On finite difference schemes for degenerate stochastic parabolic partial differential equations

Citation for published version:
Gyongy, I 2011, 'On finite difference schemes for degenerate stochastic parabolic partial differential equations' Journal of Mathematical Sciences, vol 179, no. 1, pp. 100-126. DOI: 10.1007/s10958-011-0584-3

Digital Object Identifier (DOI):
10.1007/s10958-011-0584-3

Published In:
Journal of Mathematical Sciences

ON FINITE DIFFERENCE SCHEMES FOR DEGENERATE STOCHASTIC PARABOLIC PARTIAL DIFFERENTIAL EQUATIONS

ISTVÁN GYÖNGY

ABSTRACT. Finite difference approximations in the space variable for possibly degenerate stochastic parabolic PDEs is 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 generalise some results from [5] and [6] to degenerate SPDEs.

1. Introduction

We study spatial discretisations

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

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

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

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

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

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

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

respectively. The stochastic parabolicity condition is assumed (see Assumption 2.1 below). Such equations arise in filtering theory of partially observed

2000 *Mathematics Subject Classification.* 65M06, 60H15, 65B05.

Key words and phrases. Cauchy problem, stochastic PDEs, finite differences, extrapolation to the limit, Richardson's method .

diffusion processes $Z=(X,Y)$, as equations for the unnormalised 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 $u^h_0=\psi$ on a fixed grid $\mathbb{G}_h$ of mesh-size $|h|$.

Finite difference approximations for deterministic PDEs are studied extensively in the literature. See for instance [2] and the references therein. However, there are only a few results published for degenerate equations. Sharp rate of convergence estimates are obtained in [3] for deterministic (possibly) degenerate parabolic and elliptic SPDEs with monotone finite difference schemes. Rate of convergence estimates of finite difference approximations for stochastic parabolic PDEs are are obtained under the *strong stochastic parabolicity condition*, i. e., when there is a constant $\kappa>0$ such that

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

for all $\omega\in\Omega$, $t\geq0$ 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 (see [9] and [10]). 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 proportional to the error of the method. The possibility of such expansions have been studied thoroughly in numerical analysis. See, for example, the book [8] on Richardson's idea applied to finite difference approximations for deterministic PDEs. In [6] Richardson's idea is implemented to a class of monotone finite difference schemes for (possibly) degenerate parabolic and elliptic PDEs, and in [**?**] Richardson's idea is implemented to stochastic PDEs satisfying the strong parabolicity conditions. Both in[6] and [**?**] 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 generalise some results from [6] and [**?**] 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 of when the finite difference approximations are defined by replacing the partial derivatives $\partial\partial x^i$ by centred finite differences along the basis vector e_i our main theorem reads as follows: The accuracy of the (spatial) finite difference approximations to (1.2)-(1.3) be accelerated to any order if the initial condition and free terms are sufficiently smooth in x and the matrix

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

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 unnormalised 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 [1] and [4]

The paper is organised 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 $(\Omega, \mathcal{F}, P)$, equipped with an increasing family of σ-algebras $(\mathcal{F}_t)_{t\geq 0}$, such that $\mathcal{F}_0$ contains the P-zero sets of $\mathcal{F}$. The σ-algebra of predictable subsets of $\Omega \times [0, \infty)$ is denoted by $\mathcal{P}$. We fix also a sequence of independent Wiener processes $(w_t^\rho)_{\rho=1}^\infty$, such that w_t^ρ is $\mathcal{F}_t$-measurable and $w_t^\rho - w_s^\rho$ is independent of $\mathcal{F}_s$ for $0 \leq s \leq t$, for every integer $\rho \geq 1$. 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

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

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

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

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 $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable functions on $\Omega \times H_T$ for all $\alpha, \beta \in \{0, 1, ...\}$.

A necessary condition that the Cauchy problem (2.1)-(2.2) be well-posed is the condition of *stochastic parabolicity*:

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

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

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

Let $m \geq 0$ 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

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

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

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

Assumption 2.3. We have $\psi \in L_2(\Omega, \mathcal{F}_0, W_2^m)$. 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

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

Remark 2.1. If Assumption 2.3 holds with $m > d/2$ then by Sobolev's 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 continuos 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^1_2-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 $\varphi \in C_0^\infty(\mathbb{R}^d)$

$$(u_t,\varphi) = (u_0,\varphi) + \int_0^t (-a^{ij}_s D_j u_s, D_i\varphi) + (a^j_s D_j u_s,\varphi) + (a_s u_s,\varphi)\,ds$$
$$+ \int_0^t (b^{i\rho}_s D_i u_s + b^\rho_s,\varphi)\,dw^\rho_s$$

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 [7] (see also [1]).

