Accelerated Finite Difference Schemes for Linear Stochastic Partial Differential Equations in the Whole Space

Citation for published version:
Gyongy, I & Krylov, N 2010, 'Accelerated Finite Difference Schemes for Linear Stochastic Partial Differential Equations in the Whole Space' SIAM Journal on Mathematical Analysis, vol 42, no. 5, pp. 2275-2296. DOI: 10.1137/090781395

Digital Object Identifier (DOI):
10.1137/090781395

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
SIAM Journal on Mathematical Analysis

ACCELERATED FINITE DIFFERENCE SCHEMES FOR LINEAR STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS IN THE WHOLE SPACE

ISTVÁN GYÖNGY* AND NICOLAI KRYLOV †

Abstract. We give sufficient conditions under which the convergence of finite difference approximations in the space variable of the solution to the Cauchy problem for linear stochastic PDEs of parabolic type can be accelerated to any given order of convergence by Richardson's method.

Key words. Cauchy problem, finite differences, extrapolation to the limit, Richardson's method, linear SPDEs

AMS subject classifications. 65M06, 60H15, 65B05

1. Introduction. Stochastic partial differential equations (SPDEs) play important roles in many applied fields. Here we consider linear second order nondegenerate parabolic SPDEs. These equations arise, for example, in nonlinear filtering of partially observable diffusion processes. There are various methods developed in the literature to solve them numerically. In this paper we apply the method of finite differences in the space variable, while the time variable changes continuously. It is known that in general the error of the finite difference approximations in the space variable is proportional to the parameter h of the finite difference, see, e.g., [22] or [23]. Our aim is to show that the convergence of these approximations can be accelerated by an implementation of Richardson's idea to SPDEs. We prove that for linear parabolic stochastic PDEs driven by Wiener processes the finite difference approximations u^h admit power series expansions in the parameter h. This is Theorem 2.2, one of the main results of the paper. Hence we get Theorem 2.4, our first result on acceleration of finite difference schemes for SPDEs. It says that if the coefficients and the data are sufficiently regular then the convergence of finite difference approximations can be accelerated to any high order by taking appropriate mixture of approximations with different step sizes. In the special case of symmetric finite difference schemes, Example 2.2 below, the coefficients of odd powers in the expansions vanish. Hence it follows, see Theorem 2.5, that the error of symmetric finite difference schemes is proportional to h^2 without acceleration, and we can accelerate more effectively.

The SPDEs we consider in this paper are given in $[0,T]\times\mathbb{R}^d$. The finite difference schemes are given in $[0,T]\times\mathbb{G}_h$, where $\mathbb{G}_h$ are grids in the space variable. The supremum in $t\in[0,T]$ and $x\in\mathbb{G}_h$ of the remainder terms and of the approximation errors in the expansions in Theorem 2.2 and Theorems 2.4-2.5, respectively are estimated. To prove these results we consider the finite difference schemes given not only on the grids, but on the whole $\mathbb{R}^d$, and obtain a more general theorem, Theorem 4.1, that establishes a power expansion in h for the L_2-solutions of the schemes on $[0,T]\times\mathbb{R}^d$, with the remainder estimated in terms of Sobolev norms in the whole $\mathbb{R}^d$. Hence we estimate the sup norm and also discrete Sobolev norms of the remainder by Sobolev's embedding theorems, and get our theorems on accelerated finite difference

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

†School of Mathematics, 127 Vincent Hall, University of Minnesota, Minneapolis, MN, 55455 (krylov@math.umn.edu). This author's research was partially supported by NSF grant DMS-0653121.

schemes, formulated in terms of supremum norm and also in discrete Sobolev norms of functions over $\mathbb{G}_h$.

In the special case when the stochastic terms in the equations vanish the above mentioned theorems are results on accelerated finite difference schemes for deterministic PDEs. Similar results on *monotone* finite difference schemes for parabolic and elliptic PDEs, which may degenerate, are proved in [9] on the basis of derivative estimates in the supremum norm obtained in [7]-[8] for solutions to monotone finite difference schemes. The finite difference schemes in the present article are not necessarily monotone.

