I}_{initial\_\sigma}$. By the Malliavin integration by parts formula (2.11) and the chain rule (see e.g. (Nualart, 2006, Proposition 1.5.1)), we obtain

$$\begin{aligned}
&\mathscr{R}^i_{initial_\sigma}(r,y)\\
&=\int_0^{t_{i-1}}\int_0^1\mathbb{E}\big[f''(\varphi^x_T(t_i,Y^\tau_i))D_{\theta,\xi}\varphi^x_T(t_i,Y^\tau_i)\langle\mathscr{D}\varphi^x_T(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle b'(Z^\beta_i(r,y))\big]\{G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)\}\sigma\mathrm{d}\xi\mathrm{d}\theta\\
&\quad+\int_0^{t_{i-1}}\int_0^1\mathbb{E}\big[f'(\varphi^x_T(t_i,Y^\tau_i))D_{\theta,\xi}\langle\mathscr{D}\varphi^x_T(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle b'(Z^\beta_i(r,y))\big]\{G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)\}\sigma\mathrm{d}\xi\mathrm{d}\theta\\
&\quad+\int_0^{t_{i-1}}\int_0^1\mathbb{E}\big[f'(\varphi^x_T(t_i,Y^\tau_i))\langle\mathscr{D}\varphi^x_T(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle b''(Z^\beta_i(r,y))D_{\theta,\xi}Z^\beta_i(r,y)\big]\{G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)\}\sigma\mathrm{d}\xi\mathrm{d}\theta\\
&=:\mathscr{R}^{i,1}_{initial_\sigma}(r,y)+\mathscr{R}^{i,2}_{initial_\sigma}(r,y)+\mathscr{R}^{i,3}_{initial_\sigma}(r,y).
\end{aligned}$$

(c1) Estimate of $\mathscr{R}^{i,1}_{initial\_\sigma}(r,y)$: We apply (6.3) with $g=b$, $\alpha=2$, $Q_i(r,y)=D_{\theta,\xi}\varphi^x_T(t_i,Y^\tau_i)$ to obtain

$$\begin{aligned}
&|\mathbb{E}\big[f''(\varphi^x_T(t_i,Y^\tau_i))D_{\theta,\xi}\varphi^x_T(t_i,Y^\tau_i)\langle\mathscr{D}\varphi^x_T(t_i,Y^\tau_i),G_{t_i-r}(\cdot,y)\rangle b'(Z^\beta_i(r,y))\big]|\\
&\leqslant CG_{T-r}(x,y)\|D_{\theta,\xi}\varphi^x_T(t_i,Y^\tau_i)\|_{3,8}\leqslant CG_{T-r}(x,y)G_{T-\theta}(x,\xi),
\end{aligned}\tag{6.4}$$

where we have also used (5.19) in the last step. By Lemma 2.1 and

$$G_s(x,y)\leqslant KP_s(x,y)\leqslant Cs^{-\frac{1}{2}},\quad s\in(0,T],\tag{6.5}$$

we obtain for $\nu\in(\frac{1}{3},1)$,

$$\sum_{i=1}^{N}\int_0^1\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1|\mathscr{R}^{i,1}_{initial\_\sigma}(r,y)|\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\mathrm{d}\tau$$

$$\leqslant C\sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}\int_0^1\int_0^{t_{i-1}}\int_0^1 G_{T-r}(x,y)G_{T-\theta}(x,\xi)|G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)|\mathrm{d}\xi\mathrm{d}\theta\mathrm{d}y\mathrm{d}r$$
$$\leqslant C\sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}\int_0^1\int_0^{t_{i-1}}(T-r)^{-\frac{1}{2}}G_{T-\theta}(x,\xi)\int_0^1|G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)|\mathrm{d}y\mathrm{d}\theta\mathrm{d}\xi\mathrm{d}r$$
$$\leqslant C\delta^\nu\sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}\int_0^{t_{i-1}}(T-r)^{-\frac{1}{2}}(t_{i-1}-\theta)^{-\nu}\mathrm{d}\theta\mathrm{d}r\leqslant C\delta^\nu\sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}(T-r)^{-\frac{1}{2}}\mathrm{d}r\int_0^{t_{i-1}}(t_{i-1}-\theta)^{-\nu}\mathrm{d}\theta\leqslant C\delta^\nu.$$