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^{m-1}_2, and there exists a constant N depending only on T, d, m and $K_j, j \le \max(m,2)$, such that*

$$E\sup_{t\le T}\|u_t\|^2_{W^m_2} \le N\mathcal{K}^2_m. \tag{2.3}$$

Remark 2.2. We are going to assume that $m > d/2$. Then by 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, the so constructed 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

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

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

$$du^h_t = (L^h_t u^h_t + f_t)\,dt + (M^{h,\rho}_t u^h_t + g^\rho_t)\,dw^\rho_t, \tag{2.4}$$
$$u^h_0 = \psi, \tag{2.5}$$

with

$$L^h_t = \mathfrak{a}^{\lambda\mu}_t\delta^h_\lambda\delta^h_\mu + \sum_{\lambda\in\Lambda_0}(\mathfrak{p}^\lambda\delta_{h,\lambda} - \mathfrak{q}^\lambda\delta_{-h,\lambda}),\quad M^{h,\rho}_t = \mathfrak{b}^{\lambda\rho}_t\delta^h_\lambda,$$

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 $\mathfrak{a}^{\lambda\mu} = \mathfrak{a}^{\lambda\mu}_t(x)$, $p^\lambda = p^\lambda_t(x)$, $q^\lambda = q^\lambda_t(x)$ are real-valued, and $\mathfrak{b}^\lambda = (\mathfrak{b}^{\lambda\rho}_t(x))^\infty_{\rho=1}$ are l_2-valued, $\mathcal{P}\times\mathcal{B}(\mathbb{R}^d)$-measurable bounded functions on $\Omega\times H_T$, for all $\lambda,\mu\in\Lambda$.

Introduce

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

and 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. Notice 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 $u^h_0 \in l_2(\mathbb{G}_h)$ (a.s.). By Sobolev's embedding of W^r_2 into C_b we have $W^r_2 \subset l_2(\mathbb{G}_h)$ if $r > d/2$, see Lemma 4.2 below. Therefore if

$$\|\psi\|^2_{W^r_2} + \int_0^T \|f(s)\|^2_{W^r_2} + \|g(s)\|^2_{W^r_2}\, ds < \infty \quad (a.s.),$$

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

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 $\rho = 1, 2, ...$

$$\sum_{\lambda,\mu\in\Lambda_0} \mathfrak{a}^{\lambda\mu}_t \lambda^i \mu^j = a^{ij}_t, \qquad \sum_{\lambda\in\Lambda_0} \mathfrak{b}^{\lambda\rho}_t \lambda^i = b^{i\rho}_t, \tag{2.6}$$

$$\sum_{\lambda\in\Lambda_0} \mathfrak{a}^{\lambda 0}\lambda^i + \sum_{\mu\in\Lambda_0} \mathfrak{a}^{0\mu}\mu^i + \sum_{\mu\in\Lambda_0} \mathfrak{p}^{\lambda}\lambda^i - \sum_{\lambda\in\Lambda_0} \mathfrak{q}^{\lambda}\mu^i = a^{i0}_t + a^{0i}_t,$$

$$\mathfrak{a}^{00}_t = a^{00}_t, \quad \mathfrak{b}^{0\rho}_t = b^{0\rho}_t.$$

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

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

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

$$\mathfrak{q}^{e_\alpha} = \mathfrak{p}^{e_\alpha} = 0 \quad \alpha, \beta = 1, ..., d.$$

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

$$\mathfrak{a}^{00} = a^{00}, \quad \mathfrak{a}_t^{e_\alpha e_\beta} = a_t^{\alpha\beta}, \qquad \alpha, \beta = 1, ..., d,$$

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

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

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

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

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

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

(ii) For some integer $d_1 \geq 1$ for each $\lambda \in \Lambda_0$ there are $\mathcal{F} \otimes \mathcal{B}(H_T)$-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 $l \geq 1$ 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 we have

$$|D^j\sigma^{\lambda k}| + |D^j\mathfrak{b}^\lambda| + |D^j\mathfrak{b}^0| \leq A_j \quad \text{for } j \leq l+1,$$

$$|D^j\mathfrak{a}^{\lambda 0}| + |D^j\mathfrak{a}^{0\lambda}| + |D^j\mathfrak{a}^{00}||D^j\mathfrak{p}^\lambda| + |D^j\mathfrak{q}^\lambda| \leq A_j \quad \text{for } j \leq l,$$

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

Remark 2.4. Clearly, Assumption 2.5 (ii) implies

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

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 generalised solution, which by virtue of Sobolev's 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

$$du_t(x) = (L_t u_t(x) + f_t(x))\, dt + M_t^\rho u_t(x) + g^\rho(x)\, dw_t^\rho, \quad u_0(x) = \psi(x)$$

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

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

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

where N is a constant depending only on T, Λ, l, d, m, K_0,....,K_m and A_0,....,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 $k \geq 0$ 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

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

with a constant N independent of h.

Assumption 2.6. Let $\mathfrak{m} \geq 0$ 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*

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

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

If $p^\lambda = q^\lambda = 0$ for $\lambda \in \Lambda_0$ then (2.9)-(2.10) *hold for all $h \neq 0$. Moreover, $u^{(j)} = 0$ for odd $j \leq k$, 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 $u_t^h(x)$ 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

$$\delta_{h,\lambda} u_t^h(x) = \sum_{j=0}^{k} \frac{h^j}{j!} \delta_{h,\lambda} u_t^{(j)}(x) + \delta_{h,\lambda} R_t^h(x)$$

for any $\lambda=(\lambda_1,...,\lambda_n)\in\Lambda^n$ and integer $n\geq 0$, 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 $n\geq 0$. 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 $h>0$ we have (2.9) *and*