The idea of accelerating the convergence of finite difference approximations to deterministic PDEs by suitable mixtures of approximations with different step-sizes is due to L.F. Richardson, see [18] and [19]. This method is often called *Richardson's method* or *extrapolation to the limit*, and is applied to various types of approximations. It is used in [5]-[6] to accelerate splitting up approximations for a large class of deterministic evolution equations, including second order parabolic equations and symmetric hyperbolic system of first order PDEs. Richardson's idea is implemented to the law of Euler's approximations for stochastic differential equations in [21], [1] and [15]. There is a lot of other applications of Richardson's method. The reader is referred to the survey papers [2] and [4] for a review on the method, and to textbooks (for instance, [16] and [17]) concerning finite difference methods and their accelerations. We note that previous extrapolation results for stochastic equations, i.e. in [21], and in its generalizations [15] and [12], are concerned with *week approximations* of stochastic differential equations. In contrast our main results are error expansions for *strong convergence* of finite difference approximations in the space variable for stochastic parabolic equations, and as far as we know these are the first results in this direction.

In light of the results of the present paper it is natural to look for accelerated space *and time* discretized schemes, say by using time discretization to solve the systems of ordinary stochastic equations which we obtain after discretizing the space. However, one knows that if the values of the driving multidimensional Wiener process are available only at the grid points, then in general one cannot construct a scheme with (strong) rate of convergence better than $\sqrt{\tau}$, where τ is the mesh-size of the time grid. On the other hand, in some particular cases, e.g., when the Wiener process is one-dimenional, or some special data, like iterated stochastic integrals of the components of the Wiener processes are available, then one can have accelerated fully discretised numerical schemes for SPDEs. (See, e.g., [11] for high order strong approximations of stochastic differential equations when appropriate iterated stochastic integrals of the Wiener processes are used in the numerical schemes.)

We did not try to make our results as sharp or as general as possible. The main goal of the article is to show a method of approximating. We plan to extend our results to the case of degenerate parabolic SPDEs in the continuation of this paper.

In conclusion we introduce some notation used everywhere below. Throughout the paper $\mathbb{R}^d$ is a Euclidean space of points $x = (x^1, ..., x^d)$, and $T > 0$ is a fixed finite constant. We set

$$D_i = \partial/\partial x^i, \quad i = 1, \dots, d.$$

Also let D_0 be the unit operator. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, $\mathcal{F}_t$, $t \geq 0$, be an increasing family of sub σ-fields of $\mathcal{F}$, such that $\mathcal{F}_0$ is complete with respect to $(\mathcal{F}, P)$. By $\mathcal{P}$ we denote the σ-field of predictable subsets of $\Omega \times$

$[0,\infty)$ generated by $\mathcal{F}_t$, and $\mathcal{B}(\mathbb{R}^d)$ is the σ-algebra of the Borel subsets of $\mathbb{R}^d$. We assume that on Ω we are given a sequence of $\mathcal{F}_t$-adapted independent Wiener processes $\{w^\rho\}_{\rho=1}^\infty$ such that for every integer $\rho \geq 1$ and for all $0 \leq s \leq t$ the increments $w_t^\rho - w_s^\rho$ are independent of $\mathcal{F}_s$. Unless otherwise stated throughout the paper we use the summation convention over repeated integer valued indices. For functions $u = u(\omega,t,x)$, $\omega \in \Omega$, $t \in [0,T]$, $x \in \mathbb{R}^d$, we use the notation $D^l u = D^l u(x)$ for the collection of lth order derivatives of u with respect to x and $|D^l u|^2 = |D^l u(x)|^2$ is the sum of squares of all lth order derivatives at x. If u is an l_2-valued function then the differentiability of it is understood in the sense of l_2-valued functions and $|D^l u(x)|_{l_2}^2$ means the the sum of squares of the l_2-norm of all lth order derivatives at x. For basic notions and notation concerning the theory of linear stochastic partial differential equations we refer to [20].