(c2) Estimate of $\mathscr{R}^{i,2}_{initial\_\sigma}(r,y)$: We use Lemma 3.5, Proposition 2.2, Corollary 4.4 and Lemma 5.8 to get

$$\begin{aligned}&|\mathbb{E}\big[f'(\varphi_T^x(t_i,Y_i^\tau))D_{\theta,\xi}\langle\mathscr{D}\varphi_T^x(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\rangle b'(Z_i^\beta(r,y))\big]|\\&\leqslant C\|D_{\theta,\xi}\langle\mathscr{D}\varphi_T^x(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\rangle b'(Z_i^\beta(r,y))\|_{2,2}\\&\leqslant C\|D_{\theta,\xi}\langle\mathscr{D}\varphi_T^x(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\rangle\|_{2,4}\|b'(Z_i^\beta(r,y))\|_{2,4}\\&\leqslant C\|D_{\theta,\xi}\langle\mathscr{D}\varphi_T^x(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\rangle\|_{2,4}\leqslant CG_{T-\theta}(x,\xi).\end{aligned}\tag{6.6}$$

Thus, it follows from the definition of $\mathscr{R}^{i,2}_{initial\_\sigma}(r,y)$ and (6.6) that

$$|\mathscr{R}^{i,2}_{initial\_\sigma}(r,y)|\leqslant C\int_0^{t_{i-1}}\int_0^1 G_{T-\theta}(x,\xi)|G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)|\mathrm{d}\xi\mathrm{d}\theta.$$

Using Lemma 2.1, we have for $\nu\in(\frac{1}{3},1)$,

$$\begin{aligned}&\int_{t_{i-1}}^{t_i}\int_0^1\int_0^{t_{i-1}}\int_0^1 G_{T-\theta}(x,\xi)|G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)|\mathrm{d}\xi\mathrm{d}\theta\mathrm{d}y\mathrm{d}r\\&\leqslant\int_{t_{i-1}}^{t_i}\int_0^{t_{i-1}}\int_0^1 CG_{T-\theta}(x,\xi)(r-t_{i-1})^\nu(t_{i-1}-\theta)^{-\nu}\mathrm{d}\xi\mathrm{d}\theta\mathrm{d}r\\&\leqslant C\delta^\nu\int_{t_{i-1}}^{t_i}\int_0^{t_{i-1}}(t_{i-1}-\theta)^{-\nu}\mathrm{d}\theta\mathrm{d}r\leqslant C\delta^{1+\nu},\end{aligned}$$

which shows

$$\sum_{i=1}^{N}\int_0^1\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1|\mathscr{R}^{i,2}_{initial\_\sigma}(r,y)|\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\mathrm{d}\tau\leqslant C\delta^\nu.$$

(c3) Estimate of $\mathscr{R}^{i,3}_{initial\_\sigma}(r,y)$: Notice that by (6.5), $G_{r-\theta}(y,\xi)\leqslant C(r-\theta)^{-\frac{1}{2}}$. By (6.3) (with $Q_i(r,y)=D_{\theta,\xi}Z_i^\beta(r,y)$, $\alpha=1$, and $g=b'$) and (5.20),

$$\begin{aligned}&\big|\mathbb{E}\big[f'(\varphi_T^x(t_i,Y_i^\tau))\langle\mathscr{D}\varphi_T^x(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\rangle b''(Z_i^\beta(r,y))D_{\theta,\xi}Z_i^\beta(r,y)\big]\big|\\&\leqslant CG_{T-r}(x,y)G_{r-\theta}(y,\xi)\leqslant C(r-\theta)^{-\frac{1}{2}}G_{T-r}(x,y).\end{aligned}\tag{6.7}$$