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

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 $j\leq k$, 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 $k\geq 0$ and define

$$\bar u^h=\sum_{j=0}^{\tilde k}\bar b_j u^{2^{-j}h},\quad \tilde u^h=\sum_{j=0}^{\tilde k}\tilde b_j u^{2^{-j}h}, \tag{2.14}$$

where

$$(\bar b_0,\bar b_1,...,\bar b_k):=(1,0,0,...,0)\bar V^{-1},$$
$$(\tilde b_0,\tilde b_1,...,\tilde b_{\tilde k}):=(1,0,0,...,0)\tilde V^{-1},\quad \tilde k=[\tfrac{k}{2}],$$

$\bar V^{-1}$ is the inverse of the matrix

$$\bar V^{ij}:=2^{-(i-1)(j-1)},\ i,j=1,...,k+1,$$

and $\tilde V^{-1}$ is the inverse of matrix

$$V^{ij}:=4^{-(i-1)(j-1)},\quad i,j=1,...,\tilde k+1.$$

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

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

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

$$E\sup_{t\leq T}\sup_{x\in\mathbb{G}_h}|\bar u^h_t(x)-u^{(0)}_t(x)|^2\leq Nh^{2(k+1)}\mathcal{K}^2_m \tag{2.16}$$

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

$$E\sup_{t\leq T}\sup_{x\in\mathbb{G}_h}|\tilde u^h_t(x)-u^{(0)}_t(x)|^2\leq N|h|^{2(k+1)}\mathcal{K}^2_m \tag{2.17}$$

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 estimate (2.16) can be obtained in the same way. By Theorem 2.3

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

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

$$\tilde{u}^h = \sum_{j=0}^{\tilde{k}} \tilde{b}_j u^{2^{-j}h} = (\sum_{j=0}^{\tilde{k}} \tilde{b}_j) u^{(0)} + \sum_{j=0}^{\tilde{k}} \sum_{i=1}^{\tilde{k}} \tilde{b}_j \frac{h^{2i}}{(2i)!4^{ij}} u^{(2i)} + h^{k+1} \tilde{r}^h$$

$$= u^{(0)} + \sum_{i=1}^{\tilde{k}} \frac{h^{2i}}{(2i)!} u^{(2i)} \sum_{j=0}^{\tilde{k}} \frac{\tilde{b}_j}{4^{ij}} + h^{k+1} \tilde{r}^h = u^{(0)} + h^{k+1} \tilde{r}^h,$$

since

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

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

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

and the theorem is proved. □

Remark 2.6. Notice 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

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

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

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

respectively. These are sharp estimates by virtue of Remark 2.21 in [3] on finite difference approximations for deterministic parabolic PDEs.

Remark 2.7. Let $\mathfrak{p}^\lambda = \mathfrak{q}^\lambda = 0$ for $\lambda \in \Lambda_0$. Let $n \geq 0$ and assume the conditions of Theorem 2.3 with

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

with an integer $n \geq 0$. Then for $\lambda \in \Lambda^n$ (2.17) holds with $\delta_{h,\lambda} \tilde{u}^h$ and $\delta_{h,\lambda} u^{(0)}$ in place of $\tilde{u}^h$ 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 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 $\bar{u}^h$.

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

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

satisfies

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

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. Notice 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

$$u_1(0) \approx -0.4161468365,$$

$$u_1^h(0) \approx -0.4131150562, \quad u_1^{h/2}(0) \approx -0.415389039,$$

$$\tilde{u}_1^h(0) = \tfrac{4}{3}u_1^{h/2}(0) - \tfrac{1}{3}u_1^h(0) =\approx 0.4161470333.$$

Such level of accuracy by $u_1^{\tilde{h}}(0)$ is achieved with $\tilde{h} = 0.0008$, 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 $\mathbb{R}^d$. 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.

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

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

For linear operators A and B we use the notation

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

Lemma 3.2.

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

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

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

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

Let $l \geq 0$ be an integer and $K \geq 0$ be a constant. In the next lemma $\mathcal{M}$ and $\mathcal{N}$ denote difference operators of the form $\mathcal{M} = \sum_{\lambda\in\Lambda_0} b^\lambda\delta_{h,\lambda}$ and $\mathcal{N} = \sum_{\lambda\in\Lambda_0} b^\lambda\delta_\lambda$, 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 multi-indices α, $|\alpha| \le l$, and functions $\varphi \in W_2^l$ on $\mathbb{R}^d$.*

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

$$(D^\alpha \mathcal{M}\varphi, D^\alpha\varphi) \le N\|\varphi\|^2_{W_2^l}. \tag{3.7}$$

(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 $h \ne 0$*

$$|(D^\alpha \mathcal{N}\varphi, D^\alpha\varphi)| \le N\|\varphi\|^2_{W_l^2}. \tag{3.8}$$

(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 $h \ne 0$*

$$|(D^\alpha \mathcal{N}\mathcal{N}\varphi, D^\alpha\varphi) + (D^\alpha\mathcal{N}\varphi, D^\alpha\mathcal{N}\varphi)| \le N\|\varphi\|^2_{W_l^2}. \tag{3.9}$$

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

Proof. To prove (i) notice that by (3.1)

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

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

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

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

$$\sum_{1\le|\gamma|,\gamma+\beta=\alpha}\sum_{\lambda\in\Lambda_0}|((D^\gamma b^\lambda)\delta_{h,\lambda}D^\beta\varphi, D^\alpha\varphi)| \le N\|\varphi\|^2_{W_2^l}.$$

