Journal of Mathematical Sciences, Vol. 179, No. 1, November, 2011

ON FINITE DIFFERENCE SCHEMES FOR DEGENERATE STOCHASTIC PARABOLIC PARTIAL DIFFERENTIAL EQUATIONS

I. Gyöngy

School of Mathematics and Maxwell Institute, University of Edinburgh
King's Buildings, Edinburgh, EH9 3JZ, United Kingdom
gyongy@maths.ed.ac.uk

UDC 517.95

Dedicated to Professor N. V. Krylov
on the occasion of his 70th birthday
with gratitude and admiration

Finite difference approximations in the space variable for possibly degenerate stochastic parabolic partial differential equations are investigated. Sharp estimates for the rate of convergence are obtained, and sufficient conditions are presented under which the speed of approximations can be accelerated to any given order of convergence by Richardson's method. The main theorems generalize some results of the author with N. V. Krylov. Bibliography: 10 *titles.*

1 Introduction

We study spatial discretizations

$$du^h_t(x) = (L^h_t(x)u^h_t(x) + f_t(x))\,dt + \sum_{\rho=1}^{\infty}(M^{h\rho}_t u^h_t(x) + g^\rho_t(x))\,dw^\rho_t, \tag{1.1}$$

$t \in [0,T]$, $x \in \mathbb{G}_h$, for stochastic parabolic partial differential equations

$$du_t(x) = (L_t u_t(x) + f_t(x))\,dt + \sum_{\rho=1}^{\infty}(M^\rho_t u_t(x) + g^\rho_t(x))\,dw^\rho_t, \tag{1.2}$$

$t \in [0,T]$, $x \in \mathbb{R}^d$, with initial condition

$$u_0(x) = \psi(x), \quad x \in \mathbb{R}^d. \tag{1.3}$$

Here $(w^\rho)_{\rho=1}^\infty$ is a sequence of independent $\mathscr{F}_t$-Wiener processes carried on a probability space $(\Omega, \mathscr{F}, P)$, equipped with the filtration $\mathbb{F} = (\mathscr{F}_t)_{t\geqslant 0}$. The operators L and M^ρ, $\rho = 1, 2, ...$, are differential operators in x, with random time dependent coefficients, adapted to the filtration $\mathbb{F}$, such that L is a second order differential operator and M^ρ are first order operators, of the form

$$L = \sum_{\alpha,\beta=0}^{d} a^{\alpha\beta}_t(x) D_\alpha D_\beta \quad \text{and} \quad M^\rho = \sum_{\alpha=0}^{d} b^{\alpha\rho} b_t(x) D_\alpha, \quad \rho = 1, 2, ...,$$

Translated from *Problems in Mathematical Analysis* **61**, October 2011, pp. 87–108.

respectively. The stochastic parabolicity condition is assumed (cf. Assumption 2.1 below). Such equations arise in filtering theory of partially observed diffusion processes $Z = (X, Y)$, as equations for the unnormalized conditional density of the *signal process* X at time t, given the *observation process* Y until time t. Therefore, effective numerical algorithms for solving (1.2)-(1.3) are of great practical importance. There are many methods introduced to solve (1.2)-(1.3) numerically. We take here finite difference operators L^h and $M^{h\rho}$ to approximate the solution u of (1.2)-(1.3) by the solution u^h of (1.1) with initial condition $u_0^h = \psi$ on a fixed grid $\mathbb{G}_h$ of mesh-size $|h|$.

Finite difference approximations for deterministic partial differential equations are studied extensively in the literature (cf., for instance, [1] and the references therein). However, there are only a few results published for degenerate equations. Sharp rate of convergence estimates are obtained in [2] for deterministic (possibly) degenerate parabolic and elliptic SPDEs with monotone finite difference schemes. Rate of convergence estimates of finite difference approximations for stochastic parabolic partial differential equations are obtained under the *strong stochastic parabolicity condition*, i.e., when there is a constant $\kappa > 0$ such that

$$(2a^{ij} - b^{i\rho}b^{j\rho})z^i z^j \geqslant \kappa z^i z^i$$

for all $\omega \in \Omega$, $t \geqslant 0$, and $x \in \mathbb{R}^d$.

About hundred years ago L. F. Richardson suggested a method of accelerating the convergence of numerical approximations depending on a parameter, for example on the mesh-size $|h|$ of the grid in the case of finite difference approximations (cf. [3] and [4]). He demonstrated that the accuracy of the approximations can be dramatically increased if one takes suitable mixtures of approximations with different step-sizes. His idea is based on the existence of an expansion of the finite difference approximation in powers of the step-size, which makes it possible to find such mixtures where the lower order powers are cancelled out. Therefore, it is important to find sufficient conditions under which numerical approximations admit power expansions with respect to a parameter which is related to the error of the method. The possibility of such expansions have been studied thoroughly in numerical analysis (cf., for example, the book [5] on Richardson's idea applied to finite difference approximations for deterministic partial differential equations). In [6], Richardson's idea is implemented to a class of monotone finite difference schemes for (possibly) degenerate parabolic and elliptic partial differential equations, and, in [7], Richardson's idea is implemented to stochastic partial differential equations satisfying the strong parabolicity conditions. Both in[6] and [7], general conditions are obtained under which the accuracy of finite difference approximations in the supremum norm can be made as high as desired. In the present paper, we generalize some results from [6] and [7] to SPDEs satisfying only the stochastic parabolicity conditions. We present sharp rate of convergence estimate and give sufficient conditions under which the accuracy of the accelerated schemes is as high as we wish. In the special case where the finite difference approximations are defined by replacing the partial derivatives ∂x^i by centered finite differences along the basis vector e_i, our main theorem reads as follows: The rate of convergence of the (spatial) finite difference approximations to (1.2)-(1.3) can be accelerated to any order if the initial condition, coefficients, and free terms are sufficiently smooth in x and the matrix

$$\widetilde{a}_t(x) := (2a^{ij} - b^{i\rho}b^{j\rho})$$

can be decomposed as

$$a_t(x) = \sigma_t(x)\sigma_t^T(x) \tag{1.4}$$

by a sufficiently smooth matrix σ in x. Clearly, requiring a sufficiently smooth factorization (1.4) is a rather restrictive condition. Nevertheless this condition is easily applicable to the equation of the unnormalized conditional density in nonlinear filtering since this factorization condition is satisfied even in the general setting of correlated signal and observation noises when the diffusion coefficients of the signal noise is sufficiently smooth.

For survey papers on the application of Richardson's method to various numerical approximations we refer to [8] and [9].

The paper is organized as follows. In Section 2, basic notions and notation are introduced and the main results are presented. In Section 3, the main tools are given. The proof of the main theorems are given in the last section, Section 4

We fix a probability space $(\Omega, \mathscr{F}, P)$ equipped with an increasing family of σ-algebras $(\mathscr{F}_t)_{t\geqslant 0}$ such that $\mathscr{F}_0$ contains the P-zero sets of $\mathscr{F}$. The σ-algebra of predictable subsets of $\Omega \times [0,\infty)$ is denoted by $\mathscr{P}$. We fix also a sequence of independent Wiener processes $(w_t^\rho)_{\rho=1}^\infty$, such that w_t^ρ is $\mathscr{F}_t$-measurable and $w_t^\rho - w_s^\rho$ is independent of $\mathscr{F}_s$ for $0 \leqslant s \leqslant t$, for every integer $\rho \geqslant 1$. Unless otherwise stated, the summation convention with respect to repeated integer-valued indices is used throughout the paper.

2 Formulation of the Main Results

We consider the equation

$$du_t = (\mathscr{L}_t u_t + f_t)\,dt + (\mathscr{M}_t^\rho u_t + g_t^\rho)\,dw_t^\rho \tag{2.1}$$

for $\omega \in \Omega$, $(t,x) \in [0,T] \times \mathbb{R}^d =: H_T$ with some initial condition

$$u_0(x) = \psi(x), \quad x \in \mathbb{R}^d, \tag{2.2}$$

where

$$\mathscr{L}_t\phi = a_t^{\alpha\beta} D_\alpha D_\beta \phi, \quad \mathscr{M}_t^\rho \phi = b_t^{\alpha\rho} D_\alpha \phi.$$

Here and below, the summation with respect to α and β is performed over the set $\{0,1,...,d\}$ and with respect to ρ, over the positive integers $\{1,2,...\}$. Assume that $a_t^{\alpha\beta} = a_t^{\alpha\beta}(x)$ are real-valued, $b_t^\alpha = (b_t^{\alpha\rho}(x))_{\rho=1}^\infty$ are l_2-valued $\mathscr{P} \times \mathscr{B}(\mathbb{R}^d)$-measurable functions on $\Omega \times H_T$ for all $\alpha, \beta \in \{0,1,...\}$.

A necessary condition for the well-posedness of the Cauchy problem (2.1)-(2.2) is the condition of *stochastic parabolicity*:

Assumption 2.1. For all $(\omega,t,x) \in \Omega \times H_T$ and $z \in \mathbb{R}^d$

$$\sum_{i,j=1}^d (2a_t^{ij} - b_t^{i\rho} b_t^{j\rho}) z^i z^j \geqslant 0.$$

To formulate an existence and uniqueness theorem for the *generalized solution*, we also require smoothness conditions on the coefficients $a^{\alpha\beta}$, b^α, the initial value ψ, and free terms f and g.

Let $m \geqslant 0$ be an integer, and let W_2^m be the usual Hilbert-Sobolev space of functions on $\mathbb{R}^d$ with norm $\|\cdot\|_{W_2^m}$.

Assumption 2.2. For each (ω, t) the functions a_t^{ij} are $\max(m,2)$ times, the functions a_t^{0i}, a_t^{0i}, a_t^{00} are m times continuously differentiable in x for $i,j \in \{1,...,d\}$. The l_2-valued functions $b_t^{\alpha} = (b^{\alpha\rho})_{\rho=1}^{\infty}$ are m-times continuously differentiable in x. There are constants K_l, $l = 0,..., \max(m,2)$ such that

$$|D^l a_t^{ij}| \leqslant K_l \quad \text{for } l \leqslant \max(m,2),$$

$$|D^l a^{\alpha 0}| \leqslant K_l, \quad |D^l a^{0\alpha}| \leqslant K_l, \quad |D^l b_t^{\alpha}|_{l_2} \leqslant K_l, \quad |D^l b_t|_{l_2} \leqslant K_l \quad \text{for } l \leqslant m,$$

for all $\alpha \in \{0,1...,d\}$ and $i,j \in \{1,...,d\}$.

Assumption 2.3. We have $\psi \in L_2(\Omega, \mathscr{F}_0, W_2^m)$. The function f_t is W_2^m-valued, g_t^{ρ}, $\rho = 1,2,...$, are W_2^{m+1}-valued predictable functions given on $\Omega \times [0,T]$. Moreover, for $g_t := (g_t^{\rho})_{\rho=1}^{\infty}$ and

$$\|g_t\|_{W_2^l}^2 := \sum_{\rho=1}^{\infty} \|g_t^{\rho}\|_{W_2^l}^2$$

we have

$$E\int_0^T (\|f_t\|_{W_2^m}^2 + \|g_t\|_{W_2^{m+1}}^2)\,dt + E\|u_0\|_{W_2^m}^2 =: \mathscr{K}_m^2 < \infty.$$

Remark 2.1. If Assumption 2.3 holds with $m > d/2$, then, by the Sobolev embedding of W_2^m into C_b, the space of bounded continuous functions, for almost all ω we can find a continuous function of x which equals to u_0 almost everywhere. Furthermore, for each t and ω we have continuous functions of x which coincide with f_t and g_t, for almost every $x \in \mathbb{R}^d$. Therefore, when Assumption 2.3 holds with $m > d/2$, we always assume that ψ, f_t, and g_t are continuous in x for all t.

We look for the solution of (2.1)-(2.2) in $\mathbb{H}^m(T)$, the Banach space of W_2^m-valued weakly continuous predictable processes $u = (u_t)_{t\in[0,T]}$ with the norm defined by

$$\|u\|_{H_2^m(T)}^2 = E\sup_{t\in[0,T]} \|u(t)\|_{W_2^m}^2 < \infty.$$

We use the notation (φ, ϕ) for the inner product of φ and ϕ in $L_2(\mathbb{R}^d)$.

Definition 2.1. A W_2^1-valued weakly continuous predictable process $u = (u_t)_{t\in[0,T]}$ is a solution to (2.1)-(2.2) if almost surely for all $\varphi \in C_0^{\infty}(\mathbb{R}^d)$

$$(u_t, \varphi) = (u_0, \varphi) + \int_0^t (-a_s^{ij} D_j u_s, D_i\varphi) + (a_s^j D_j u_s, \varphi) + (a_s u_s, \varphi)\,ds + \int_0^t (b_s^{i\rho} D_i u_s + b_s^{\rho}, \varphi)\,dw_s^{\rho}$$

for all $t \in [0,T]$, where $a^j := -D_i a^{ij} + a^{0j} + a^{j0}$ and the summation in the repeated indices i,j is performed over their range $\{1,2...,d\}$.

The following result is known from [10] (cf. also [8]).

Theorem 2.1. *Let Assumptions* 2.2, 2.3, *and* 2.1 *hold. Then* (2.1)-(2.2) *has a unique solution* u. *Moreover,* $u \in \mathbb{H}^m$, *it is a strongly continuous process with values in* W_2^{m-1}, *and there exists a constant* N *depending only on* T, d, m, *and* $K_j, j \leqslant \max(m,2)$, *such that*

$$E\sup_{t\leqslant T} \|u_t\|_{W_2^m}^2 \leqslant N\mathscr{K}_m^2. \tag{2.3}$$

Remark 2.2. We are going to assume that $m > d/2$. Then, by the Sobolev embedding theorems, the solution $u_t(x)$ from Theorem 2.1 is a continuous function of (t, x) (a.s). More precisely, with probability one, for any t one can find a continuous function of x which equals $u_t(x)$ for almost all x and, in addition, this modification is continuous with respect to the couple (t, x).