Acknowledgments. The first version of this paper was presented at the ICMS conference on 'Numerical Analysis of Stochastic PDEs' (Edinburgh, May 2009), organised by Evelyn Buckwar and Gabriel Lord, and at the 'Stochastic Analysis' session of the 7th International ISAAC Congress (Imperial College, London, August 2009) organized by Dan Crisan and Terry Lyons. We thank the organisers for the invitations.

We are sincerely grateful to the referees for their careful work which helped improve the presentation of 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 $(t,x) \in [0,T] \times \mathbb{R}^d =: H_T$ with some initial condition where

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

and $\{w^\rho\}_{\rho=1}^\infty$ is a sequence of independent Wiener processes given on a probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $(\mathcal{F})_{t\geq 0}$ such that w_t^ρ is $\mathcal{F}_t$-measurable and $w_t^\rho - w_s^\rho$ is independent of $\mathcal{F}_s$ for all $0 \leq s \leq t$ and integers $\rho \geq 1$. Here and below the summation with respect to α and β is performed over the set $\{0,1,...,d\}$ and with respect to ρ in the range $\{1,2,...\}$. Assume that, for $\alpha,\beta \in \{0,1,...,d\}$, we have $a_t^{\alpha\beta} = a_t^{\beta\alpha}$ and $a_t^{\alpha\beta} = a_t^{\alpha\beta}(x)$ are real-valued and $b_t^\alpha = (b_t^{\alpha\rho})_{\rho=1}^\infty$ are l_2-valued $\mathcal{P} \times \mathcal{B}(\mathbb{R}^d)$-measurable functions on $\Omega \times H_T$.

Let $m \geq 1$ 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 2.1. (i) For each (ω,t) the functions $a_t^{\alpha\beta}$ are m times and the functions b_t^α are $m+1$ times continuously differentiable in x. There exist constants K_l, $l = 0,...,m+1$, such that for all values of indices and arguments we have

$$|D^l a_t^{\alpha\beta}| \leq K_l, \quad l \leq m, \quad |D^l b_t^\alpha|_{l_2} \leq K_l, \quad l \leq m+1.$$

(ii) There is a constant $\kappa > 0$ such that for all $(\omega,t,x) \in \Omega \times H_T$ and $z \in \mathbb{R}^d$

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

Assumption 2.2. We have $u_0 \in L_2(\Omega, \mathcal{F}_0, W_2^{m+1})$. The function f_t is W_2^m-valued, g_t^ρ, $\rho = 1, 2, ...$, are W_2^{m+1}-valued functions given on $\Omega \times [0, T]$ and they are predictable. 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

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

Remark 2.1. If Assumption 2.2 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 2.2 holds with $m > d/2$, we always assume that u_0, f_t and g_t are continuous in x for all t.

The solutions of (2.1) will be looked for in the Hilbert space

$$\mathbb{W}_2^{m+2}(T) = L_2(\Omega \times [0, T], \mathcal{P}, W_2^{m+2}).$$

One knows, see e.g., [14] or [20], how to define stochastic integrals of Hilbert-space valued processes and equation (2.1) is understood accordingly. Observe that since $u_0 \in L_2(\Omega, \mathcal{F}_0, W_2^m)$ the solutions of (2.1) automatically are continuous W_2^m-valued processes (a.s.).

We are going to use the following classical result (see, for instance, Theorem 5.1, Remark 5.6, and Theorem 7.1 of [13]).