Hence

$$(D^\alpha\mathcal{M}\varphi, D^\alpha\varphi) \le N\|\varphi\|^2_{W_2^l} + (\mathcal{M}D^\alpha\varphi, D^\alpha\varphi),$$

which yields (3.7), since by (3.10)

$$(\mathcal{M}D^\alpha\varphi, D^\alpha\varphi) \le N\|\varphi\|^2_{W_2^l}.$$

To prove (ii) notice notice that

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

with

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

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

$$\begin{aligned}|(\mathcal{N}\varphi,\varphi)| &\leq \frac{K}{4}\sum_{\lambda\in\Lambda_0}|\lambda|(\|T_{h,\lambda}\varphi\|_{L_2}+\|T_{h,-\lambda}\varphi\|_{L_2})\|\varphi\|_{L_2}\\ &= \frac{K}{2}\sum_{\lambda\in\Lambda_0}|\lambda|\|\varphi\|^2_{L_2}, \qquad (3.13)\end{aligned}$$

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

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

Hence

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

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

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

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

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

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

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

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

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

with a constant N depending only on Λ, l, d and K. Clearly,

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

For multi-indices γ, $|\gamma|\leq m$, set

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

and notice that for multi-indices $\beta\neq 0$, $\gamma\neq 0$, ρ, such that $\beta+\gamma+\rho=\alpha$ we have

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

Similarly, for multi-indices $\beta\neq 0$, $\gamma\neq 0$, $\bar\beta$ and $\bar\gamma$ such that $\beta+\bar\beta=\alpha$ and $\gamma+\bar\gamma=\alpha$ we have

$$(\mathcal{N}^{(\beta)}D^{\bar\beta}\varphi)\mathcal{N}^{(\gamma)}D^{\bar\gamma}\varphi\sim 0,$$

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

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

owing to

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

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

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

with

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

where

$$\beta+\gamma+\rho = \alpha, \quad \beta+\bar\beta = \alpha, \quad \gamma+\bar\gamma=\alpha.$$

By (3.9) with $\alpha = 0$

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

Using (3.14) we have

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

that for $\gamma\le\alpha$, $|\gamma| = 1$ yields

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

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

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

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

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

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

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

and similarly, for $\beta<\alpha$, $|\beta|\ge 2$ 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

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

Clearly,

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

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 $(\mathcal{N}^\rho)_{\rho=1}^\infty$ be a sequence of operators of the form $\mathcal{N}^\rho = \sum_{\lambda\in\Lambda_0} \mathfrak{b}^{\lambda\rho}\delta_\lambda$, 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 $l \geq 0$ be an integer. Then the following statements hold for all multi-indices α, $|\alpha| \leq l$ 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

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

(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

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

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

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

Lemma 3.4. *Let Assumption 2.5 hold. Then for multi-indices α, $|\alpha| \leq l$ we have*

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

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

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

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

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

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

with

$$\tilde{\mathcal{M}} = \frac{1}{2}\sum_{\lambda,\mu\in\Lambda_0} \mathfrak{b}^{\mu\rho}(\delta_\mu \mathfrak{b}^{\lambda\rho})(\delta_{\lambda+\mu} + \delta_{\lambda-\mu})$$

$$+\frac{1}{8}\sum_{\lambda,\mu\in\Lambda_0} ((\delta_{h,\mu} + \delta_{h,-\mu})\mathfrak{b}^{\lambda\rho})(T_{h,\lambda+\mu} - T_{h,\mu-\lambda} - T_{h,\lambda-\mu} + T_{h,-\lambda-\mu})$$

$$+\mathfrak{b}^{0\rho}\mathfrak{b}^{0\rho}+\sum_{\lambda\in\Lambda_0}\mathfrak{b}^{0\rho}\mathfrak{b}^{\lambda\rho}\delta_\lambda+\sum_{\mu\in\Lambda_0}\mathfrak{b}^{\mu\rho}\mathfrak{b}^{0\rho}\delta_\mu$$

$$+\frac{1}{2}\sum_{\mu\in\Lambda_0}\{\mathfrak{b}^{\mu\rho}(\delta_{h,\mu}\mathfrak{b}^{0\rho})T_{h,\mu}+\mathfrak{b}^{\mu\rho}(\delta_{h,-\mu}\mathfrak{b}^{0\rho})T_{h,-\mu}\},$$

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

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

and

$$|(D^\alpha\tilde{\mathcal{M}}\varphi,D^\alpha\varphi)|\le N\|\varphi\|_{W^l_2},\quad|(D^\alpha\varphi,D^\alpha(\mathfrak{a}^{0\lambda}+\mathfrak{a}^{\lambda0})\delta_\lambda\varphi)|\le N\|\varphi\|_{W^l_2},$$

$$(D^\alpha b^{0\rho}\varphi,D^\alpha\mathcal{M}^{h\rho}\varphi)\le N\|\varphi\|_{W^l_2}$$

for $h\neq0$. Here, and everywhere in this proof, N stands for constants depending only on l, K, d and Λ. Hence

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

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

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

with

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

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

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

By Lemma 3.3 (ii) and (iii) for $h\neq0$

$$|(D^\alpha\tilde{\mathcal{N}}\varphi,D^\alpha\varphi)|\le N\|\varphi\|^2_{W^l_2},$$