We are interested in approximating the solution by means of solving a semidiscretized version of (2.1) when partial derivatives are replaced with finite differences. For $\lambda \in \mathbb{R}^d \setminus \{0\}$ and $h \in \mathbb{R} \setminus \{0\}$ define

$$\delta_{h,\lambda} u(x) = \frac{u(x + h\lambda) - u(x)}{h}, \quad \delta_\lambda = \delta_\lambda^h = \frac{1}{2}(\delta_{h,\lambda} + \delta_{-h,\lambda}),$$

and let $\delta_{h,0}$ be the unit operator.

Let $\Lambda \subset \mathbb{R}^d$ be a finite set containing the zero vector and consider the following finite difference equation

$$du_t^h = (L_t^h u_t^h + f_t)\, dt + (M_t^{h,\rho} u_t^h + g_t^\rho)\, dw_t^\rho, \tag{2.4}$$

$$u_0^h = \psi, \tag{2.5}$$

with

$$L_t^h = \mathfrak{a}_t^{\lambda\mu} \delta_\lambda^h \delta_\mu^h + \sum_{\lambda \in \Lambda_0} (\mathfrak{p}^\lambda \delta_{h,\lambda} - \mathfrak{q}^\lambda \delta_{-h,\lambda}), \quad M_t^{h,\rho} = \mathfrak{b}_t^{\lambda\rho} \delta_\lambda^h,$$

where the summation is performed over $\lambda, \mu \in \Lambda$ and in (2.4) also with respect to $\rho \in \{1, 2, \dots\}$. Assume that $\mathfrak{a}^{\lambda\mu} = \mathfrak{a}_t^{\lambda\mu}(x)$, $\mathfrak{p}^\lambda = \mathfrak{p}_t^\lambda(x)$, $\mathfrak{q}^\lambda = \mathfrak{q}_t^\lambda(x)$ are real-valued, and $\mathfrak{b}^\lambda = (\mathfrak{b}_t^{\lambda\rho}(x))_{\rho=1}^\infty$ are l_2-valued, $\mathscr{P} \times \mathscr{B}(\mathbb{R}^d)$-measurable bounded functions on $\Omega \times H_T$, for all $\lambda, \mu \in \Lambda$.

Introduce

$$\mathbb{G}_h = \{\lambda_1 h + ... + \lambda_n h : n = 1, 2, ..., \lambda_i \in \Lambda \cup (-\Lambda)\}.$$

Let $l_2(\mathbb{G}_h)$ be the set of real-valued functions u on $\mathbb{G}_h$ such that

$$|u|^2_{l_2(\mathbb{G}_h)} := |h|^d \sum_{x \in \mathbb{G}_h} |u(x)|^2 < \infty.$$

The notation $l_2(\mathbb{G}_h)$ will also be used for l_2-valued functions like g.

Remark 2.3. Note that Equation (2.4) is just an infinite system of ordinary Itô equations for $\{u_t(x) : x \in \mathbb{G}_h\}$. Therefore, if, for instance, (a.s.)

$$\int_0^T (|f_t|^2_{l_2(\mathbb{G}_h)} + |g_t|^2_{l_2(\mathbb{G}_h)})\, dt < \infty,$$

and Assumption 2.5 (i) holds, then Equation (2.4) has a unique solution with continuous trajectories in $l_2(\mathbb{G}_h)$ provided that the initial data $u_0^h \in l_2(\mathbb{G}_h)$ (a.s.). By the Sobolev embedding of W_2^r into C_b, we have $W_2^r \subset l_2(\mathbb{G}_h)$ if $r > d/2$ (cf. Lemma 4.2 below). Therefore, if

$$\|\psi\|^2_{W_2^r} + \int_0^T \|f(s)\|^2_{W_2^r} + \|g(s)\|^2_{W_2^r}\, ds < \infty \quad (a.s.),$$

then (2.4), (2.5) has a unique $l_2(\mathbb{G}_h)$-valued $\mathscr{F}_t$-adapted continuous solution $(u_t^h)_{t \in [0,T]}$.

It is easy to see that in order u^h approximate the solution of (2.1)-(2.2) the following *consistency condition* is necessary.

Assumption 2.4. For all $i,j=1,...,d$ and $\rho=1,2,...$

$$
\begin{aligned}
&\sum_{\lambda,\mu\in\Lambda_0}\mathfrak{a}_t^{\lambda\mu}\lambda^i\mu^j=a_t^{ij},\quad \sum_{\lambda\in\Lambda_0}\mathfrak{b}_t^{\lambda\rho}\lambda^i=b_t^{i\rho},\\
&\sum_{\lambda\in\Lambda_0}\mathfrak{a}^{\lambda 0}\lambda^i+\sum_{\mu\in\Lambda_0}\mathfrak{a}^{0\mu}\mu^i+\sum_{\mu\in\Lambda_0}\mathfrak{p}^{\lambda}\lambda^i-\sum_{\lambda\in\Lambda_0}\mathfrak{q}^{\lambda}\mu^i=a_t^{i0}+a_t^{0i},\\
&\mathfrak{a}_t^{00}=a_t^{00},\quad \mathfrak{b}_t^{0\rho}=b_t^{0\rho}.
\end{aligned}
\tag{2.6}
$$

There are many ways of constructing appropriate coefficients $\mathfrak{a}$, $\mathfrak{p}$, $\mathfrak{q}$, and $\mathfrak{b}$, satisfying this condition.

Example 2.1. We set $\Lambda=\{e_0,e_1,...,e_d\}$, where $e_0=0$ and e_i is the ith basis vector. Let

$$
\mathfrak{a}_t^{e_\alpha e_\beta}=a_t^{\alpha\beta},\quad \mathfrak{b}_t^{e_\alpha\rho}=b_t^{\alpha\rho},\quad \alpha,\beta=0,1,...,d,\quad \mathfrak{q}^{e_\alpha}=\mathfrak{p}^{e_\alpha}=0\quad \alpha,\beta=1,...,d.
$$

Thus, each derivative D_i in (2.1) is approximated by the *symmetric* finite difference $\delta^h_{e_i}$.

Example 2.2. We take the same set Λ as in the previous example, and define $\mathfrak{p}^{e_\alpha}$, $\mathfrak{q}^{e_\alpha}$ for $\alpha\in\{1,2,...,d\}$ and define

$$
\mathfrak{a}^{00}=a^{00},\quad \mathfrak{a}_t^{e_\alpha e_\beta}=a_t^{\alpha\beta},\quad \alpha,\beta=1,...,d, \mathfrak{b}_t^{e_\alpha\rho}=b_t^{\alpha\rho},\quad \alpha,\beta=0,1,...,d.
$$

We also take nonnegative $\mathscr{P}\otimes\mathscr{B}(\mathbb{R}^d)$-measurable functions $\mathfrak{p}^{e_\alpha}$, $\mathfrak{q}^{e_\alpha}$ for $\alpha\in\{1,...,d\}$, such that

$$
\mathfrak{p}^{e_\alpha}-\mathfrak{q}^{e_\alpha}=a^{0\alpha}+a^{\alpha 0},\quad \alpha\in\{1,2,...,d\}.
$$

To formulate our theorem on the accuracy of the approximation u^h, we fix an integer $l\geqslant 1$, constants A_0,..., A_{l+1} and impose the following condition.

Assumption 2.5. (i) For each (ω,t), $x\in\mathbb{R}^d$

$$
\mathfrak{p}^{\lambda}\geqslant 0,\quad \mathfrak{q}^{\lambda}\geqslant 0,\quad \lambda\in\Lambda_0.
$$

(ii) For some integer $d_1\geqslant 1$ for each $\lambda\in\Lambda_0$ there are $\mathscr{F}\otimes\mathscr{B}(H_T)$-measurable real functions $\sigma^{\lambda 1}$,..., $\sigma^{\lambda d_1}$ on $\Omega\times H_T$ such that for all $(\omega,t,x)\in\Omega\times H_T$

$$
\tilde{\mathfrak{a}}_t^{\lambda\mu}:=2\mathfrak{a}_t^{\lambda\mu}-\mathfrak{b}_t^{\lambda\rho}\mathfrak{b}_t^{\mu\rho}=\sum_{k=1}^{d_1}\sigma^{\lambda k}\sigma^{\mu k},\quad \lambda,\ \mu\in\Lambda_0.
\tag{2.7}
$$

(iii) Let $l\geqslant 1$ be an integer. For $\lambda\in\Lambda_0$ the functions $\sigma^{\lambda k}$, $\mathfrak{b}^{\lambda}$, and $\mathfrak{b}^0$ are $l+1$ times continuously differentiable in x, and $\mathfrak{a}^{0\lambda}$, $\mathfrak{a}^{\lambda 0}$ $\mathfrak{a}^{00}$ $\mathfrak{p}^{\lambda}$, and $\mathfrak{q}^0$ are l times continuously differentiable in x. For all values of arguments

$$
\begin{aligned}
&|D^j\sigma^{\lambda k}|+|D^j\mathfrak{b}^{\lambda}|+|D^j\mathfrak{b}^0|\leqslant A_j\quad \text{for } j\leqslant l+1,\\
&|D^j\mathfrak{a}^{\lambda 0}|+|D^j\mathfrak{a}^{0\lambda}|+|D^j\mathfrak{a}^{00}||D^j\mathfrak{p}^{\lambda}|+|D^j\mathfrak{q}^{\lambda}|\leqslant A_j\quad \text{for } j\leqslant l,
\end{aligned}
$$

for all $\lambda\in\Lambda_0$, $k=1,...,d_1$.

Remark 2.4. It is clear that Assumption 2.5 (ii) implies

$$\sum_{\lambda,\mu\in\Lambda_0} \widetilde{\mathfrak{a}}^{\lambda\mu} z_\lambda z_\mu \geqslant 0 \quad \text{for } (\omega,t,x)\in\Omega\times H_T,\ z_\lambda\in\mathbb{R},\ \lambda\in\Lambda_0,$$

which, together with (2.6), implies Assumption 2.1. If, in addition, Assumptions 2.2 and 2.3 are also satisfied with $m > 2 + d/2$, then (2.1)-(2.2) admits a unique generalized solution, which, by the Sobolev embedding, almost surely equals to a function u for every $t\in[0,T]$ and almost every $x\in\mathbb{R}^d$, such that u and its derivatives in x up to second order are continuous functions on H_T and almost surely

$$du_t(x) = (\mathscr{L}_t u_t(x) + f_t(x))\,dt + (\mathscr{M}_t^\rho u_t(x) + g^\rho(x))\,dw_t^\rho, \quad u_0(x) = \psi(x)$$

for all $t\in[0,T]$ and $x\in\mathbb{R}^d$.

Theorem 2.2. *Let Assumptions* 2.2 *through* 2.5 *hold with* $m\geqslant 3+l$ *and* $l>d/2$. *Then for* $h>0$

$$E\sup_{t\in[0,T]}\sup_{x\in\mathbb{G}_h}|u_t^h(x)-u(t,x)|^2 \leqslant Nh^2\mathscr{K}_m, \tag{2.8}$$

where N *is a constant depending only on* T, Λ, l, d, m, $K_0,\dots,K_m$ *and* $A_0,\dots,A_{l+1}$.

We prove this theorem after the next section. Now we are going to formulate the main result of the paper. Namely, that under additional smoothness conditions, for a given integer $k\geqslant 0$ there exist random fields $u_t^{(j)}(x)$, $(t,x)\in H_T$, such that they are independent of h, $u^{(0)}$ is the solution of (2.1)-(2.2), and for $h\neq 0$ almost surely

$$u_t^h(x) = \sum_{j=0}^k \frac{h^j}{j!}u_t^{(j)}(x) + R_t^h(x) \tag{2.9}$$

for all $t\in[0,T]$ and $x\in\mathbb{G}_h$, where u_t^h is the solution to (2.4), (2.5), and R^h is an $l_2(\mathbb{G}_h)$-valued adapted continuous process such that

$$E\sup_{t\in[0,T]}\sup_{x\in\mathbb{G}_h}|R_t^h(x)|^2 \leqslant Nh^{2(k+1)}\mathscr{K}_m^2 \tag{2.10}$$

with a constant N independent of h.

Assumption 2.6. Let $\mathfrak{m}\geqslant 0$ be a fixed integer. For $\lambda,\mu\in\Lambda$ the derivatives in $x\in\mathbb{R}^d$ of $\mathfrak{a}^{\lambda\mu}$ and the l_2-valued functions $\mathfrak{b}^\lambda$ up to order $\max(\mathfrak{m}-4,0)$, and for $\lambda\in\Lambda_0$ the derivatives in x of $\mathfrak{p}^\lambda$, $\mathfrak{q}^\lambda$ up to order $\max(\mathfrak{m}-2,0)$ are functions, bounded by a constant $C_\mathfrak{m}$, for all $\omega\in\Omega$ and $(t,x)\in H_T$.

Theorem 2.3. *Let Assumptions* 2.2 *through* 2.6 *hold with*

$$m = \mathfrak{m} \geqslant 2k+3+l \tag{2.11}$$

and $l>d/2$, *where* $k\geqslant 0$ *is an integer. Then for* $h>0$ *the expansion* (2.9) *and the estimate* (2.10) *hold with a constant* N *depending only on* d, m, l, T, Λ, $K_0,\dots,K_m$, $A_0,\dots,A_{l+1}$, *and* C_m.