Hence, it follows that

$$|\mathscr{R}^{i,3}_{initial\_\sigma}(r,y)|\leqslant\int_0^{t_{i-1}}\int_0^1 CG_{T-r}(x,y)(r-\theta)^{-\frac{1}{2}}|G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)|\mathrm{d}\xi\mathrm{d}\theta,$$

which yields

$$\sum_{i=1}^{N}\int_0^1\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1|\mathscr{R}^{i,3}_{initial\_\sigma}(r,y)|\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\mathrm{d}\tau$$

$$\leqslant \sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}\int_0^1 CG_{T-r}(x,y)\int_0^{t_{i-1}}(r-\theta)^{-\frac{1}{2}}\int_0^1 |G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)|\mathrm{d}\xi\mathrm{d}\theta\mathrm{d}y\mathrm{d}r. \tag{6.8}$$

Applying Lemma 2.1 with $\nu\in(\frac{1}{2},1)$ yields

$$\begin{aligned}
&\sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}\int_0^1 CG_{T-r}(x,y)\int_0^{t_{i-2}}(r-\theta)^{-\frac{1}{2}}\int_0^1 |G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)|\mathrm{d}\xi\mathrm{d}\theta\mathrm{d}y\mathrm{d}r\\
&\leqslant \sum_{i=1}^{N}\int_{t_{i-1}}^{t_i} C\int_0^{t_{i-2}}(r-\theta)^{-\frac{1}{2}}(r-t_{i-1})^{\nu}(t_{i-1}-\theta)^{-\nu}\mathrm{d}\theta\mathrm{d}r\\
&\leqslant C\delta^{\nu}\sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}\int_0^{t_{i-2}}(r-\theta)^{-\frac{1}{2}}(t_{i-1}-\theta)^{-\nu}\mathrm{d}\theta\mathrm{d}r\leqslant C\delta^{\frac{1}{2}},
\end{aligned} \tag{6.9}$$

since

$$\begin{aligned}
&\int_{t_{i-1}}^{t_i}\int_0^{t_{i-2}}(r-\theta)^{-\frac{1}{2}}(t_{i-1}-\theta)^{-\nu}\mathrm{d}\theta\mathrm{d}r=\int_0^{t_{i-2}}\int_{t_{i-1}}^{t_i}(r-\theta)^{-\frac{1}{2}}\mathrm{d}r(t_{i-1}-\theta)^{-\nu}\mathrm{d}\theta\\
&=\int_0^{t_{i-2}}\frac{2\delta}{(t_i-\theta)^{\frac{1}{2}}+(t_{i-1}-\theta)^{\frac{1}{2}}}(t_{i-1}-\theta)^{-\nu}\mathrm{d}\theta\leqslant\delta\int_0^{t_{i-2}}(t_{i-1}-\theta)^{-\nu-\frac{1}{2}}\mathrm{d}\theta\leqslant C\delta^{-\nu+\frac{3}{2}}.
\end{aligned}$$

Besides, (2.4) implies $\int_0^1|G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)|\mathrm{d}\xi\leqslant 2$, and thus

$$\begin{aligned}
&\sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}\int_0^1 CG_{T-r}(x,y)\int_{t_{i-2}}^{t_{i-1}}(r-\theta)^{-\frac{1}{2}}\int_0^1 |G_{r-\theta}(y,\xi)-G_{t_{i-1}-\theta}(y,\xi)|\mathrm{d}\xi\mathrm{d}\theta\mathrm{d}y\mathrm{d}r\\
&\leqslant C\sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}\int_0^1\int_{t_{i-2}}^{t_{i-1}}G_{T-r}(x,y)(r-\theta)^{-\frac{1}{2}}\mathrm{d}\theta\mathrm{d}y\mathrm{d}r\leqslant C\sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}\int_{t_{i-2}}^{t_{i-1}}(r-t_{i-1})^{-\frac{1}{2}}\mathrm{d}\theta\mathrm{d}r\leqslant C\delta^{\frac{1}{2}}.
\end{aligned} \tag{6.10}$$