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

Hence

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

which along (3.22) finishes 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

$$\|f\|^2_{\mathbb{W}_2^m(T)}=\int_0^T\|f(t)\|^2_{W_2^m}\,dt,\quad \|g\|^2_{\mathbb{W}_2^m(T,l_2)}=\int_0^T\sum_{\rho=1}^\infty\|g^\rho(t)\|^2_{W_2^m}\,dt.$$

For short 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 $\mathcal{F}_0$-measurable random variable, $f\in\mathbb{W}_2^l(T)$ and $g\in\mathbb{W}_2^{l+1}(T,l_2)$. Then for each $h\neq 0$ 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 $h>0$*

$$E\sup_{t\le T}\|u_t^h\|^2_{W_2^l}\le NE\|u_0\|^2_{W_2^l}+NE\int_0^T\left(\|f_t\|^2_{W_2^l}+\|g_t\|^2_{W_2^{l+1}}\right)dt, \tag{3.23}$$

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 $h\neq 0$.

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 $h\neq 0$. 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 estimate (3.23) with a constant N which depends on h. In particular we have that the solution is in $\mathbb{W}_2^{0,l}(T)$. We assume $E\|u\|^2_{W_2^l}<\infty$, otherwise (3.23) is trivial. To prove (3.23) with a constant N independent of h, we take any multi-index α, $|\alpha|\le l$ and use Itô's formula for the L_2-valued process $D^\alpha u^h$ to find

$$d\|D^\alpha u_t^h\|^2_{L_2}$$

$$=\{\mathbb{Q}_t^\alpha(u_t^h)+2(D^\alpha u_t^h,D^\alpha f_t)+2(D^\alpha b^{\lambda\rho}\delta_\lambda u_t^h,D^\alpha g_t^\rho)+\sum_\rho\|D^\alpha g_t^\rho\|^2_{L_2}\}\,dt$$

$$+2(D^\alpha u_t^h,D^\alpha M^{h,\rho}u_t^h+D^\alpha g_t^\rho)\,dw_t^\rho, \tag{3.24}$$

where $\mathbb{Q}^\alpha$ is defined in Lemma 3.4. Clearly,

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

and by integration by parts

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

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

$$d\sum_{|\alpha|\le l}\|D^\alpha u^h_t\|^2_{L_2} \le N\big(\|u^h_t\|^2_{W^l_2}+\|f_t\|^2_{W^l_2}+\|g_t\|^2_{W^l_2}\big)\,dt + 2\sum_{|\alpha|\le l}(D^\alpha u^h_t, D^\alpha M^{h,\rho}u^h_t + D^\alpha g^\rho_t)\,dw^\rho_t \tag{3.25}$$

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

$$E\|u^h_t\|^2_{W^l_2} \le E\|u_0\|^2_{W^l_2} + NE\int_0^t \big(\|u^h_s\|^2_{W^l_2}+\|f_s\|^2_{W^l_2}+\|g_s\|^2_{W^{l+1}_2}\big)\,ds<\infty, \tag{3.26}$$

which by Gronwall's lemma yields

$$E\|u^h_t\|^2_{W^l_2} \le NE\|u_0\|^2_{W^l_2} + NE\int_0^t \big(\|f_s\|^2_{W^l_2}+\|g_s\|^2_{W^{l+1}_2}\big)\,ds \tag{3.27}$$

for $t\in[0,T]$. Now we return to (3.25) and use Davis's inequality to get

$$E\sup_{t\le T}\|u^h_t\|^2_{W^l_2} \le E\|u_0\|^2_{W^l_2} + NE\int_0^T \big(\|f_t\|^2_{W^l_2}+\|g_t\|^2_{W^{l+1}_2}\big)\,dt + N_1J, \tag{3.28}$$

where

$$J = E\Big(\int_0^T\sum_{\rho=1}^\infty\Big|\sum_{|\alpha|\le l}(D^\alpha u^h_t, D^\alpha M^{h,\rho}u^h_t + D^\alpha g^\rho_t)\Big|^2\,dt\Big)^{1/2}.$$

By the Cauchy-Bunyakovsky-Schwarz inequality

$$\sum_{\rho=1}^\infty\Big|\sum_{|\alpha|\le l}(D^\alpha u^h_t, D^\alpha g^\rho_t)\Big|^2 \le \|u^h_t\|^2_{W^l_2}\|g_t\|^2_{W^l_2},$$

and by Remark 3.1 (i)

$$\sum_{\rho=1}^\infty\Big|\sum_{|\alpha|\le l}(D^\alpha u^h_t, D^\alpha M^{h,\rho}u^h_t)\Big|^2 \le N\|u_t\|^4_{W^l_2}.$$