If $\mathfrak{p}^\lambda=\mathfrak{q}^\lambda=0$ *for* $\lambda\in\Lambda_0$, *then* (2.9)-(2.10) *hold for all* $h\neq 0$. *Moreover,* $u^{(j)}=0$ *for odd* $j\leqslant k$, *and if* k *is odd then to have* (2.9) *and* (2.10) *we need only*

$$m > 2k+2+l$$

instead of (2.11).

Remark 2.5. Actually $u^h_t(x)$ is defined for all $x \in \mathbb{R}^d$ rather than only on $\mathbb{G}_h$ and, as we will see from the proof of Theorem 2.3, one can replace $\mathbb{G}_h$ in (2.10) with $\mathbb{R}^d$.

Equality (2.9) clearly yields

$$\delta_{h,\lambda} u^h_t(x) = \sum_{j=0}^{k} \frac{h^j}{j!} \delta_{h,\lambda} u^{(j)}_t(x) + \delta_{h,\lambda} R^h_t(x)$$

for any $\lambda = (\lambda_1, ..., \lambda_n) \in \Lambda^n$ and integer $n \geqslant 0$, where $\Lambda^0 = \{0\}$ and

$$\delta_{h,\lambda} := \delta_{h,\lambda_1} \cdot ... \cdot \delta_{h,\lambda_n}.$$

Theorem 2.3 can be generalized as follows.

Theorem 2.4. *Let $\lambda \in \Lambda^n$ for an integer $n \geqslant 0$. Let Assumptions 2.2 through 2.6 hold with*

$$\mathfrak{m} = m > n + 2k + 3 + l \tag{2.12}$$

and $l > d/2$. Then for $h > 0$ (2.9) and

$$E \sup_{t\in[0,T]} \sup_{x\in\mathbb{G}_h} |\delta_{h,\lambda} R^h_t(x)|^2 \leqslant N h^{2(k+1)} \mathscr{K}^2_m, \tag{2.13}$$

with a constant N depending only on d, m, n, k, l, T, Λ, K_0, ..., K_m, C_m, A_0,...,A_{l+1}.

If $\mathfrak{p}^\lambda = \mathfrak{q}^\lambda = 0$ for $\lambda \in \Lambda_0$, then $u^{(j)} = 0$ for odd $j \leqslant k$, and if k is odd, then, instead of (2.12), we need only

$$m > n + 2k + 2 + l$$

to have (2.9) and the estimate (2.13).

We prove Theorem 2.4 in Section 4 after some preliminaries presented in Section 3.

To accelerate the rate of convergence of u^h, we fix an integer $k \geqslant 0$ and define

$$\overline{u}^h = \sum_{j=0}^{\widetilde{k}} \overline{b}_j u^{2^{-j}h}, \quad \widetilde{u}^h = \sum_{j=0}^{\widetilde{k}} \widetilde{b}_j u^{2^{-j}h}, \tag{2.14}$$

where

$$(\overline{b}_0, \overline{b}_1, ..., \overline{b}_k) := (1,0,0,...,0)\overline{V}^{-1}, \quad (\widetilde{b}_0, \widetilde{b}_1, ..., \widetilde{b}_{\widetilde{k}}) := (1,0,0,...,0)\widetilde{V}^{-1}, \quad \widetilde{k} = [\tfrac{k}{2}],$$

$\overline{V}^{-1}$ is the inverse of the matrix $\overline{V}^{ij} := 2^{-(i-1)(j-1)}$, $i,j = 1, ..., k+1$, and
$\widetilde{V}^{-1}$ is the inverse of matrix $V^{ij} := 4^{-(i-1)(j-1)}$, $i,j = 1, ..., \widetilde{k}+1$.

Theorem 2.5. *Let Assumptions 2.2 through 2.6 hold with*

$$m = \mathfrak{m} \geqslant 2k + 3 + l \tag{2.15}$$

and $l > d/2$, where $k \geqslant 0$ is an integer. Then for $h > 0$

$$E \sup_{t\leqslant T} \sup_{x\in\mathbb{G}_h} |\overline{u}^h_t(x) - u^{(0)}_t(x)|^2 \leqslant N h^{2(k+1)} \mathscr{K}^2_m \tag{2.16}$$

with a constant $N = N(T, m, k, d, l, \Lambda, K_0, ..., K_m, A_0, ..., A_{l+1}, C_m)$.

If $\mathfrak{p}^\lambda = \mathfrak{q}^\lambda = 0$ *for* $\lambda \in \Lambda_0$, *then*

$$E \sup_{t \leqslant T} \sup_{x \in \mathbb{G}_h} |\widetilde{u}_t^h(x) - u_t^{(0)}(x)|^2 \leqslant N|h|^{2(k+1)} \mathscr{K}_m^2 \tag{2.17}$$

for $h \neq 0$, *and if* k *is odd, then to have* (2.17), *we need only*

$$\mathfrak{m} = m > 2k + 2 + l$$

instead of (2.15).

Proof. We prove only (2.17) since the estimate (2.16) can be obtained in the same way. By Theorem 2.3,

$$u^{2^{-j}h} = u^{(0)} + \sum_{i=1}^{\widetilde{k}} \frac{h^{2i}}{(2i)!4^{ji}} u^{(2i)} + h^{k+1} r^{2^{-j}h}, \quad j = 0, 1, ..., k,$$

with $r^{2^{-j}h} := h^{-(k+1)} R^{2^{-j}h}$. Hence with $\widetilde{r}^h := \sum_{j=0}^{\widetilde{k}} r^{2^{-j}h}$

$$\begin{aligned} \widetilde{u}^h = \sum_{j=0}^{\widetilde{k}} \widetilde{b}_j u^{2^{-j}h} &= \left(\sum_{j=0}^{\widetilde{k}} \widetilde{b}_j \right) u^{(0)} + \sum_{j=0}^{\widetilde{k}} \sum_{i=1}^{\widetilde{k}} \widetilde{b}_j \frac{h^{2i}}{(2i)!4^{ij}} u^{(2i)} + h^{k+1} \widetilde{r}^h \\ &= u^{(0)} + \sum_{i=1}^{\widetilde{k}} \frac{h^{2i}}{(2i)!} u^{(2i)} \sum_{j=0}^{\widetilde{k}} \frac{\widetilde{b}_j}{4^{ij}} + h^{k+1} \widetilde{r}^h = u^{(0)} + h^{k+1} \widetilde{r}^h \end{aligned}$$

since

$$\sum_{j=0}^{\widetilde{k}} \widetilde{b}_j = 1, \quad \sum_{j=0}^{\widetilde{k}} \widetilde{b}_j 4^{-ij} = 0, \quad i = 1, 2, ...\widetilde{k},$$

by the definition of $(\widetilde{b}_0, ..., \widetilde{b}_{\widetilde{k}})$. Thus, owing to (2.10), we have

$$E \sup_{t \in [0,T]} \sup_{x \in \mathbb{G}_h} |\widetilde{u}^h - u|^2 \leqslant h^{2(k+1)} E \sup_{t \in [0,T]} \sup_{x \in \mathbb{G}_h} |\widetilde{r}_t^h(x)|^2 \leqslant N h^{2(k+1)} \mathscr{K}_m^2$$

and the theorem is proved. □

Remark 2.6. Note that without acceleration, i.e., when $k = 0$ and $k = 1$ in (2.15) and (2.16) respectively, in the above theorem for $h > 0$ we have

$$E \sup_{t \in [0,T]} \sup_{x \in \mathbb{G}_h} |u^h - u^0|^2 \leqslant N h^2 \mathscr{K}_m^2,$$

and when $\mathfrak{p}^\lambda = \mathfrak{q}^\lambda = 0$ for $\lambda \in \Lambda_0$ we have

$$E \sup_{t \in [0,T]} \sup_{x \in \mathbb{G}_h} |u^h - u^0|^2 \leqslant N h^4 \mathscr{K}_m^2$$

respectively. These are sharp estimates by virtue of Remark 2.21 in [2] on finite difference approximations for deterministic parabolic partial differential equations.

Remark 2.7. Let $\mathfrak{p}^\lambda = \mathfrak{q}^\lambda = 0$ for $\lambda \in \Lambda_0$. Let $n \geqslant 0$. Suppose that the assumptions of Theorem 2.3 hold with

$$m > n + 2k + 3 + d/2,$$

with an integer $n \geqslant 0$. Then for $\lambda \in \Lambda^n$ (2.17) holds with $\delta_{h,\lambda}\widetilde{u}^h$ and $\delta_{h,\lambda}u^{(0)}$ in place of $\widetilde{u}^h$ and $u^{(0)}$, respectively, with a constant N depending on T, m, k, n, d, b, Λ, K_0,...,K_m, A_0,...,A_{l+1} and C_m.

Proof. This assertion follows from Theorem 2.4 in the same way as Theorem 2.5 follows from Theorem 2.3. □

By the above remark, one can construct fast approximations for the derivatives of $u^{(0)}$ via suitable linear combinations of finite differences of $\widetilde{u}^h$.

Example 2.3. Assume that $d = 2$, $m = 10$, and $\mathfrak{p}^\lambda = \mathfrak{q}^\lambda = 0$ in $\lambda \in \Lambda_0$. Then

$$\widetilde{u}^h := \tfrac{4}{3}u^{h/2} - \tfrac{1}{3}u^h$$

satisfies

$$E \sup_{t\leqslant T} \sup_{x\in\mathbb{G}_h} |u_t^{(0)}(x) - \widetilde{u}_t^h(x)|^2 \leqslant Nh^8.$$

Example 2.4. Consider the SPDE

$$du_t = aD^2u_t\,dt + bDu_t\,dw_t \quad t > 0,\, x \in \mathbb{R}$$

with initial data $u_0(x) = \cos x$, $x \in \mathbb{R}$, coefficients $a = b = 2$, and a one-dimensional Wiener process w. Note that $2a - b^2/2 = 0$, i.e., this is a degenerate parabolic SPDE. The unique bounded solution is

$$u_t(x) = \cos(x + 2w_t).$$

The finite difference equation (2.4) is the following:

$$du_t^h(x) = \frac{u_t^h(x+2h) - 2u_t^h(x) + u_t^h(x-2h)}{2h^2}\,dt + \frac{u_t^h(x+h) - u_t^h(x-h)}{h}\,dw_t.$$

Its unique bounded solution with initial condition $u_0(x) = \cos x$ is

$$u_t^h(x) = \cos(x + 2\phi_h w_t),$$

where $\phi_h = \sin h/h$. For $t = 1$, $h = 0.1$, and $w_t = 1$ we have

$$u_1(0) \approx -0.4161468365, \quad u_1^h(0) \approx -0.4131150562, \quad u_1^{h/2}(0) \approx -0.415389039,$$
$$\widetilde{u}_1^h(0) = \tfrac{4}{3}u_1^{h/2}(0) - \tfrac{1}{3}u_1^h(0) \approx 0.4161470333.$$

Such a level of accuracy by $u_1^{\widetilde{h}}(0)$ is achieved with $\widetilde{h} = 0.0008$, which is more than 60 times smaller than $h/2$.

Note that this example does not quite fit into our scheme because u_0 is not square summable over $\mathbb{R}$, but one can extend our approach to weighted Sobolev spaces and then the above example can be included formally.

3 Auxiliary Facts

Recall the notation

$$\delta_{h,\lambda} = \frac{1}{h}(T_{h,\lambda} - I), \quad \delta_\lambda = \delta_\lambda^h = \frac{1}{2}(\delta_{h,\lambda} + \delta_{-h,\lambda}) = \frac{1}{2h}(T_{h,\lambda} - T_{h,-\lambda}),$$

for $h \neq 0$, $\lambda \in \mathbb{R}^d$, where for all $h \in \mathbb{R}$

$$T_{h,\lambda}\varphi(x) = \varphi(x + h\lambda), \quad x \in \mathbb{R}^d$$

for functions φ on $\mathbb{R}^d$. We set

$$I_\lambda = I_\lambda^h = \frac{1}{2}(T_{h,\lambda} + T_{h,-\lambda}),$$

$$\Delta_\lambda = \Delta_\lambda^h = \frac{1}{h}(\delta_{h,\lambda} + \delta_{h,-\lambda}) = \delta_{h,\lambda}\delta_{-h,\lambda} = (\delta_\lambda^{h/2})^2.$$

The following useful identities can easily be verified.