Combining (6.8), (6.9) and (6.10), we obtain

$$\sum_{i=1}^{N}\int_0^1\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1|\mathscr{R}^{i,3}_{initial\_\sigma}(r,y)|\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\mathrm{d}\tau\leqslant C\delta^{\frac{1}{2}}.$$

(d) Estimate of $\mathscr{I}_b$. By (6.3) (with $Q_i=E_b^i$, $\alpha=1$ and $g=b$), (6.1), Corollary 4.4, the Minkowski inequality, and (2.4),

$$\begin{aligned}
|\mathscr{R}^i_b(r,y)|&\leqslant CG_{T-r}(x,y)\left\|\int_{t_{i-1}}^{r}\int_0^1 G_{r-\theta}(y,\xi)b\left(\varphi^{\xi}_{\theta}(t_{i-1},\Phi_{t_{i-1}}(0,u_0))\right)\mathrm{d}\xi\mathrm{d}\theta\right\|_{2,8}\\
&\leqslant CG_{T-r}(x,y)\int_{t_{i-1}}^{r}\int_0^1 G_{r-\theta}(y,\xi)\left\|b(\varphi^{\xi}_{\theta}(t_{i-1},\Phi_{t_{i-1}}(0,u_0)))\right\|_{2,8}\mathrm{d}\xi\mathrm{d}\theta\\
&\leqslant CG_{T-r}(x,y)\int_{t_{i-1}}^{r}\int_0^1 G_{r-\theta}(y,\xi)\mathrm{d}\xi\mathrm{d}\theta\leqslant CG_{T-r}(x,y)(r-t_{i-1}).
\end{aligned}$$

It follows from (6.2) that

$$\mathscr{I}_b\leqslant C\sum_{i=1}^{N}\int_{t_{i-1}}^{t_i}\int_0^1 G_{T-r}(x,y)(r-t_{i-1})\mathrm{d}y\mathrm{d}r\leqslant C\delta.$$

(e) Estimate of $\mathscr{I}_\sigma$. We apply the Malliavin integration by parts formula (2.11) to get

$$\mathscr{R}^i_\sigma(r,y)=\sigma\int_{t_{i-1}}^{r}\int_0^1\mathbb{E}\big[f''(\varphi^x_T(t_i,Y^{\tau}_i))D_{\theta,\xi}\varphi^x_T(t_i,Y^{\tau}_i)\langle\mathscr{D}\varphi^x_T(t_i,Y^{\tau}_i),G_{t_i-r}(\cdot,y)\rangle b'(Z^{\beta}_i(r,y))\big]G_{r-\theta}(y,\xi)\mathrm{d}\xi\mathrm{d}\theta$$

$$+\sigma\int_{t_{i-1}}^{r}\int_0^1\mathbb{E}\big[f'(\varphi_T^x(t_i,Y_i^\tau))D_{\theta,\xi}\langle\mathscr{D}\varphi_T^x(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\rangle b'(Z_i^\beta(r,y))\big]G_{r-\theta}(y,\xi)\mathrm{d}\xi\mathrm{d}\theta$$
$$+\sigma\int_{t_{i-1}}^{r}\int_0^1\mathbb{E}\big[f'(\varphi_T^x(t_i,Y_i^\tau))\langle\mathscr{D}\varphi_T^x(t_i,Y_i^\tau),G_{t_i-r}(\cdot,y)\rangle b''(Z_i^\beta(r,y))D_{\theta,\xi}Z_i^\beta(r,y)\big]G_{r-\theta}(y,\xi)\mathrm{d}\xi\mathrm{d}\theta$$
$$=:\mathscr{R}_\sigma^{i,1}(r,y)+\mathscr{R}_\sigma^{i,2}(r,y)+\mathscr{R}_\sigma^{i,3}(r,y).$$