Hence

$$J\le J_1+J_2,$$

where

$$J_1 = E\Big(\int_0^T\sum_{\rho=1}^\infty\Big|\sum_{|\alpha|\le l}(D^\alpha u^h_t, D^\alpha M^{h,\rho}u^h_t)\Big|^2\,dt\Big)^{1/2} \le NE\Big(\int_0^T \|u^h_t\|^4_{W^l_2}\,dt\Big)^{1/2}$$

$$\le NE\left(\sup_{t\le T}\|u^h_t\|_{W^l_2}\Big(\int_0^T\|u^h_t\|^2_{W^l_2}\,dt\Big)^{1/2}\right),$$

$$\leq \frac{1}{4N_1} E \sup_{t\leq T} \|u_t^h\|^2_{W_2^l} + N_2 E \int_0^T \|u_t^h\|^2_{W_2^l}\, ds.$$

$$J_2 = E\big(\int_0^T \sum_{\rho=1}^{\infty} |\sum_{|\alpha|\leq l} (D^\alpha u_t^h, D^\alpha g_t^\rho))|^2\, dt\big)^{1/2} \leq NE\big(\int_0^T \|u_t^h\|^2_{W_2^l} \|g_t\|^2_{W_2^l}\, dt\big)^{1/2}$$

$$\leq E\left(\sup_{t\leq T} \|u_t^h\|_{W_2^l} \big(\int_0^T \|g_t^h\|^2_{W_2^l}\, dt\big)^{1/2}\right),$$

$$\leq \frac{1}{4N_1} E \sup_{t\leq T} \|u_t^h\|^2_{W_2^l} + N_2 E \int_0^T \|g_t\|^2_{W_2^l}\, ds.$$

Thus from (3.28) we get

$$E \sup_{t\leq T} \|u_t^h\|^2_{W_2^l} \leq E\|u_0\|^2_{W_2^l} + \frac{1}{2} E \sup_{t\leq T} \|u_t^h\|^2_{W_2^l}$$

$$+NE \int_0^T \|f_t\|^2_{W_2^l} + \|g_t\|^2_{W_2^{l+1}}\big)\, ds,$$

which proves (3.23). □

Lemma 3.6. *Let $n \geq 0$ be an integer, let $\phi \in W_2^{n+1}$, $\psi \in W_2^{n+2}$, 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 we have

$$\frac{\partial^n}{(\partial h)^n} \delta_{h,\lambda} \phi(x) = \int_0^1 \theta^n \partial_\lambda^{n+1} \phi(x + h\theta\lambda)\, d\theta, \tag{3.29}$$

$$\frac{\partial^n}{(\partial h)^n} \delta_\lambda^h \phi(x) = \frac{1}{2} \int_{-1}^1 \theta^n \partial_\lambda^{n+1} \phi(x + h\theta\lambda)\, d\theta, \tag{3.30}$$

$$\frac{\partial^n}{(\partial h)^n} \delta_\lambda \delta_\mu \psi(x)$$

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

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

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

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

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 $l \geq 0$ is an integer and $\phi \in W_2^{n+2+l}$ and $\psi \in W_2^{n+3+l}$, then

$$\big\|\delta_{h,\lambda}\phi - \sum_{i=0}^{n} \frac{h^i}{(i+1)!}\partial_\lambda^{i+1}\phi\big\|_{W_2^l} \leq \frac{|h|^{n+1}}{(n+2)!}\|\partial_\lambda^{n+2}\phi\|_{W_2^l} \tag{3.35}$$

$$\big\|\delta_\lambda^h\phi - \sum_{i=0}^{n} \frac{h^i}{(i+1)!}B_i\partial_\lambda^{i+1}\phi\big\|_{W_2^l} \leq \frac{|h|^{n+1}}{(n+2)!}\|\partial_\lambda^{n+2}\phi\|_{W_2^l} \tag{3.36}$$

$$\big\|\delta_\lambda^h\delta_\mu^h\psi - \sum_{i=0}^{n} h^i \sum_{r=0}^{i} A_{i,r}\partial_\lambda^{r+1}\partial_\mu^{i-r+1}\psi\big\|_{W_2^l} \leq N|h|^{n+1}\|\psi\|_{W_2^{l+n+3}}, \tag{3.37}$$

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 $n \geq 1$ one obtains (3.29)–(3.31) by differentiating both parts of these equations written with $n = 1$.

Next by Taylor's 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

$$\delta_{h,\lambda}\phi(x) = \sum_{i=0}^{n} \frac{h^i}{(i+1)!}\partial_\lambda^{i+1}\phi(x)$$

$$+\frac{h^{n+1}}{n!}\int_0^1\int_0^1 (1-\theta_2)^n\theta_1^{n+1}\partial_\lambda^{n+2}\phi(x + h\theta_1\theta_2\lambda)\,d\theta_1 d\theta_2, \tag{3.38}$$

$$\delta_\lambda^h\phi(x) = \sum_{i=0}^{n} \frac{h^i}{(i+1)!}B_i\partial_\lambda^{i+1}\phi(x)$$

$$+\frac{h^{n+1}}{2n!}\int_0^1\int_{-1}^1 (1-\theta_2)^n\theta_1^{n+1}\partial_\lambda^{n+2}\phi(x + h\theta_1\theta_2\lambda)\,d\theta_1 d\theta_2. \tag{3.39}$$

Hence we get (3.35) and (3.36) by noting that by Minkowski's 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 Minkowski's inequality. □

Remark 3.2. Formula (3.29) with $n=1$ and Minkowski's inequality imply that

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

By 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 $l\ge 0$ and $r\ge 1$ 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