Lemma 3.1.

$$\delta_{h,\lambda}(uv) = (\delta_{h,\lambda}u)v + (\delta_{h,\lambda}v)T_{h,\lambda}u = (\delta_{h,\lambda}u)v + (\delta_{h,\lambda}v)u + h(\delta_{h,\lambda}u)(\delta_{h,\lambda}v), \tag{3.1}$$

$$\begin{aligned} \delta_\lambda(uv) &= (\delta_\lambda u)v + \frac{1}{2}\{(\delta_{h,\lambda}v)T_{h,\lambda}u + (\delta_{-h,\lambda}v)T_{h,-\lambda}u\} \\ &= (\delta_\lambda u)v + (\delta_\lambda v)I_\lambda u + \frac{h^2}{2}(\Delta_\lambda v)\delta_\lambda u \\ &= (\delta_\lambda u)v + (\delta_\lambda v)u + \frac{h^2}{2}\{(\delta_\lambda u)\Delta_\lambda v + (\Delta_\lambda u)\delta_\lambda v\} \end{aligned} \tag{3.2}$$

For linear operators A and B we use the notation

$$[A, B] = BA - AB.$$

Lemma 3.2.

$$\delta_\mu(a\delta_\lambda) = a\delta_\mu\delta_\lambda + \frac{1}{2}(\delta_\mu a)(\delta_{\lambda+\mu} + \delta_{\lambda-\mu}) + \frac{h^2}{2}(\Delta_\mu a)\delta_\lambda\delta_\mu, \tag{3.3}$$

$$\begin{aligned} [a\delta_\lambda, b\delta_\mu] = \frac{1}{2}(b(\delta_\mu a) - a(\delta_\lambda b))\delta_{\lambda+\mu} + \frac{1}{2}(b(\delta_\mu a) + a(\delta_\lambda b))\delta_{\lambda-\mu} \\ + \frac{h^2}{2}(b(\Delta_\mu a) - a(\Delta_\lambda b)\delta_\lambda\delta_\mu, \end{aligned} \tag{3.4}$$

$$[a\delta_{h,\mu}, bT_{h,\lambda}] = (b(\delta_{h,\lambda}a) - a(\delta_{h,\mu}b))T_{h,\lambda+\mu} - b(\delta_{h,\lambda}a)T_{h,\lambda}, \tag{3.5}$$

$$[a\delta_\mu, bT_{h,\lambda}] = \frac{1}{2}(b(\delta_{h,\lambda}a) - a(\delta_{h,\mu}b))(T_{h,\lambda+\mu} - T_{h,\lambda-\mu}) - a(\delta_{h,\mu}b)T_{h,\lambda-\mu} \tag{3.6}$$

Let $l \geqslant 0$ be an integer and $K \geqslant 0$ be a constant. In the next lemma $\mathscr{M}$ and $\mathscr{N}$ denote difference operators of the form $\mathscr{M} = \sum_{\lambda\in\Lambda_0} b^\lambda\delta_{h,\lambda}$ and $\mathscr{N} = \sum_{\lambda\in\Lambda_0} b^\lambda\delta_\lambda$, with functions b_λ on $\mathbb{R}^d$, and $(,)$ denotes the inner product in $L_2(\mathbb{R}^d)$.

Lemma 3.3. *The following estimates hold for all multiindices α, $|\alpha| \leqslant l$, and functions $\varphi \in W_2^l$ on $\mathbb{R}^d$.*

(i) *If $b^\lambda \geqslant 0$ for $\lambda \in \Lambda_0$, and they, together with their derivatives up to order $l \vee 1$ are functions, bounded by K, then for $h > 0$*

$$(D^\alpha \mathscr{M}\varphi, D^\alpha\varphi) \leqslant N\|\varphi\|^2_{W^l_2}. \tag{3.7}$$

(ii) *If for each $\lambda \in \Lambda_0$, b^λ and its derivatives up to order $l \vee 1$ are functions, bounded by K, then for $h \neq 0$*

$$|(D^\alpha \mathscr{N}\varphi, D^\alpha\varphi)| \leqslant N\|\varphi\|^2_{W^2_l}. \tag{3.8}$$

(iii) *If for $\lambda \in \Lambda_0$ the coefficients b^λ and its derivatives up to order $(l+1) \vee 2$ are functions on $\mathbb{R}^d$, bounded by K, and b^0 and its derivatives up to order $l+1$ are functions, bounded by K, then for $h \neq 0$*

$$|(D^\alpha \mathscr{N}\mathscr{N}\varphi, D^\alpha\varphi) + (D^\alpha \mathscr{N}\varphi, D^\alpha\mathscr{N}\varphi)| \leqslant N\|\varphi\|^2_{W^2_l}. \tag{3.9}$$

In these estimates N denotes a constant that depends only on Λ_0, d, l, and K.

Proof. To prove (i), we note that, by (3.1),

$$\sum_{\lambda\in\Lambda_0} \varphi b^\lambda \delta_{h,\lambda}\varphi = \frac{1}{2}\sum_{\lambda\in\Lambda_0} b^\lambda \delta_{h,\lambda}(\varphi^2) - \frac{h}{2}\sum_{\lambda\in\Lambda_0} b^\lambda(\delta_{h,\lambda}\varphi)^2 \leqslant \frac{1}{2}\sum_{\lambda\in\Lambda_0} b^\lambda\delta_{h,\lambda}(\varphi^2).$$

Hence, taking into account that $\delta^*_{h,\lambda}$, the adjoint of $\delta_{h,\lambda}$ in L_2, is $\delta_{h,-\lambda}$, we have

$$(\mathscr{M}\varphi, \varphi) \leqslant \frac{1}{2}\sum_{\lambda\in\Lambda_0}(\delta_{h,-\lambda}b^\lambda, \varphi^2) \tag{3.10}$$

which yields (3.7) for $l = 0$. For $|\alpha| = l \geqslant 1$

$$\sum_{1\leqslant|\gamma|,\gamma+\beta=\alpha}\ \sum_{\lambda\in\Lambda_0} |((D^\gamma b^\lambda)\delta_{h,\lambda}D^\beta\varphi, D^\alpha\varphi)| \leqslant N\|\varphi\|^2_{W^l_2}.$$

Hence

$$(D^\alpha\mathscr{M}\varphi, D^\alpha\varphi) \leqslant N\|\varphi\|^2_{W^l_2} + (\mathscr{M}D^\alpha\varphi, D^\alpha\varphi),$$

which yields (3.7) since, by (3.10),

$$(\mathscr{M}D^\alpha\varphi, D^\alpha\varphi) \leqslant N\|\varphi\|^2_{W^l_2}.$$

To prove (ii), we note that

$$(\mathscr{N}\varphi, \varphi) = (T\varphi, \varphi) \tag{3.11}$$

with

$$T = \frac{1}{2}(\mathscr{N} + \mathscr{N}^*) = -\frac{1}{4}\sum_{\lambda\in\Lambda_0}((\delta_{h,\lambda}b_\lambda)T_{h,\lambda} + (\delta_{-h,\lambda}b_\lambda)T_{h,-\lambda}), \tag{3.12}$$

where $\mathscr{N}^*$ denotes the adjoint of $\mathscr{N}$ in L_2. Hence

$$|(\mathscr{N}\varphi, \varphi)| \leqslant \frac{K}{4}\sum_{\lambda\in\Lambda_0}|\lambda|(\|T_{h,\lambda}\varphi\|_{L_2} + \|T_{h,-\lambda}\varphi\|_{L_2})\|\varphi\|_{L_2} = \frac{K}{2}\sum_{\lambda\in\Lambda_0}|\lambda|\|\varphi\|^2_{L_2}, \tag{3.13}$$

which proves (3.8) for $\alpha = 0$. For $|\alpha| = l \geqslant 1$

$$\sum_{1\leqslant|\gamma|, \gamma+\beta=\alpha} \sum_{\lambda\in\Lambda_0} |((D^\gamma b^\lambda)\delta_{h,\lambda} D^\beta \varphi, D^\alpha \varphi)| \leqslant N\|\varphi\|^2_{W^l_2}.$$

Hence

$$(D^\alpha \mathscr{N}\varphi, D^\alpha\varphi) \leqslant N\|\varphi\|^2_{W^l_2} + (\mathscr{N} D^\alpha\varphi, D^\alpha\varphi),$$

which implies (3.8) since due to (3.10) we have

$$(\mathscr{N} D^\alpha\varphi, D^\alpha\varphi) \leqslant N\|\varphi\|^2_{W^l_2}.$$

Now we prove (iii). From (3.11) by polarization we get

$$(\mathscr{N}\psi, \phi) + (\mathscr{N}\varphi, \psi) = 2(T\varphi, \psi)$$

for functions $\varphi, \psi \in L_2$. Substituting here $\mathscr{N}\varphi$ in place of ψ, using $T^* = T$ and $\mathscr{N}^* = 2T - \mathscr{N}$, we obtain

$$\begin{aligned}(\mathscr{N}\mathscr{N}\varphi, \varphi) + (\mathscr{N}\varphi, \mathscr{N}\varphi) &= 2(T\varphi, \mathscr{N}\varphi) = ((T\mathscr{N} + \mathscr{N}^* T)\varphi, \varphi)\\ &= ((T\mathscr{N} - \mathscr{N}T + 2T^2)\varphi, \varphi) = ([\mathscr{N}, T]\varphi, \varphi) + 2(T\varphi, T\varphi).\end{aligned}$$

Hence, using (3.12) and the identity (3.5), we easily get (3.9) for $\alpha = 0$. To deal with the case $\alpha \neq 0$, we fix $\varphi \in W^l_2$ and use the notation $f \sim g$ for functions $f, g \in L_1$, which may depend also on the parameter h if

$$\left|\int_{\mathbb{R}^d} (f(x) - g(x))\, dx\right| \leqslant N|\varphi|^2_{W^2_l}$$

with a constant N depending only on Λ, l, d and K. It is clear that

$$(D^\alpha \mathscr{N}\varphi) D^\alpha\varphi \sim (\mathscr{N} D^\alpha\varphi) D^\alpha\varphi.$$

For multiindices γ, $|\gamma| \leqslant m$, we set

$$\mathscr{N}^{(\gamma)} = \sum_{\lambda\in\Lambda_0} (D^\gamma b_\lambda)\delta_\lambda,$$

and note that for multiindices $\beta \neq 0$, $\gamma \neq 0$, ρ, such that $\beta + \gamma + \rho = \alpha$ we have

$$(\mathscr{N}^{(\beta)} \mathscr{N}^{(\gamma)} D^\rho\varphi) D^\alpha\varphi \sim 0.$$

Similarly, for multiindices $\beta \neq 0$, $\gamma \neq 0$, $\overline{\beta}$ and $\overline{\gamma}$ such that $\beta + \overline{\beta} = \alpha$ and $\gamma + \overline{\gamma} = \alpha$ we have

$$(\mathscr{N}^{(\beta)} D^{\overline{\beta}}\varphi) \mathscr{N}^{(\gamma)} D^{\overline{\gamma}}\varphi \sim 0,$$

and if $\beta = 0$ and $0 < \gamma < \alpha$ we have

$$(\mathscr{N} D^\alpha\varphi) \mathscr{N}^{(\gamma)} D^{\overline{\gamma}}\varphi \sim (D^\alpha\varphi) \mathscr{N}^* \mathscr{N}^{(\gamma)} D^{\overline{\gamma}}\varphi \sim 0,$$

owing to

$$\mathscr{N}^* = -\mathscr{N} + \frac{1}{2} \sum_{\lambda\in\Lambda_0} \{(\delta_{h,\lambda} c_\lambda) T_{h,\lambda} + (\delta_{-h,\lambda} c_\lambda) T_{h,-\lambda}\}. \tag{3.14}$$

Thus, for

$$J := (D^\alpha \mathcal{N}\mathcal{N}\varphi)D^\alpha\varphi + (D^\alpha\mathcal{N}\varphi)D^\alpha\mathcal{N}\varphi$$

we get

$$J \sim J_{00} + \sum_{0<\gamma\leqslant\alpha} J_{0\gamma} + \sum_{0<\beta\leqslant\alpha} J_{\beta 0}, \tag{3.15}$$

with

$$J_{\beta\gamma} := (\mathcal{N}^{(\beta)}\mathcal{N}^{(\gamma)}D^\rho\varphi)D^\alpha\varphi + (\mathcal{N}^{(\beta)}D^{(\overline{\beta})}\varphi)\mathcal{N}^{(\gamma)}D^{\overline{\gamma}}\varphi,$$

where $\beta+\gamma+\rho=\alpha$, $\beta+\overline{\beta}=\alpha$, and $\gamma+\overline{\gamma}=\alpha$. By (3.9) with $\alpha=0$, we have

$$J_{00} \sim 0. \tag{3.16}$$

Using (3.14), we find

$$(\mathcal{N}\mathcal{N}^{(\gamma)}D^{\overline{\gamma}}\varphi)D^\alpha\varphi \sim (\mathcal{N}^{(\gamma)}D^{\overline{\gamma}}\varphi)\mathcal{N}^*D^\alpha\varphi \sim -(\mathcal{N}^{(\gamma)}D^{\overline{\gamma}}\varphi)\mathcal{N}D^\alpha\varphi,$$

which for $\gamma\leqslant\alpha$, $|\gamma|=1$ yields

$$J_{0\gamma} = (\mathcal{N}\mathcal{N}^{(\gamma)}D^{\overline{\gamma}}\varphi)D^\alpha\varphi + (\mathcal{N}D^\alpha\varphi)\mathcal{N}^{(\gamma)}D^{\overline{\gamma}}\varphi \sim 0. \tag{3.17}$$

For $\beta\leqslant\alpha$, $|\beta|=1$ the identity (3.4) yields

$$(\mathcal{N}^{(\beta)}\mathcal{N}D^{\overline{\beta}}\varphi)D^\alpha\varphi \sim (\mathcal{N}\mathcal{N}^{(\beta)}D^{\overline{\beta}}\varphi)D^\alpha\varphi,$$

which for $\beta<\alpha$, $|\beta|=1$ implies

$$J_{\beta 0} \sim J_{0\beta} \sim 0. \tag{3.18}$$

For $\gamma<\alpha$, $|\gamma|\geqslant 2$ it is easy to see that

$$J_{0\gamma} \sim 0, \tag{3.19}$$

and similarly, for $\beta<\alpha$, $|\beta|\geqslant 2$ it is easy to see that

$$J_{\beta 0} \sim 0. \tag{3.20}$$

Using (3.14) we get

$$(\mathcal{N}\mathcal{N}^{(\alpha)}\varphi)D^\alpha\varphi \sim -(\mathcal{N}^{(\alpha)}\varphi)\mathcal{N}D^\alpha\varphi,$$

that yields

$$J_{0\alpha} \sim 0. \tag{3.21}$$

It is clear that

$$(\mathcal{N}^{(\alpha)}\mathcal{N}\varphi)D^\alpha\varphi \sim 0,$$

and

$$(\mathcal{N}^{(\alpha)}\varphi)\mathcal{N}D^\alpha\varphi \sim (\mathcal{N}^*\mathcal{N}^{(\alpha)}\varphi)D^\alpha\varphi \sim -(\mathcal{N}\mathcal{N}^{(\alpha)}\varphi)D^\alpha\varphi \sim 0.$$

(This is the only place where we need that the coefficients of $\mathcal{N}$ have bounded derivatives up to $l+1$, not only up to $l\vee 2$ as in the rest of the proof.) Hence $J_{\alpha 0}\sim 0$, that together with (3.15)–(3.16) and (3.17)–(3.21) implies $J\sim 0$, which proves (3.9). ☐

Remark 3.1. Let $(\mathscr{N}^\rho)_{\rho=1}^\infty$ be a sequence of operators of the form $\mathscr{N}^\rho = \sum_{\lambda\in\Lambda_0} \mathfrak{b}^{\lambda\rho}\delta_\lambda$, where $\mathfrak{b}^\lambda = (\mathfrak{b}^{\lambda\rho})_{\rho=1}^\infty$ is an l_2-valued Borel function on $\mathbb{R}^d$ for each $\lambda \in \Lambda_0$. Let $l \geqslant 0$ be an integer. Then the following statements hold for all multiindices α, $|\alpha| \leqslant l$ and functions $\varphi \in W_2^l$.