(e1) Estimate of $\mathscr{R}_\sigma^{i,1}(r,y)$: Using (6.4), we get

$$|\mathscr{R}_\sigma^{i,1}(r,y)|\leqslant C\int_{t_{i-1}}^{r}\int_0^1 G_{T-r}(x,y)G_{r-\theta}(y,\xi)G_{T-\theta}(x,\xi)\mathrm{d}\xi\mathrm{d}\theta.$$

Therefore, (2.5) and (2.8) give

$$\sum_{i=1}^N\int_0^1\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1|\mathscr{R}_\sigma^{i,1}(r,y)|\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\mathrm{d}\tau\leqslant C\sum_{i=1}^N\int_{t_{i-1}}^{t_i}\int_0^1\int_{t_{i-1}}^{r}\int_0^1 G_{T-r}(x,y)G_{r-\theta}(y,\xi)G_{T-\theta}(x,\xi)\mathrm{d}\xi\mathrm{d}\theta\mathrm{d}y\mathrm{d}r$$
$$\leqslant C\sum_{i=1}^N\int_{t_{i-1}}^{t_i}\int_{t_{i-1}}^{t_i}\int_0^1 G_{T-\theta}^2(x,\xi)\mathrm{d}\xi\mathrm{d}\theta\mathrm{d}r\leqslant C\delta\sum_{i=1}^N\int_{t_{i-1}}^{t_i}\int_0^1 G_{T-\theta}^2(x,\xi)\mathrm{d}\xi\mathrm{d}\theta\leqslant C\delta.$$

(e2) Estimate of $\mathscr{R}_\sigma^{i,2}(r,y)$: By the definition of $\mathscr{R}_\sigma^{i,2}(r,y)$ and (6.6), we have

$$|\mathscr{R}_\sigma^{i,2}(r,y)|\leqslant C\int_{t_{i-1}}^{r}\int_0^1 G_{T-\theta}(x,\xi)G_{r-\theta}(y,\xi)\mathrm{d}\xi\mathrm{d}\theta.$$

In view of (2.5),

$$\int_{t_{i-1}}^{t_i}\int_0^1\int_{t_{i-1}}^{r}\int_0^1 G_{T-\theta}(x,\xi)G_{r-\theta}(y,\xi)\mathrm{d}\xi\mathrm{d}\theta\mathrm{d}y\mathrm{d}r=\int_{t_{i-1}}^{t_i}\int_{t_{i-1}}^{r}\int_0^1 G_{T-\theta}(x,\xi)\mathrm{d}\xi\int_0^1 G_{r-\theta}(y,\xi)\mathrm{d}y\mathrm{d}\theta\mathrm{d}r\leqslant C\delta^2,$$

from which we deduce that

$$\sum_{i=1}^N\int_0^1\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1|\mathscr{R}_\sigma^{i,2}(r,y)|\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\mathrm{d}\tau\leqslant C\delta.$$

(e3) Estimate of $\mathscr{R}_\sigma^{i,3}(r,y)$: Due to (6.7) and the definition of $\mathscr{R}_\sigma^{i,3}(r,y)$,

$$|\mathscr{R}_\sigma^{i,3}(r,y)|\leqslant\int_{t_{i-1}}^{r}\int_0^1 CG_{T-r}(x,y)(r-\theta)^{-\frac12}G_{r-\theta}(y,\xi)\mathrm{d}\xi\mathrm{d}\theta$$
$$\leqslant CG_{T-r}(x,y)\int_{t_{i-1}}^{r}(r-\theta)^{-\frac12}\int_0^1 G_{r-\theta}(y,\xi)\mathrm{d}\xi\mathrm{d}\theta\leqslant CG_{T-r}(x,y)\int_{t_{i-1}}^{r}(r-\theta)^{-\frac12}\mathrm{d}\theta.$$