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

We also set $W_{h,2}^{l,0}=W_2^l$. Then for any $\phi\in W_2^{l+r}$ we have

$$\|\varphi\|_{l,h,r}\le N\|\varphi\|_{W_2^{l+r}},$$

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

Set

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

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

$$\mathcal{L}_t^{(n)}=\sum_{\lambda,\mu\in\Lambda_0}\mathfrak{a}_t^{\lambda\mu}\sum_{r=0}^n A_{n,r}\partial_\lambda^{r+1}\partial_\mu^{n-r+1}+(n+1)^{-1}\sum_{\lambda\in\Lambda_0}(\mathfrak{a}_t^{\lambda0}+\mathfrak{a}_t^{0\lambda})B_n\partial_\lambda^{n+1}$$

$$+(n+1)^{-1}\sum_{\lambda\in\Lambda_0}(\mathfrak{p}_t^\lambda+(-1)^{n+1}\mathfrak{q}_t)\partial_\lambda^{n+1},$$

$$\mathcal{M}_t^{(n)\rho}=(n+1)^{-1}\sum_{\lambda\in\Lambda_0}\mathfrak{b}_t^{\lambda\rho}B_n\partial_\lambda^{n+1},$$

$$\mathcal{O}_t^{h(n)}=L_t^h-\sum_{i=0}^n\frac{h^i}{i!}\mathcal{L}_t^{(i)},\quad \mathcal{R}_t^{h(n)\rho}=M_t^{h,\rho}-\sum_{i=0}^n\frac{h^i}{i!}\mathcal{M}_t^{(i)\rho},$$

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

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

Remark 3.4. Owing to Assumption 2.4 we have

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

Notice also that by (3.35)-(3.37) under Assumptions 2.2 and 2.6 with $\mathfrak{m} = m$, for $\phi \in W_2^{n+2+l}$ and $\psi \in W_2^{n+3+l}$ we have for $l \leq m$

$$\|\mathcal{O}_t^{h(n)}\psi\|_{W_2^l} \leq N|h|^{n+1}\|\psi\|_{W_2^{l+n+3}},$$
$$\|\mathcal{R}_t^{h(n)}\phi\|_{W_2^l} \leq N|h|^{n+1}\|\phi\|_{W_2^{l+n+2}}, \tag{3.42}$$

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

Let $k \in [1, m/2]$ be an integer. The functions $u^{(1)}$,..., $u^{(k)}$ we need in 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 $v_t^{(1)}$,...,$v_t^{(k)}$ 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 PDEs:

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

Theorem 3.7. *Let Assumptions 2.1, 2.2, 2.3 and 2.6 hold with $\mathfrak{m} = m \geq 2k$. 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 $n = 1, ..., k$*

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

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 $n \leq k$.

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 $n \leq k$. Denote

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

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

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

which by Theorem 2.1 implies that there exists a unique $v^{(1)} \in \mathbb{H}^{m-2}(T)$ satisfying (3.43) with zero initial condition. Furthermore, $v^{(1)}_t$ is a continuous W^{m-3}_2-valued function (a.s.) and (3.44) holds with $n = 1$. Passing to higher n we assume that $m \geq k \geq 2$ and for an $n \in \{2, ..., k\}$ we have found $v^{(1)}$,...,$v^{(n-1)}$ with the asserted properties. Then $\mathcal{M}^{(1)} v^{(n-1)} = 0$ and

$$\|\mathcal{L}^{(1)}_t v^{(n-1)}_t\|_{W^{m-2n}_2} \leq N\|v^{(n-1)}_t\|_{W^{m-2n+2}_2} = N\|v^{(n-1)}_t\|_{W^{m-2(n-1)}_2},$$

and for $i \geq 2$

$$\|\mathcal{L}^{(i)}_t v^{(n-i)}_t\|_{W^{m-2n}_2} \leq N\|v^{(n-i)}_t\|_{W^{m-2n+(i+2)}_2} \leq N\|v^{(n-i)}_t\|_{W^{m-2(n-i)}_2},$$

$$\sum_{k=1}^{\infty} \|\mathcal{M}^{(i)\rho} v^{(n-i)}\|^2_{W^{m-2n+1}_2} \leq N\|v^{(n-i)}\|^2_{W^{m-2n+1+(i+1)}_2}$$
$$\leq N\|v^{(n-i)}_t\|^2_{W^{m-2(n-i)}_2}.$$

It follows by the induction hypothesis that

$$E\int_0^T \|S^{(n)}_t\|^2_{W^{m-2n}_2}\,dt \leq N\mathcal{K}^2_m, \quad E\int_0^T \|R^{(n)}_t\|^2_{W^{m-2n+1}_2}\,dt \leq N\mathcal{K}^2_m, \quad (3.45)$$

which by Theorem 2.1 yields that there exists 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 $v^{(n)}_t$ 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.

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

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 $k \geq 0$, $l \geq 0$. Set

$$r^h_t = v^h_t - \sum_{j=1}^{k} \frac{h^j}{j!} v^{(j)}_t, \quad (3.46)$$

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 $(v^{(n)})^k_{n=1}$ is the solution of* (3.43)*, given by Theorem 3.7. Then $r^h_0 = 0$, $r^h \in \mathbb{W}^{0,l}_2(T)$, and*

$$dr^h_t = (L^h_t r^h_t + F^h_t)\,dt + (M^{h,\rho}_t r^h_t + G^{h,\rho}_t)\,dw^\rho_t, \quad (3.47)$$

where

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

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 $j \leq k$ 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 $\hat{F}$ and $\hat{G}$ in place of F and G, respectively, where

$$\hat{F}^h = L^h v^{(0)} - \mathcal{L} v^{(0)} + \sum_{1 \leq j \leq k} L^h v^{(j)} \frac{h^j}{j!} - \sum_{1 \leq j \leq k} \mathcal{L} v^{(j)} \frac{h^j}{j!} - I^h,$$