(i) If $\mathfrak{b}^\lambda$ and their derivatives up to order $\max(l,1)$ are functions, bounded by K for all $\lambda \in \Lambda_0$, then

$$\sum_{\rho=1}^\infty |(D^\alpha \mathscr{N}^\rho\varphi, D^\alpha\varphi)|^2 \leqslant N\|\varphi\|_{W_2^l}^4.$$

(ii) If $\mathfrak{b}^\lambda$ and their derivatives up to order $(l+1)\vee 2$ are l_2-valued functions bounded by K for all $\lambda \in \Lambda_0$, then

$$|(D^\alpha \sum_{\rho=1}^\infty \mathscr{N}^\rho\mathscr{N}^\rho\varphi, D^\alpha\varphi) + \sum_{\rho=1}^\infty (D^\alpha\mathscr{N}^\rho\varphi, D^\alpha\mathscr{N}^\rho\varphi)| \leqslant N\|\varphi\|_{W_2^l}^2.$$

In these estimates N is a constant depending only on K, l and d.

Proof. Taking into account that $\sum_\rho |D^\alpha b^{\lambda\rho}|^2 \leqslant K$ for $|\alpha| \leqslant \max(l,1)$ and for $|\alpha| \leqslant \max(l+1,2)$ respectively, we can get these estimates in the same way as the estimates (3.8) and (3.9) are obtained. □

Lemma 3.4. *Let Assumption* 2.5 *hold. Then for multiindices* α, $|\alpha| \leqslant l$, *we have*

$$\mathbb{Q}_t^\alpha(\varphi) := \int_{\mathbb{R}^d} 2D^\alpha\varphi(x)D^\alpha L^h\varphi(x) + \sum_{\rho=1}^\infty |D^\alpha M^{h\rho}\varphi(x)|^2\,dx \leqslant N\|\varphi\|_{W_2^l}^2,$$

where N *depends only on* l, K, d, *and* Λ.

Proof. We set $\mathscr{M}^{h\rho} = \sum_{\lambda\in\Lambda_0} b^{\lambda\rho}\delta_\lambda$. Then $M^{h\rho} = \mathscr{M}^{h\rho} + b^{0\rho}$, and, by Remark 3.1,

$$\sum_\rho (D^\alpha\mathscr{M}^{h\rho}\varphi, D^\alpha\mathscr{M}^{h\rho}\varphi) \leqslant -\Big(D^\alpha\sum_\rho \mathscr{M}^{h\rho}\mathscr{M}^{h\rho}, D^\alpha\varphi\Big) + N\|\varphi\|_{W_2^l}^2,$$

where $(,)$ denotes the inner product in L_2 and N is a constant depending only on K, l d, and Λ. By the equality (3.3),

$$\mathscr{M}^{h\rho}\mathscr{M}^{h\rho} = \sum_{\lambda,\mu\in\Lambda_0} \mathfrak{b}^{\lambda\rho}\mathfrak{b}^{\mu\rho}\delta_\lambda\delta_\mu + \widetilde{\mathscr{M}}$$

with

$$\begin{aligned}
\widetilde{\mathscr{M}} = {} & \frac{1}{2}\sum_{\lambda,\mu\in\Lambda_0} \mathfrak{b}^{\mu\rho}(\delta_\mu\mathfrak{b}^{\lambda\rho})(\delta_{\lambda+\mu} + \delta_{\lambda-\mu}) \\
& + \frac{1}{8}\sum_{\lambda,\mu\in\Lambda_0} ((\delta_{h,\mu} + \delta_{h,-\mu})\mathfrak{b}^{\lambda\rho})(T_{h,\lambda+\mu} - T_{h,\mu-\lambda} - T_{h,\lambda-\mu} + T_{h,-\lambda-\mu}) \\
& + \mathfrak{b}^{0\rho}\mathfrak{b}^{0\rho} + \sum_{\lambda\in\Lambda_0} \mathfrak{b}^{0\rho}\mathfrak{b}^{\lambda\rho}\delta_\lambda + \sum_{\mu\in\Lambda_0} \mathfrak{b}^{\mu\rho}\mathfrak{b}^{0\rho}\delta_\mu \\
& + \frac{1}{2}\sum_{\mu\in\Lambda_0} \{\mathfrak{b}^{\mu\rho}(\delta_{h,\mu}\mathfrak{b}^{0\rho})T_{h,\mu} + \mathfrak{b}^{\mu\rho}(\delta_{h,-\mu}\mathfrak{b}^{0\rho})T_{h,-\mu}\},
\end{aligned}$$

where the summation convention with respect to the repeated index ρ is used. By Lemma (3.3) (i) and (ii), for $h>0$ we have

$$(D^\alpha \sum_{\lambda\in\Lambda_0} \mathfrak{p}^\lambda \delta_{h,\lambda}\varphi, D^\alpha\varphi) \leqslant \|\varphi\|^2_{W^l_2},$$
$$(D^\alpha \sum_{\lambda\in\Lambda_0} \mathfrak{q}^\lambda \delta_{h,-\lambda}\varphi, D^\alpha\varphi) \leqslant \|\varphi\|^2_{W^l_2}$$

and

$$|(D^\alpha \widetilde{\mathscr{M}}\varphi, D^\alpha\varphi)| \leqslant N\|\varphi\|_{W^l_2},$$
$$|(D^\alpha\varphi, D^\alpha(\mathfrak{a}^{0\lambda}+\mathfrak{a}^{\lambda 0})\delta_\lambda\varphi)| \leqslant N\|\varphi\|_{W^l_2},$$
$$(D^\alpha b^{0\rho}\varphi, D^\alpha \mathscr{M}^{h\rho}\varphi) \leqslant N\|\varphi\|_{W^l_2}$$

for $h\neq 0$. Here and everywhere in this proof, N stands for constants depending only on l, K, d, and Λ. Hence

$$\mathbb{Q}^\alpha(\varphi) \leqslant (D^\alpha\varphi, D^\alpha \sum_{\lambda,\mu\in\Lambda_0} \widetilde{\mathfrak{a}}^{\lambda\mu}\delta_\lambda\delta_\mu\varphi) + N\|\varphi\|^2_{W^l_2}. \tag{3.22}$$

Owing to Assumption 2.5 (ii) and the equality (3.3), we have

$$\sum_{\lambda,\mu\in\Lambda_0} \widetilde{\mathfrak{a}}^{\lambda\mu}\delta_\lambda\delta_\mu = \sum_{r=1}^{d_1} \mathscr{N}^r\mathscr{N}^r - \widetilde{\mathscr{N}}$$

with

$$\mathscr{N}^r = \sum_{\lambda\in\Lambda_0} \sigma^{\lambda r}\delta_\lambda, \quad \widetilde{\mathscr{N}} = \sum_{r=1}^{d_1} \widetilde{\mathscr{N}^r},$$

where for each $r=1,...,d_1$

$$\widetilde{\mathscr{N}^r} = \frac{1}{2}\sum_{\lambda,\mu\in\Lambda_0} \sigma^{\lambda r}(\delta_\lambda\sigma^{\mu r})\{\delta_{\lambda+\mu}+\delta_{\mu-\lambda}\}$$
$$+\frac{1}{8}\sum_{\lambda,\mu\in\Lambda_0} ((\delta_{h,\lambda}+\delta_{h,-\lambda})\sigma^{\mu r})(T_{h,\lambda+\mu}-T_{h,-\mu-\lambda}-T_{h,\lambda-\mu}+T_{h,-\lambda-\mu}).$$

By Lemma 3.3 (ii) and (iii) for $h\neq 0$

$$|(D^\alpha\widetilde{\mathscr{N}}\varphi, D^\alpha\varphi)| \leqslant N\|\varphi\|^2_{W^l_2},$$
$$(D^\alpha\sum_r \mathscr{N}^r\mathscr{N}^r, D^\alpha\varphi) = \sum_r (D^\alpha\mathscr{N}^r\mathscr{N}^r, D^\alpha\varphi) \leqslant N\|\varphi\|^2_{W^l_2}.$$

Hence

$$(D^\alpha\varphi, D^\alpha \sum_{\lambda,\mu\in\Lambda_0} \widetilde{\mathfrak{a}}^{\lambda\mu}\delta_\lambda\delta_\mu\varphi) \leqslant N\|\varphi\|^2_{W^l_2},$$

which along (3.22) completes the proof. □

We consider the finite difference scheme (2.4), (2.5) now on $[0,T]\times\mathbb{R}^d$ rather than on $[0,T]\times\mathbb{G}_h$. We use the notation $\mathbb{W}_2^m(T)$ and $\mathbb{W}_2^m(T,l_2)$ for the Banach spaces of W_2^m-valued predictable processes $(f_t)_{t\in[0,T]}$ and sequences of W_2^m valued processes $g_t=(g_t^\rho)_{t\in[0,T]}$, $\rho=1,2,...$, respectively, with the norms defined by

$$\|f\|^2_{\mathbb{W}_2^m(T)}=\int\limits_0^T\|f(t)\|^2_{W_2^m}\,dt,$$

$$\|g\|^2_{\mathbb{W}_2^m(T,l_2)}=\int\limits_0^T\sum_{\rho=1}^\infty\|g^\rho(t)\|^2_{W_2^m}\,dt.$$

For the sake of brevity, we write also $\mathbb{W}_2^m(T)$ in place of $\mathbb{W}_2^m(T,l_2)$.

Theorem 3.5. *Let Assumption 2.5 hold. Let ψ be a W_2^l-valued $\mathscr{F}_0$-measurable random variable, $f\in\mathbb{W}_2^l(T)$ and $g\in\mathbb{W}_2^{l+1}(T,l_2)$. Then for each $h\neq0$ there exists a unique continuous L_2-valued solution $u^h=(u_t^h)_{t\in[0,T]}$ to (2.4), (2.5). Moreover, u^h is a W_2^l-valued continuous process, and for $h>0$*

$$E\sup_{t\leqslant T}\|u_t^h\|^2_{W_2^l}\leqslant NE\|u_0\|^2_{W_2^l}+NE\int\limits_0^T\big(\|f_t\|^2_{W_2^l}+\|g_t\|^2_{W_2^{l+1}}\big)\,dt, \tag{3.23}$$

with a constant N depending only on d, l, Λ, $A_0,\dots,A_{l+1}$, and T. If $\mathfrak{p}^\lambda=\mathfrak{q}^\lambda=0$ for $\lambda\in\Lambda_0$, then this estimate holds for all $h\neq0$.

Proof. Since (2.4) is an ordinary Itô equation with Lipschitz continuous coefficients for L_2-valued processes, it has a unique L_2-valued continuous solution u^h for each $h\neq0$. Similarly, it has a unique W_2^l-valued continuous solution and, since $W_2^l\subset L_2$, it follows that u^h is actually a continuous W_2^l-valued adapted process. One can easily get the estimate (3.23) with a constant N which depends on h. In particular, the solution belongs to $\mathbb{W}_2^{0,l}(T)$. We assume that $E\|u\|^2_{W_2^l}<\infty$; otherwise, (3.23) is trivial. To prove (3.23) with a constant N independent of h, we take any multiindex α, $|\alpha|\leqslant l$, and use the Itô formula for the L_2-valued process $D^\alpha u^h$ to find

$$\begin{aligned}d\|D^\alpha u_t^h\|^2_{L_2}&=\{\mathbb{Q}_t^\alpha(u_t^h)+2(D^\alpha u_t^h,D^\alpha f_t)+2(D^\alpha b^{\lambda\rho}\delta_\lambda u_t^h,D^\alpha g_t^\rho)+\sum_\rho\|D^\alpha g_t^\rho\|^2_{L_2}\}\,dt\\&\quad+2(D^\alpha u_t^h,D^\alpha M^{h,\rho}u_t^h+D^\alpha g_t^\rho)\,dw_t^\rho,\end{aligned} \tag{3.24}$$

where $\mathbb{Q}^\alpha$ is defined in Lemma 3.4. It is clear that

$$2|(D^\alpha u_t^h,D^\alpha f_t)|\leqslant\|u_t\|^2_{W_2^l}+\|f_t\|^2_{W_2^l}$$

and, by integration by parts,

$$2|(D^\alpha b^{\lambda\rho}\delta_\lambda u_t^h,D^\alpha g_t^\rho)|\leqslant N\|u_t\|_{W_2^l}\|g\|_{W_2^{l+1}}\leqslant N(\|u_t\|^2_{W_2^l}+\|g\|^2_{W_2^{l+1}}).$$