Therefore, it holds that

$$\sum_{i=1}^N\int_0^1\int_0^1\int_{t_{i-1}}^{t_i}\int_0^1|\mathscr{R}_\sigma^{i,3}(r,y)|\mathrm{d}y\mathrm{d}r\mathrm{d}\beta\mathrm{d}\tau\leqslant C\sum_{i=1}^N\int_{t_{i-1}}^{t_i}\int_0^1 G_{T-r}(x,y)\int_{t_{i-1}}^{r}(r-\theta)^{-\frac12}\mathrm{d}\theta\mathrm{d}y\mathrm{d}r$$
$$\leqslant C\sum_{i=1}^N\int_{t_{i-1}}^{t_i}\int_{t_{i-1}}^{r}(r-\theta)^{-\frac12}\mathrm{d}\theta\mathrm{d}r\leqslant C\delta^{\frac12}.$$

Gathering all above estimates, we complete the proof of (3.4).

If $b(u)=b_1u+c$ is an affine function, then $b''(Z_i^\beta(r,y))\equiv 0$. Therefore $\mathscr{R}_\sigma^{i,3}(r,y)=\mathscr{R}_{initial\_\sigma}^{i,3}(r,y)=0$, $i=1,\dots,N$. In this case, by combining the estimates **(a)**-**(e)**, we have, instead of (3.4), that (3.5) holds for every $\mu\in(\frac12,1)$. The proof is completed. □

Acknowledgements

This work is funded by National Natural Science Foundation of China (No. 11971470, No. 11871068, No. 12031020, No. 12022118) and by Youth Innovation Promotion Association CAS. The research of J. C. is partially supported by start-up funds from Hong Kong Polytechnic University and the CAS AMSS-PolyU Joint Laboratory of Applied Mathematics. The authors are very grateful to Charles-Edouard Bréhier (Université Lyon 1) for his comments.

REFERENCES

ANDERSSON, A. & LARSSON, S. (2016) Weak convergence for a spatial approximation of the nonlinear stochastic heat equation. *Math. Comp.*, **85**, 1335–1358.

BALLY, V. & CARAMELLINO, L. (2014) On the distances between probability density functions. *Electron. J. Probab.*, **19**, no. 110, 33.

BALLY, V. & PARDOUX, E. (1998) Malliavin calculus for white noise driven parabolic SPDEs. *Potential Anal.*, **9**, 27–64.

BALLY, V. & TALAY, D. (1996) The law of the Euler scheme for stochastic differential equations. II. Convergence rate of the density. *Monte Carlo Methods Appl.*, **2**, 93–128.

BEHMARDI, D. & NAYERI, E. D. (2008) Introduction of Fréchet and Gâteaux derivative. *Appl. Math. Sci. (Ruse)*, **2**, 975–980.

BERTINI, L. & CANCRINI, N. (1995) The stochastic heat equation: Feynman-Kac formula and intermittence. *J. Statist. Phys.*, **78**, 1377–1401.

BRÉHIER, C.-E. (2020) Influence of the regularity of the test functions for weak convergence in numerical discretization of SPDEs. *J. Complexity*, **56**, 101424, 15.

BRÉHIER, C.-E. & DEBUSSCHE, A. (2018) Kolmogorov equations and weak order analysis for SPDEs with nonlinear diffusion coefficient. *J. Math. Pures Appl. (9)*, **119**, 193–254.

COX, S. & VAN NEERVEN, J. (2010) Convergence rates of the splitting scheme for parabolic linear stochastic Cauchy problems. *SIAM J. Numer. Anal.*, **48**, 428–451.

CUI, J., HONG, J. & SHENG, D. (2019) Convergence in density of splitting AVF scheme for stochastic Langevin equation. *arXiv:1906.03439*.