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

with

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

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

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

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

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

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

and similarly,

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

After that the fact that $\hat{F} = F$ and $\hat{G} = G$ follows by simple arithmetics. To prove the last assertion notice that for $j = 0, 1, ..., k$

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

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

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

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

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

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

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 $j \leq k$, and if k is odd then it is sufficient to assume $m \geq l + 2k + 2$ in place of $m = l + 2k + 3$ in* (4.1) *to have 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

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

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

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

and by Remark 3.4 we have

$$\begin{aligned} \|\mathcal{O}_t^{h(k-j)} v_t^{(j)}\|_{W_2^l} &\leq N|h|^{k-j+1}\|v_t^{(j)}\|_{W_2^{l+k-j+3}} \leq N|h|^{k-j+1}\|v_t^{(j)}\|_{W_2^{m-2j}}, \\ \|\mathcal{R}_t^{h(k-j)} v_t^{(j)}\|_{W_2^l} &\leq N|h|^{k-j+1}\|v^{(j)}\|_{W_2^{l+k-j+2}} \leq N|h|^{k-j+1}\|v^{(j)}\|_{W_2^{m-2j-1}}. \end{aligned} \tag{4.5}$$

Hence, using Theorem 3.7 we see that

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

which by virtue of Theorem 3.5 implies estimate (4.2). If $p^\lambda = q^\lambda = 0$ for $\lambda \in \Lambda_0$ then by Theorem 3.7 we know that $v^{(j)} = 0$ for odd $j \leq k$. (Remark that this 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 $j \leq k - 1$ we need only $m = l + 2k + 2$. □

By Sobolev's theorem on embedding 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

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

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

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

where N is a constant depending only on d.

Proof. This lemma is a straightforward consequence of Sobolev's theorem on embedding 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. (See, e.g., [6].) □

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

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

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

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. Set $j = n - l$. Then $j > d/2$ and by Sobolev's theorem on embedding W_2^j into C_b and by Remark 3.2, from Theorem 4.1 we get

$$E \sup_{t\in[0,T]} \sup_{x\in\mathbb{R}^d} |\delta_{h,\lambda} R_t^h(x)|^2 \le C_1 E \sup_{t\in[0,T]} \|R_t^h\|^2_{j,h,n}$$

$$\le C_2 E \sup_{t\in[0,T]} \|R_t^h\|^2_{W_2^l} \le N h^{2(k+1)} \mathcal{K}_m^2.$$

Similarly, by Lemma 4.2 and Remark 3.2

$$E \sup_{t\in[0,T]} \sum_{x\in\mathbb{G}_h} |\delta_{h,\lambda} R_t^h(x)|^2 h^d \le C_1 E \sup_{t\in[0,T]} \|\delta_{h,\lambda} R_t^h\|^2_{W_2^j}$$

$$\le C_2 E \sup_{t\in[0,T]} \|R_t^h\|^2_{W_2^l} \le N h^{2(k+1)} \mathcal{K}_m^2.$$

□

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

$$\hat{u}^h = Iv^h, \quad u^{(j)} = Iv^{(j)}, \quad j = 0, ..., k,$$

where v^h is the unique $\mathcal{F}_t$-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 virtue of 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 $\hat{u}^h$ 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

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

for each $x \in \mathbb{G}_h$, where u^h is the unique $\mathcal{F}_t$-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 $\hat{u}^h$ is a continuous L_2-valued solution of (2.4), for each ε almost surely

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

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

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

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 $\hat{u}_t$ 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 $\mathcal{F}_t$-adapted solution of (2.4) for any l_2-valued $\mathcal{F}_0$-measurable initial condition, we have (4.7), that finishes the proof Theorem 2.4.

References

[1] C. Brezinski, *Convergence acceleration during the 20th century*, Journal of Computational and Applied Mathematics, Vol. 122 (2000), 1–21.

[2] P.G. Ciarlet and J.L. Lions (eds), Handbook of Numerical Analysis, Elsevier, 2003.

[3] Hongjie 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., Vol. 2005(2005), No. 102, pp. 1-25. http://ejde.math.txstate.edu

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

[5] 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.

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

[7] N.V. Krylov and B.L. Rozovskii, *Characteristics of degenerating second-order parabolic Itô equations*, Trudy Seminara imeni I. G. Petrovskogo, N.8 pp. 153-168, 1982.

[8] G.I. Marschuk and V.V. Shaidurov, Difference methods and their extrapolations, New York Berlin Heidelberg, Springer 1983.

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

[10] L.F. Richardson and J.A. Gaunt, *The Deferred Approach to the Limit*, Phil. Trans. Roy. Soc. London Ser. A, Vol. 226 (1927), 299-361.

[11] B.L. Rozovskii, Stochastic evolution systems. Linear theory and applications to nonlinear filtering. Kluwer, Dordrecht, 1990.

[12] Yoo, H.: An analytic approach to stochastic partial differential equations and its applications. Thesis, University of Minnesota (1998).

School of Mathematics and Maxwell Institute, University of Edinburgh, King's Buildings, Edinburgh, EH9 3JZ, United Kingdom

E-mail address: gyongy@maths.ed.ac.uk