Thus, using Lemma 3.4, from (3.24) we have

$$\begin{aligned}d\sum_{|\alpha|\leqslant l}\|D^\alpha u_t^h\|^2_{L_2}&\leqslant N\big(\|u_t^h\|^2_{W_2^l}+\|f_t\|^2_{W_2^l}+\|g_t\|^2_{W_2^l}\big)\,dt\\&\quad+2\sum_{|\alpha|\leqslant l}(D^\alpha u_t^h,D^\alpha M^{h,\rho}u_t^h+D^\alpha g_t^\rho)\,dw_t^\rho\end{aligned} \tag{3.25}$$

for $h>0$ and if $\mathfrak{p}^\lambda=\mathfrak{q}^\lambda=0$ for $\lambda\in\Lambda_0$, then it holds for all $h\neq 0$. Hence

$$E\|u_t^h\|^2_{W_2^l}\leqslant E\|u_0\|^2_{W_2^l}+NE\int_0^t\big(\|u_s^h\|^2_{W_2^l}+\|f_s\|^2_{W_2^l}+\|g_s\|^2_{W_2^{l+1}}\big)\,ds<\infty, \tag{3.26}$$

which, by the Gronwall lemma, yields

$$E\|u_t^h\|^2_{W_2^l}\leqslant NE\|u_0\|^2_{W_2^l}+NE\int_0^t\big(\|f_s\|^2_{W_2^l}+\|g_s\|^2_{W_2^{l+1}}\big)\,ds \tag{3.27}$$

for $t\in[0,T]$. Now we return to (3.25) and use the Davis inequality to get

$$E\sup_{t\leqslant T}\|u_t^h\|^2_{W_2^l}\leqslant E\|u_0\|^2_{W_2^l}+NE\int_0^T\big(\|f_t\|^2_{W_2^l}+\|g_t\|^2_{W_2^{l+1}}\big)\,dt+N_1J, \tag{3.28}$$

where

$$J=E\bigg(\int_0^T\sum_{\rho=1}^\infty\Big|\sum_{|\alpha|\leqslant l}(D^\alpha u_t^h,D^\alpha M^{h,\rho}u_t^h+D^\alpha g_t^\rho))\Big|^2\,dt\bigg)^{1/2}.$$

By the Cauchy–Bunyakovsky–Schwarz inequality,

$$\sum_{\rho=1}^\infty\Big|\sum_{|\alpha|\leqslant l}(D^\alpha u_t^h,D^\alpha g_t^\rho))\Big|^2\leqslant\|u_t^h\|^2_{W_2^l}\|g_t\|^2_{W_2^l},$$

and, by Remark 3.1 (i),

$$\sum_{\rho=1}^\infty\Big|\sum_{|\alpha|\leqslant l}(D^\alpha u_t^h,D^\alpha M^{h,\rho}u_t^h)\Big|^2\leqslant N\|u_t\|^4_{W_2^l}.$$

Hence

$$J\leqslant J_1+J_2,$$

where

$$J_1=E\bigg(\int_0^T\sum_{\rho=1}^\infty\Big|\sum_{|\alpha|\leqslant l}(D^\alpha u_t^h,D^\alpha M^{h,\rho}u_t^h)\Big|^2\,dt\bigg)^{1/2}\leqslant NE\Big(\int_0^T\|u_t^h\|^4_{W_2^l}\,dt\Big)^{1/2}$$
$$\leqslant NE\bigg(\sup_{t\leqslant T}\|u_t^h\|_{W_2^l}\bigg(\int_0^T\|u_t^h\|^2_{W_2^l}\,dt\bigg)^{1/2}\bigg)\leqslant\frac{1}{4N_1}E\sup_{t\leqslant T}\|u_t^h\|^2_{W_2^l}+N_2E\int_0^T\|u_t^h\|^2_{W_2^l}\,ds$$

and

$$J_2=E\bigg(\int_0^T\sum_{\rho=1}^\infty\Big|\sum_{|\alpha|\leqslant l}(D^\alpha u_t^h,D^\alpha g_t^\rho))\Big|^2\,dt\bigg)^{1/2}\leqslant NE\bigg(\int_0^T\|u_t^h\|^2_{W_2^l}\|g_t\|^2_{W_2^l}\,dt\bigg)^{1/2}$$
$$\leqslant E\bigg(\sup_{t\leqslant T}\|u_t^h\|_{W_2^l}\bigg(\int_0^T\|g_t^h\|^2_{W_2^l}\,dt\bigg)^{1/2}\bigg)\leqslant\frac{1}{4N_1}E\sup_{t\leqslant T}\|u_t^h\|^2_{W_2^l}+N_2E\int_0^T\|g_t\|^2_{W_2^l}\,ds.$$

Thus, from (3.28) we get

$$E\sup_{t\leqslant T}\|u^h_t\|^2_{W^l_2}\leqslant E\|u_0\|^2_{W^l_2}+\frac{1}{2}E\sup_{t\leqslant T}\|u^h_t\|^2_{W^l_2}+NE\int\limits_0^T\left(\|f_t\|^2_{W^l_2}+\|g_t\|^2_{W^{l+1}_2}\right)ds,$$

which proves (3.23). □

Lemma 3.6. *Let $n\geqslant 0$ be an integer, let $\phi\in W^{n+1}_2$, $\psi\in W^{n+2}_2$, and $\lambda,\mu\in\mathbb{R}^d\setminus\{0\}$. Set*

$$\partial_\lambda\phi=\lambda^i D_i\phi,\quad \partial_{\lambda\mu}=\partial_\lambda\partial_\mu.$$

Then

$$\frac{\partial^n}{(\partial h)^n}\delta_{h,\lambda}\phi(x)=\int\limits_0^1\theta^n\partial^{n+1}_\lambda\phi(x+h\theta\lambda)\,d\theta, \tag{3.29}$$

$$\frac{\partial^n}{(\partial h)^n}\delta^h_\lambda\phi(x)=\frac{1}{2}\int\limits_{-1}^1\theta^n\partial^{n+1}_\lambda\phi(x+h\theta\lambda)\,d\theta, \tag{3.30}$$

$$\frac{\partial^n}{(\partial h)^n}\delta_\lambda\delta_\mu\psi(x)=\frac{1}{4}\int\limits_{-1}^1\int\limits_{-1}^1(\theta_1\partial_\lambda+\theta_2\partial_\mu)^n\partial_{\lambda\mu}\psi(x+h(\theta_1\lambda+\theta_2\mu))\,d\theta_1 d\theta_2 \tag{3.31}$$

for almost all $x\in\mathbb{R}^d$ and for each $h\in\mathbb{R}$. Hence

$$\frac{\partial^n}{(\partial h)^n}\delta_{h,\lambda}\phi\big|_{h=0}=\frac{1}{n+1}\partial^{n+1}_\lambda\phi,\quad \frac{\partial^n}{(\partial h)^n}\delta^h_\lambda\phi\big|_{h=0}=\frac{B_n}{n+1}\partial^{n+1}_\lambda\phi, \tag{3.32}$$

$$\frac{\partial^n}{(\partial h)^n}\delta_\lambda\delta_\mu\psi\big|_{h=0}=\sum_{r=0}^n A_{n,r}\partial^{r+1}_\lambda\partial^{n-r+1}_\mu\psi, \tag{3.33}$$

where

$$B_n=\begin{cases}0 & \text{if } n \text{ is odd}\\ 1 & \text{if } n \text{ is even}\end{cases},\quad A_{nr}=\begin{cases}0 & \text{if } n \text{ or } r \text{ is odd}\\ \frac{n!}{(r+1)!(n-r+1)!} & \text{if } n \text{ and } r \text{ are even}\end{cases}. \tag{3.34}$$

Furthermore, if $l\geqslant 0$ is an integer and $\phi\in W^{n+2+l}_2$ and $\psi\in W^{n+3+l}_2$, then

$$\Big\|\delta_{h,\lambda}\phi-\sum_{i=0}^n\frac{h^i}{(i+1)!}\partial^{i+1}_\lambda\phi\Big\|_{W^l_2}\leqslant\frac{|h|^{n+1}}{(n+2)!}\|\partial^{n+2}_\lambda\phi\|_{W^l_2}, \tag{3.35}$$

$$\Big\|\delta^h_\lambda\phi-\sum_{i=0}^n\frac{h^i}{(i+1)!}B_i\partial^{i+1}_\lambda\phi\Big\|_{W^l_2}\leqslant\frac{|h|^{n+1}}{(n+2)!}\|\partial^{n+2}_\lambda\phi\|_{W^l_2}, \tag{3.36}$$

$$\Big\|\delta^h_\lambda\delta^h_\mu\psi-\sum_{i=0}^n h^i\sum_{r=0}^i A_{i,r}\partial^{r+1}_\lambda\partial^{i-r+1}_\mu\psi\Big\|_{W^l_2}\leqslant N|h|^{n+1}\|\psi\|_{W^{l+n+3}_2}, \tag{3.37}$$

where $N=N(|\lambda|,|\mu|,d,n)$.

Proof. Clearly, it suffices to prove the lemma for $\phi, \psi \in C_0^\infty(\mathbb{R}^d)$. For $n = 0$ formulas (3.29) and (3.30) are obtained by applying the Newton–Leibnitz formula to $\phi(x + \theta h\lambda)$ and $\phi(x + \theta h\lambda) - \phi(x - \theta h\lambda)$ as functions of θ from $[0, 1]$ and $[-1, 1]$, respectively. Applying (3.30) one more time derives (3.31) from (3.30) for $n = 0$. After that for $n \geqslant 1$ one obtains (3.29)–(3.31) by differentiating both parts of these equations written with $n = 1$.

By the Taylor formula, for smooth $f(h)$ we have

$$f(h) = \sum_{i=0}^{n} \frac{h^i}{i!} \frac{d^i}{(dh)^i} f(0) + \frac{h^{n+1}}{n!} \int_0^1 (1-\theta)^n \frac{d^{n+1}}{(dh)^{n+1}} f(\theta h)\, d\theta.$$

Applying this to

$$\delta_{h,\lambda}\phi(x) = \int_0^1 \partial_\lambda \phi(x + h\theta\lambda)\, d\theta, \quad \delta_\lambda^h \phi(x) = \frac{1}{2} \int_{-1}^1 \partial_\lambda \phi(x + h\theta\lambda)\, d\theta$$

as functions of h and verifying (3.32), we see that

$$\delta_{h,\lambda}\phi(x) = \sum_{i=0}^{n} \frac{h^i}{(i+1)!} \partial_\lambda^{i+1}\phi(x) + \frac{h^{n+1}}{n!} \int_0^1 \int_0^1 (1-\theta_2)^n \theta_1^{n+1} \partial_\lambda^{n+2} \phi(x + h\theta_1\theta_2\lambda)\, d\theta_1 d\theta_2, \quad (3.38)$$

$$\delta_\lambda^h\phi(x) = \sum_{i=0}^{n} \frac{h^i}{(i+1)!} B_i \partial_\lambda^{i+1}\phi(x) + \frac{h^{n+1}}{2n!} \int_0^1 \int_{-1}^1 (1-\theta_2)^n \theta_1^{n+1} \partial_\lambda^{n+2} \phi(x + h\theta_1\theta_2\lambda)\, d\theta_1 d\theta_2. \quad (3.39)$$

Hence we get (3.35) and (3.36) by noting that, by the Minkowski inequality, the W_2^l-norm of the last terms in Equations (3.38) and (3.39) is less than the W_2^l-norm of $\partial_\lambda^{n+2}\phi$ times

$$\frac{|h|^{n+1}}{n!} \int_0^1 \int_0^1 (1-\theta_2)^n \theta_1^{n+1}\, d\theta_1 d\theta_2 = \frac{|h|^{n+1}}{(n+2)!}.$$

Similarly, to get (3.37) from (3.31), we need only verify (3.33) and see that the left-hand side of (3.37) is the W_2^l-norm of

$$\frac{h^{n+1}}{4n!} \int_0^1 \int_{-1}^1 \int_{-1}^1 (1-\theta_3)^n (\theta_1\partial_\lambda + \theta_2\partial_\mu)^{n+1} \partial_{\lambda\mu}\psi(x + h\theta_3(\theta_1\lambda + \theta_2\mu))\, d\theta_1 d\theta_2 d\theta_3,$$

and apply the Minkowski inequality. □

Remark 3.2. Formula (3.29) with $n = 1$ and the Minkowski inequality implies

$$\|\delta_{h,\lambda}\phi\|_{L_2} \leqslant \|\partial_\lambda\phi\|_{L_2}.$$

Applying this inequality to finite differences of ϕ and using induction, we easily conclude that $W_2^{l+r} \subset W_{h,2}^{l,r}$, where for integers $l \geqslant 0$ and $r \geqslant 1$ we denote by $W_{h,2}^{l,r}$ the Hilbert space of functions φ on $\mathbb{R}^d$ with the norm $\|\varphi\|_{l,r,h}$ defined by

$$\|\varphi\|_{l,r,h}^2 = \sum_{\lambda_1,\ldots,\lambda_r \in \Lambda} \|\delta_{h,\lambda_1} \cdot \ldots \cdot \delta_{h,\lambda_r}\varphi\|_{W_2^l}^2. \quad (3.40)$$

We also set $W^{l,0}_{h,2} = W^l_2$. Then for any $\phi \in W^{l+r}_2$ we have

$$\|\varphi\|_{l,h,r} \leqslant N\|\varphi\|_{W^{l+r}_2},$$

where N depends only on $|\Lambda_0|^2 := \sum\limits_{\lambda\in\Lambda_0} |\lambda|^2$ and r.