CUI, J., HONG, J. & SUN, L. (2021) Weak convergence and invariant measure of a full discretization for parabolic SPDEs with non-globally Lipschitz coefficients. *Stochastic Process. Appl.*, **134**, 55–93.

CUI, J. & HONG, J. (2019) Strong and weak convergence rates of a spatial approximation for stochastic partial differential equation with one-sided Lipschitz coefficient. *SIAM J. Numer. Anal.*, **57**, 1815–1841.

CUI, J. & HONG, J. (2020) Absolute continuity and numerical approximation of stochastic Cahn-Hilliard equation with unbounded noise diffusion. *J. Differential Equations*, **269**, 10143–10180.

DEBUSSCHE, A. (2011) Weak approximation of stochastic partial differential equations: the nonlinear case. *Math. Comp.*, **80**, 89–117.

DEBUSSCHE, A. & PRINTEMS, J. (2009) Weak order for the discretization of the stochastic heat equation. *Math. Comp.*, **78**, 845–863.

GYÖNGY, I. (1998) Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. I. *Potential Anal.*, **9**, 1–25.

GYÖNGY, I. (1999) Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. II. *Potential Anal.*, **11**, 1–37.

HONG, J. & WANG, X. (2019) *Invariant measures for stochastic nonlinear Schrödinger equations: Numerical approximations and symplectic structures*. Lecture Notes in Mathematics, vol. 2251. Springer, Singapore, pp. xiv+220.

JENTZEN, A., KLOEDEN, P. & WINKEL, G. (2011) Efficient simulation of nonlinear parabolic SPDEs with additive noise. *Ann. Appl. Probab.*, **21**, 908–950.

JENTZEN, A. & KLOEDEN, P. E. (2009) Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive space-time noise. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, **465**, 649–667.

KESAVAN, S. (2020) *Nonlinear functional analysis—a first course*. Texts and Readings in Mathematics, vol. 28, second edn. Hindustan Book Agency, New Delhi, pp. xii+164.

MARINELLI, C. & SCARPA, L. (2020) Fréchet differentiability of mild solutions to SPDEs with respect to the initial datum. *J. Evol. Equ.*, **20**, 1093–1130.

MISHURA, Y., RALCHENKO, K., ZILI, M. & ZOUGAR, E. (2021) Fractional stochastic heat equation with piecewise constant coefficients. *Stoch. Dyn.*, **21**, 2150002, 39.

MUELLER, C. & NUALART, D. (2008) Regularity of the density for the stochastic heat equation. *Electron. J. Probab.*, **13**, 2248–2258.

NUALART, D. (2006) *The Malliavin calculus and related topics*. Probability and its Applications (New York), second edn. Springer-Verlag, Berlin, pp. xiv+382.

NUALART, D. & QUER-SARDANYONS, L. (2009) Gaussian density estimates for solutions to quasi-linear stochastic partial differential equations. *Stochastic Process. Appl.*, **119**, 3914–3938.

SANZ-SOLÉ, M. (2005) *Malliavin calculus*. Fundamental Sciences. EPFL Press, Lausanne; distributed by CRC Press, Boca Raton, FL, pp. viii+162. With applications to stochastic partial differential equations.

SANZ-SOLÉ, M. (2008) Properties of the density for a three-dimensional stochastic wave equation. *J. Funct. Anal.*, **255**, 255–281.

SERFLING, R. J. (1980) *Approximation theorems of mathematical statistics*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, pp. xiv+371.

WALSH, J. B. (1986) An introduction to stochastic partial differential equations. *École d'été de probabilités de Saint-Flour, XIV—1984*. Lecture Notes in Math., vol. 1180. Springer, Berlin, pp. 265–439.

WANG, X. & QI, R. (2015) A note on an accelerated exponential Euler method for parabolic SPDEs with additive noise. *Appl. Math. Lett.*, **46**, 31–37.

YAN, Y. (2005) Galerkin finite element methods for stochastic parabolic partial differential equations. *SIAM J. Numer. Anal.*, **43**, 1363–1384.