We set

$$\mathscr{L}^{(0)}_t = \sum_{\lambda,\mu\in\Lambda} \mathfrak{a}^{\lambda\mu}_t \partial_\lambda\partial_\mu + \sum_{\lambda\in\Lambda_0} (\mathfrak{p}^\lambda_t - \mathfrak{q}^\lambda_t)\partial_\lambda,$$
$$\mathscr{M}^{(0)\rho}_t = \sum_{\lambda\in\Lambda} \mathfrak{b}^{\lambda\rho}_t \partial_\lambda$$

and for integers $n \geqslant 1$ introduce the operators

$$\mathscr{L}^{(n)}_t = \sum_{\lambda,\mu\in\Lambda_0} \mathfrak{a}^{\lambda\mu}_t \sum_{r=0}^{n} A_{n,r}\partial^{r+1}_\lambda \partial^{n-r+1}_\mu + (n+1)^{-1} \sum_{\lambda\in\Lambda_0} (\mathfrak{a}^{\lambda 0}_t + \mathfrak{a}^{0\lambda}_t) B_n \partial^{n+1}_\lambda$$
$$+ (n+1)^{-1} \sum_{\lambda\in\Lambda_0} (\mathfrak{p}^\lambda_t + (-1)^{n+1}\mathfrak{q}_t)\partial^{n+1}_\lambda,$$
$$\mathscr{M}^{(n)\rho}_t = (n+1)^{-1} \sum_{\lambda\in\Lambda_0} \mathfrak{b}^{\lambda\rho}_t B_n \partial^{n+1}_\lambda,$$
$$\mathscr{O}^{h(n)}_t = L^h_t - \sum_{i=0}^{n} \frac{h^i}{i!}\mathscr{L}^{(i)}_t,$$
$$\mathscr{R}^{h(n)\rho}_t = M^{h,\rho}_t - \sum_{i=0}^{n} \frac{h^i}{i!}\mathscr{M}^{(i)\rho}_t,$$

where $A_{n,r}$ and B_n are defined by (3.34).

Remark 3.3. Formally, for $n \geqslant 1$ the values $\mathscr{L}^{(n)}_t\phi$ and $\mathscr{M}^{(n)\rho}_t\phi$ are obtained as the values at $h = 0$ of the n-th derivatives in h of $L^h_t\phi$ and $M^{h,\rho}_t\phi$.

Remark 3.4. Owing to Assumption 2.4, we have

$$\mathscr{L}^{(0)}_t = \mathscr{L}_t, \quad \mathscr{M}^{(0)\rho}_t = \mathscr{M}^\rho_t. \tag{3.41}$$

Note also that, by (3.35)-(3.37), under Assumptions 2.2 and 2.6 with $\mathfrak{m} = m$, for $\phi \in W^{n+2+l}_2$ and $\psi \in W^{n+3+l}_2$ we have for $l \leqslant m$

$$\begin{aligned} \|\mathscr{O}^{h(n)}_t\psi\|_{W^l_2} &\leqslant N|h|^{n+1}\|\psi\|_{W^{l+n+3}_2}, \\ \|\mathscr{R}^{h(n)}_t\phi\|_{W^l_2} &\leqslant N|h|^{n+1}\|\phi\|_{W^{l+n+2}_2}, \end{aligned} \tag{3.42}$$

where N denotes constants depending only on n, d, m, $K_0, \ldots, K_{m\vee 2}$, C_m, and Λ.

Let $k \in [1, m/2]$ be an integer. The functions $u^{(1)}$,..., $u^{(k)}$ we need in the expansion (2.9) will be obtained as the result of embedding in $C_b(\mathbb{R}^d)$ appropriate functions $v^{(1)}$,...,$v^{(k)}$, with values in certain Sobolev spaces. We determine the functions $v^{(1)}_t$,...,$v^{(k)}_t$ as follows. We define

$v_t^{(0)}$ as the solution of (2.1) from Theorem 2.1 and we are going to find $v^{(1)}$,..., $v^{(k)}$ by solving the following system of stochastic partial differential equations:

$$dv_t^{(n)} = \Big(\mathscr{L}_t v_t^{(n)} + \sum_{l=1}^{n} \tbinom{n}{l} \mathscr{L}_t^{(l)} v_t^{(n-l)}\Big)\, dt + \Big(\mathscr{M}_t^{\rho} v_t^{(n)} + \sum_{l=1}^{n} \tbinom{n}{l} \mathscr{M}_t^{(l)\rho} v_t^{(n-l)}\Big)\, dw_t^{\rho}, \quad n = 1, ..., k. \tag{3.43}$$

Theorem 3.7. *Let Assumptions 2.1, 2.2, 2.3, and 2.6 hold with $\mathfrak{m} = m \geqslant 2k$. Then there exists a unique set $v^{(1)}, ..., v^{(k)}$ of solutions of (3.43) with initial condition $v_0^{(1)} = ... = v_0^{(k)} = 0$ and such that $v^{(n)} \in \mathbb{H}^{m-2n}(T)$, $n = 1, ..., k$. Furthermore, with probability one $v^{(n)}$ are continuous W_2^{m-1-2n}-valued predictable processes and there exists a constant N depending only on T, d, Λ, m, k, K_0,...,K_m and C_m, such that for $n = 1, ..., k$*

$$E \sup_{t \leqslant T} \|v_t^{(n)}\|^2_{W_2^{m-2n}} \leqslant N \mathscr{K}_m^2 \tag{3.44}$$

for $h > 0$. Moreover, if $\mathfrak{p}_\lambda = \mathfrak{q}^\lambda = 0$ for $\lambda \in \Lambda_0$, then (3.44) holds for all $h \neq 0$, and hence $v^{(n)} = 0$ for odd $n \leqslant k$.

Proof. Since for each $n = 1, \dots, k$ the equation for $v_t^{(n)}$ does not involve the unknown functions $v^{(l+1)}$,..., $v^{(n)}$, we can prove the theorem recursively on $n \leqslant k$. Denote

$$S^{(n)} = \sum_{i=1}^{n} \tbinom{n}{i} \mathscr{L}^{(i)} v^{(n-i)}, \quad R^{(n)\rho} = \sum_{i=1}^{n} \tbinom{n}{i} \mathscr{M}^{(i)\rho} v^{(n-i)},$$

and first let $n = 1$. By Theorem 2.1, we have $v^{(0)} \in \mathbb{H}^m(T)$ such that the estimate (2.3) holds. Note that $R^{(1)} = 0$ and owing to Assumption 2.6

$$\|\mathscr{L}_t^{(1)} v_t^{(0)}\|_{W_2^{m-2}} \leqslant N \|v_t^{(0)}\|_{W_2^m},$$

which, by Theorem 2.1, implies the existence of a unique $v^{(1)} \in \mathbb{H}^{m-2}(T)$ satisfying (3.43) with zero initial condition. Furthermore, $v_t^{(1)}$ is a continuous W_2^{m-3}-valued function (a.s.) and (3.44) holds with $n = 1$. Passing to higher n, we assume that $m \geqslant k \geqslant 2$ and for an $n \in \{2, ..., k\}$ we have found $v^{(1)}$,...,$v^{(n-1)}$ with the asserted properties. Then $\mathscr{M}^{(1)} v^{(n-1)} = 0$ and

$$\|\mathscr{L}_t^{(1)} v_t^{(n-1)}\|_{W_2^{m-2n}} \leqslant N \|v_t^{(n-1)}\|_{W_2^{m-2n+2}} = N \|v_t^{(n-1)}\|_{W_2^{m-2(n-1)}},$$

and for $i \geqslant 2$

$$\|\mathscr{L}_t^{(i)} v_t^{(n-i)}\|_{W_2^{m-2n}} \leqslant N \|v_t^{(n-i)}\|_{W_2^{m-2n+(i+2)}} \leqslant N \|v_t^{(n-i)}\|_{W_2^{m-2(n-i)}},$$

$$\sum_{k=1}^{\infty} \|\mathscr{M}^{(i)\rho} v^{(n-i)}\|^2_{W_2^{m-2n+1}} \leqslant N \|v^{(n-i)}\|^2_{W_2^{m-2n+1+(i+1)}} \leqslant N \|v_t^{(n-i)}\|^2_{W_2^{m-2(n-i)}}.$$

By the induction hypothesis,

$$E \int_0^T \|S_t^{(n)}\|^2_{W_2^{m-2n}}\, dt \leqslant N \mathscr{K}_m^2, \quad E \int_0^T \|R_t^{(n)}\|^2_{W_2^{m-2n+1}}\, dt \leqslant N \mathscr{K}_m^2, \tag{3.45}$$

which, by Theorem 2.1, yields the existence of a unique $v^{(n)} \in \mathbb{H}^{m-2n}(T)$ satisfying (3.43) with zero initial condition. This theorem also yields the continuity property of $v_t^{(n)}$ and an estimate, that combined with (3.45) yields (3.44) for $h > 0$. The proof of the existence of $v^{(1)}$,...,$v^{(k)}$ with the stated properties is complete. We obtain the uniqueness by inspecting the above proof in which each $v^{(n)}$ is found uniquely.

Note that $\mathscr{M}^{(n)} = 0$ for odd $n \leqslant k$ by (3.32). Assume now that $\mathfrak{p}^\lambda = \mathfrak{q}^\lambda = 0$ for $\lambda \in \Lambda_0$. Then also $\mathscr{L}^{(n)} = 0$ for odd $n \leqslant k$ by (3.32) and (3.33). Hence $S^{(1)} = 0$ and $R^{(1)} = 0$, which implies $v^{(0)} = 0$. Assume that $k \geqslant 2$ and that for an odd $n \leqslant k$ we have $v^{(l)} = 0$ for all odd $l < n$. Then $\mathscr{L}^{(n-i)}v^{(i)} = 0$ and $\mathscr{M}^{(n-i)}v^{(i)} = 0$ for all $i = 1$,...,n since either i or $n-i$ is odd. Thus, $S^{(n)} = 0$ and $R^{(n)} = 0$. Hence $v^{(n)} = 0$ for all $n \leqslant k$. □

Lemma 3.8. *Let Assumptions* 2.1, 2.2, 2.3, 2.5 (iii), *and* 2.6, *hold with*

$$m = l + 2k + 2$$

for some integers $k \geqslant 0$ and $l \geqslant 0$. Set

$$r_t^h = v_t^h - \sum_{j=1}^{k} \frac{h^j}{j!} v_t^{(j)}, \tag{3.46}$$

where v^h is the unique L_2-valued solution of (2.4), (2.5), *$v^{(0)}$ is the solution of* (2.1)-(2.2), *and $(v^{(n)})_{n=1}^k$ is the solution of* (3.43), *given by Theorem* 3.7. *Then $r_0^h = 0$, $r^h \in \mathbb{W}_2^{0,l}(T)$, and*

$$dr_t^h = (L_t^h r_t^h + F_t^h)\,dt + (M_t^{h,\rho} r_t^h + G_t^{h,\rho})\,dw_t^\rho, \tag{3.47}$$

where

$$F_t^h := \sum_{j=0}^{k} \frac{h^j}{j!} \mathscr{O}_t^{h(k-j)} v_t^{(j)}, \quad G_t^{h,\rho} := \sum_{j=0}^{k} \frac{h^j}{j!} \mathscr{R}_t^{h(k-j)} v_t^{(j)}.$$

Finally, $F^h \in \mathbb{W}_2^l(T)$ and $G^{h,\cdot} \in \mathbb{W}_2^{l+1}(T)$.

Proof. We have $v^{(h)} \in \mathbb{H}^l(T)$ due to Assumptions 2.1 and 2.5(iii), and $v^{(j)} \in \mathbb{H}^l(T)$, for $j \leqslant k$ by Theorems 2.1 and 3.7. Hence $r^h \in \mathbb{H}^l$. Using the equations for v^h and $v^{(n)}$ for $n = 0, ..., k$, we can easily see that (3.47) holds with $\widehat{F}$ and $\widehat{G}$ in place of F and G, respectively, where

$$\widehat{F}^h = L^h v^{(0)} - \mathscr{L} v^{(0)} + \sum_{1 \leqslant j \leqslant k} L^h v^{(j)} \frac{h^j}{j!} - \sum_{1 \leqslant j \leqslant k} \mathscr{L} v^{(j)} \frac{h^j}{j!} - I^h,$$

$$G^{h,\rho} = M^{h,\rho} v^{(0)} - \mathscr{M}^\rho v^{(0)} + \sum_{1 \leqslant j \leqslant k} M^{h,\rho} v^{(j)} \frac{h^j}{j!} - \sum_{1 \leqslant j \leqslant k} \mathscr{M}^\rho v^{(j)} \frac{h^j}{j!} - J^{h,\rho}$$

with

$$I^h = \sum_{1 \leqslant j \leqslant k} \sum_{i=1}^{j} \frac{1}{i!(j-i)!} \mathscr{L}^{(i)} v^{(j-i)} h^j,$$

$$J^{h,\rho} = \sum_{1 \leqslant j \leqslant k} \sum_{i=1}^{j} \frac{1}{i!(j-i)!} \mathscr{M}^{(i)\rho} v^{(j-i)} h^j,$$

where, as usual, summations over empty sets mean zero. Note that

$$I^h = \sum_{i=1}^{k}\sum_{j=i}^{k}\frac{1}{i!(j-i)!}\mathscr{L}^{(i)}v^{(j-i)}h^j = \sum_{i=1}^{k}\sum_{l=0}^{k-i}\frac{1}{i!l!}\mathscr{L}^{(i)}v^{(l)}h^{l+i}$$

$$= \sum_{l=0}^{k-1}\frac{h^l}{l!}\sum_{i=1}^{k-l}\frac{h^i}{i!}\mathscr{L}^{(i)}v^{(l)} = \sum_{j=0}^{k}\frac{h^j}{j!}\sum_{i=1}^{k-j}\frac{h^i}{i!}\mathscr{L}^{(i)}v^{(j)}$$

and, similarly,

$$J^{h,\rho} = \sum_{j=1}^{k}\sum_{i=1}^{j}\frac{1}{i!(j-i)!}\mathscr{M}^{(i)\rho}v^{(j-i)}h^j = \sum_{j=0}^{k}\frac{h^j}{j!}\sum_{i=1}^{k-j}\frac{h^i}{i!}\mathscr{M}^{(i)\rho}v^{(j)}.$$

After that the fact that $\widehat{F} = F$ and $\widehat{G} = G$ follows by simple arithmetic. To prove the last assertion, we note that for $j = 0, 1, ..., k$

$$l + k - j + 2 \leqslant m - 2j, \quad l + k - j + 1 \leqslant m - 2j - 1.$$

Thus, by Lemma 3.6 for $j = 0, 1, ..., k$, $(t, \omega) = [0, T] \times \Omega$

$$\|\mathscr{O}_t^{h(k-j)}v_t^{(j)}\|_{W_2^l} \leqslant N\|v_t^{(j)}\|_{W_2^{l+k-j+2}} \leqslant N\|v_t^{(j)}\|_{W_2^{m-2j}},$$

$$\|\mathscr{R}_t^{h(k-j)}v_t^{(j)}\|_{W_2^l} \leqslant N\|v_t^{(j)}\|_{W_2^{l+k-j+1}} \leqslant N\|v_t^{(j)}\|_{W_2^{m-2j-1}},$$

which implies $F^h \in \mathbb{W}_2^l(T)$ and $G^{h,\cdot} \in \mathbb{W}_2^{l+1}(T)$ by Theorems 2.1 and 3.7. □

4 Proof of Theorems 2.2 and 2.4

First we present a theorem which, as we will see it later, is more general than Theorem 2.4. Theorem 2.2 can be obtained similarly.

Theorem 4.1. *Let Assumptions 2.2, 2.3, 2.4, 2.5, and 2.6 hold with*

$$\mathfrak{m} = m = l + 2k + 3 \tag{4.1}$$

for some integer $k \geqslant 0$. Then for r_t^k, defined as in Lemma 3.8, we have

$$E\sup_{t\leqslant T}\|r_t^h\|_{W_2^l}^2 \leqslant N|h|^{2(k+1)}\mathscr{K}_m^2 \tag{4.2}$$

for $h > 0$, where N depends only on T, d, Λ, m, l, K_0,...,K_m, C_m, and A_0,..., A_{l+1}. Moreover, if $\mathfrak{p}^\lambda = \mathfrak{q}^\lambda = 0$ for $\lambda \in \Lambda_0$, then $v^{(j)} = 0$ in (3.46) for odd $j \leqslant k$, and if k is odd then it is sufficient to assume $m \geqslant l+2k+2$ in place of $m = l+2k+3$ in (4.1) to have the estimate (4.2).

Proof. By Lemma 3.8, we have $F^h \in \mathbb{W}_2^l(T)$ and $G^{h,\cdot} \in \mathbb{W}_2^{l+1}(T)$, which, by Lemma 3.8 and Theorem 3.5, yields

$$E\sup_{t\leqslant T}\|r_t^h\|_{W_2^l}^2 \leqslant NE\int_0^T(\|F_t^h\|_{W_2^l}^2 + \|G_t^h\|_{W_2^{l+1}}^2)\,dt \tag{4.3}$$

with a constant N depending only on d, l, T, and $A_0, \dots, A_{l+1}$ Let (4.1) hold. Then for $j = 0, ..., k$

$$l + k - j + 3 \leqslant m - 2j, \\ l + k - j + 2 \leqslant m - 2j + 1 \tag{4.4}$$

and, by Remark 3.4,

$$\|\mathscr{O}_t^{h(k-j)} v_t^{(j)}\|_{W_2^l} \leqslant N|h|^{k-j+1} \|v_t^{(j)}\|_{W_2^{l+k-j+3}} \leqslant N|h|^{k-j+1} \|v_t^{(j)}\|_{W_2^{m-2j}}, \\ \|\mathscr{R}_t^{h(k-j)} v_t^{(j)}\|_{W_2^l} \leqslant N|h|^{k-j+1} \|v^{(j)}\|_{W_2^{l+k-j+2}} \leqslant N|h|^{k-j+1} \|v^{(j)}\|_{W_2^{m-2j-1}}. \tag{4.5}$$

Using Theorem 3.7, we see that

$$E \int_0^T \|F_t^h\|_{W_2^l}^2 + \|G^h\|_{W_2^{l+1}}^2 \, dt \leqslant N|h|^{2(k+1)} \mathscr{K}_m^2,$$

which, by Theorem 3.5, implies the estimate (4.2). If $p^\lambda = q^\lambda = 0$ for $\lambda \in \Lambda_0$, then, by Theorem 3.7, $v^{(j)} = 0$ for odd $j \leqslant k$. (Note that it follows also from (4.2) valid now for all $h \neq 0$ since $v^h = v^{-h}$ due to that v^h and v^{-h} are the unique L_2 solutions of the same problem (2.4), (2.5).) If, in addition, k is odd then $v^{(k)} = 0$. Thus (4.5) obviously holds for $j = k$ and to have it also for $j \leqslant k - 1$ we need only $m = l + 2k + 2$. □

By the Sobolev theorem on embedding of W_2^l into C_b for $l > d/2$, there exists a linear operator $I : W_2^l \to C_b$ such that $I\varphi(x) = \varphi(x)$ for almost every $x \in \mathbb{R}^d$ and

$$\sup_{\mathbb{R}^d} |I\varphi| \leqslant N \|\varphi\|_{W_2^l}$$

for all $\varphi \in W_2^l$, where N is a constant depending only on d. One has also the following lemma on the embedding $W_2^l \subset l_2(\mathbb{G}_h)$, that we have already referred to, when we used Remark 2.3 on the existence of a unique $l_2(\mathbb{G}_h)$-valued continuous solution $\{u_t(x) : x \in \mathbb{G}_h\}$ to Equation (2.4).

Lemma 4.2. *For all $\varphi \in W_2^l(\mathbb{R}^d)$, $l > d/2$, $h \in (0, 1)$*

$$\sum_{x \in \mathbb{G}_h} |I\varphi(x)|^2 h^d \leqslant N \|\varphi\|_{W_2^l}^2, \tag{4.6}$$

where N is a constant depending only on d.

Proof. This lemma is a straightforward consequence of the Sobolev theorem on embedding of W_2^l¡ functions on the unit ball B_1 of R^d, into $C(B_1)$, the space of continuous functions on B_1 (cf., for example, [6].) □

Set $R_t^h = I r_t^h$. Recall that $\Lambda^0 = \{0\}$, $\delta_{h,0}$ is the identity operator and $\delta_{h,\lambda} = \delta_{h,\lambda_1} \cdot \ldots \cdot \delta_{h,\lambda_n}$ for $(\lambda_1, \dots, \lambda_n) \in \Lambda^n$, $n \geqslant 1$. Then we have the following consequence of Theorem 4.1

Corollary 4.3. *If for some integer $n \geqslant 0$ we have $l > n + d/2$ in Theorem* 4.1. *Then for $\lambda \in \Lambda^n$*

$$E \sup_{t \in [0,T]} \sup_{x \in \mathbb{R}^d} |\delta_{h,\lambda} R_t^h(x)|^2 \leqslant N h^{2(k+1)} \mathscr{K}_m^2$$

for $h > 0$ with a constant N depending only on Λ, d, m, l, T, K_0, ..., K_m, A_0,...,A_{l+1}, and C_m.

Proof. We set $j = n - l$. Then $j > d/2$ and, by the Sobolev theorem on embedding of W_2^j into C_b and Remark 3.2, from Theorem 4.1 we get

$$E \sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^d} |\delta_{h,\lambda} R_t^h(x)|^2 \leqslant C_1 E \sup_{t\in[0,T]} \|R_t^h\|_{j,h,n}^2 \leqslant C_2 E \sup_{t\in[0,T]} \|R_t^h\|_{W_2^l}^2 \leqslant N h^{2(k+1)} \mathscr{K}_m^2.$$

Similarly, by Lemma 4.2 and Remark 3.2,

$$E \sup_{t\in[0,T]} \sum_{x\in\mathbb{G}_h} |\delta_{h,\lambda} R_t^h(x)|^2 h^d \leqslant C_1 E \sup_{t\in[0,T]} \|\delta_{h,\lambda} R_t^h\|_{W_2^j}^2 \leqslant C_2 E \sup_{t\in[0,T]} \|R_t^h\|_{W_2^l}^2 \leqslant N h^{2(k+1)} \mathscr{K}_m^2. \quad \square$$

Now we show that Theorem 2.4 follows from the above corollary. We define

$$\widehat{u}^h = I v^h, \quad u^{(j)} = I v^{(j)}, \quad j = 0, ..., k,$$

where v^h is the unique $\mathscr{F}_t$-adapted continuous $L_2(\mathbb{R}^d)$-valued solution of Equation (2.4) with initial condition ψ, the processes $v^{(0)}$,...,$v^{(k)}$ are given by Theorem 3.7, I is the embedding operator from W^l into C_b. By Theorem 3.5, v^h is a continuous W_2^l-valued process, and, by Theorem 3.7, $v^{(j)}$, $j = 1, 2, ..., k$, are W_2^{m-2k}-valued continuous processes. Since $l > d/2$ and hence $m - 2k > d/2$, the processes $\widehat{u}^h$ and $u^{(j)}$ are well-defined and clearly (3.46) implies (2.9). To show that Corollary 4.3 yields Theorem 2.4, we need only show that almost surely

$$\widehat{u}_t^{(h)}(x) = u_t^h(x) \quad \text{for all } t \in [0, T] \tag{4.7}$$

for each $x \in \mathbb{G}_h$, where u^h is the unique $\mathscr{F}_t$-adapted l_2-valued continuous solution of (2.4). To see this, let φ be a compactly supported nonnegative smooth function on $\mathbb{R}^d$ with unit integral, and for a fixed $x \in \mathbb{G}_h$ set

$$\varphi_\varepsilon(y) = \varphi((y - x)/\varepsilon)$$

for $y \in \mathbb{R}^d$ and $\varepsilon > 0$. Since $\widehat{u}^h$ is a continuous L_2-valued solution of (2.4), for each ε almost surely

$$\int_{\mathbb{R}^d} \widehat{u}_t^h(y)\varphi_\varepsilon(y)\, dy = \int_{\mathbb{R}^d} \widehat{u}(y)\varphi_\varepsilon(y)\, dy + \int_0^t \int_{\mathbb{R}^d} (L_s^h \widehat{u}_s^h(y) + f_s(y))\varphi_\varepsilon(y)\, dy\, ds$$

$$+ \int_0^t \int_{\mathbb{R}^d} (M_s^{h,\rho} \widehat{u}_s^h(y) + g_s^\rho(y))\varphi_\varepsilon(y)\, dy\, dw_s^\rho$$

for all $t \in [0, T]$. Letting here $\varepsilon \to 0$ we see that both sides converges in probability, uniformly in $t \in [0, T]$, and thus we get that almost surely

$$\widehat{u}_t^h(x) = u_0(x) + \int_0^t L_s^h \widehat{u}_s^h(x) + f_s(x)\, ds + \int_0^t M_s^{h,\rho} \widehat{u}_s^h(x) + g_s^\rho(x)\, dw_s^\rho$$

for all $t \in [0, T]$. (Remember that u_0, f and g are continuous in x by virtue of Remark 2.1.) Moreover, owing to Lemma 4.2 the restriction of $\widehat{u}_t$ onto $\mathbb{G}_h$ is a continuous $l_2(G_h)$-valued process. Hence, because of the uniqueness of the $l_2(\mathbb{G}_h)$-valued continuous $\mathscr{F}_t$-adapted solution of (2.4) for any l_2-valued $\mathscr{F}_0$-measurable initial condition, we have (4.7), which completes the proof of Theorem 2.4.

References

1. P. G. Ciarlet and J. L. Lions (eds), *Handbook of Numerical Analysis*, Elsevier (2003).

2. H. Dong and N. V. Krylov, "On the rate of convergence of finite-difference approximations for degenerate linear parabolic equations with C^1 and C^2 coefficients," *Electron. J. Diff. Eqns.* **2005** No. 102, 1-25 (2005). http://ejde.math.txstate.edu

3. L. F. Richardson, "The approximative arithmetical solution by finite differences of physical problems involving differential equations," *Philos. Trans. Roy. Soc. London, Ser. A* **210**, 307-357 (1910).

4. L. F. Richardson and J. A. Gaunt, "The deferred approach to the Limit," *Phil. Trans. Roy. Soc. London Ser. A* **226**, 299-361 (1927).

5. G. I. Marschuk and V. V. Shaidurov, *Difference Methods and Their Extrapolations*, Springer, New York etc. (1983).

6. I. Gyöngy and N.V. Krylov, "Accelerated finite difference schemes for second order degenerate elliptic and parabolic problems in the whole space," *Math. Comp.* 80 (2011), 1431-1458.

7. I. Gyöngy and N. V. Krylov, "Accelerated finite difference schemes for stochastic parabolic partial differential equations in the whole space," *SIAM J. Math. Anal.* **42**, No. 5, 2275-2296 (2010).

8. C. Brezinski, "Convergence acceleration during the 20th century," *J. Comput. Appl. Math.* **122**, 1–21 (2000).

9. D. C. Joyce, "Survey of extrapolation processes in numerical analysis," *SIAM Review* **13**, No. 4, 435–490 (1971).

10. N. V. Krylov and B. L. Rozovskii, "On characteristics of the degenerate second order Itô parabolic equations" [in Russian], *Tr. Semin. Im. I. G. Petrovskogo*, **8**, 153-168 (1982).

Submitted on September 19, 2011