THEOREM 2.1. *Under the above assumptions there exists a unique solution $u \in \mathbb{W}_2^{m+2}(T)$ of (2.1) with initial condition u_0. Furthermore, with probability one the function u_t is a continuous W_2^{m+1}-valued function and there exists a constant N depending only on T, d, κ, m, and $K_l, l \le m+1$, such that*

$$E\sup_{t\le T}\|u_t\|_{W_2^{m+1}}^2 + E\int_0^T \|u_t\|_{W_2^{m+2}}^2\, dt \le N\mathcal{K}_m^2.$$

Remark 2.2. In the future we are going to assume that $m+1 > d/2$. 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 $\lambda = 0$ set $\delta_{h,\lambda}$ to be the unit operator and for the other values of $\lambda \in \mathbb{R}^d$ let

$$\delta_{h,\lambda} u(x) = \frac{u(x + h\lambda) - u(x)}{h} \quad \text{for } h \in \mathbb{R} \setminus \{0\}.$$

We draw the reader's attention to the fact that h can be of any sign. This will be important in the future.

To introduce difference equations we take a finite set $\Lambda\subset\mathbb{R}^d$ containing the origin, and consider the equation

$$du^h_t=(L^h_tu^h_t+f_t)\,dt+(M^{h,\rho}_tu^h_t+g^\rho_t)\,dw^\rho_t, \tag{2.2}$$

with

$$L^h_t\phi=\mathfrak{a}^{\lambda\mu}_t\delta_{h,\lambda}\delta_{-h,\mu}\phi,\quad M^{h,\rho}_t\phi=\mathfrak{b}^{\lambda\rho}_t\delta_{h,\lambda}\phi,$$

where the summation is performed over $\lambda,\mu\in\Lambda$ and in (2.2) also with respect to $\rho=1,2,\dots$. Assume that, for $\lambda,\mu\in\Lambda$, $\mathfrak{a}^{\lambda\mu}=\mathfrak{a}^{\lambda\mu}_t(x)$ are real-valued and $\mathfrak{b}^\lambda=\mathfrak{b}^\lambda_t(x)=(\mathfrak{b}^{\lambda\rho}_t(x))_{\rho=1}^\infty$ are l_2-valued functions on $\Omega\times H_T$, measurable with respect to $\mathcal{P}\times\mathcal{B}(\mathbb{R}^d)$.

Set $\Lambda_0:=\Lambda\setminus\{0\}$. Let $\mathfrak{m}\geq 0$ be an integer. Set $\bar{\mathfrak{m}}=\max(\mathfrak{m},1)$, and let A_0, A_1,..., $A_{\bar{\mathfrak{m}}}$ be some constants. The functions $\mathfrak{a}$ and $\mathfrak{b}$ are supposed to possess the following properties.

Assumption 2.3. (i) For each (ω,t) and $\lambda,\mu\in\Lambda_0$ and $\nu\in\Lambda$, $\mathfrak{a}^{\lambda\mu}_t$ are $\bar{\mathfrak{m}}$ times continuously differentiable in x, $\mathfrak{a}^{0\nu}_t$, $\mathfrak{a}^{\nu 0}_t$ are $\mathfrak{m}$ times continuously differentiable in x and $\mathfrak{b}^\nu_t$ are $\mathfrak{m}$ times continuously differentiable in x as l_2-valued functions. For all values of arguments we have

$$|D^j\mathfrak{a}^{\lambda\mu}_t|\leq A_j,\quad \lambda,\mu\in\Lambda_0,\quad j\leq\bar{\mathfrak{m}},$$

$$|D^j\mathfrak{a}^{\lambda 0}_t|\leq A_j,\quad |D^j\mathfrak{a}^{0\lambda}_t|\leq A_j,\quad |D^j\mathfrak{b}^\lambda_t|_{l_2}\leq A_j,\quad \lambda\in\Lambda,\quad j\leq\mathfrak{m}.$$

(ii) For all $(\omega,t,x)\in\Omega\times H_T$ and numbers $z_\lambda,\lambda\in\Lambda_0$, we have

$$\sum_{\lambda,\mu\in\Lambda_0}(2\mathfrak{a}^{\lambda\mu}_t-\mathfrak{b}^{\lambda\rho}_t\mathfrak{b}^{\mu\rho}_t)z_\lambda z_\mu\geq\kappa\sum_{\lambda\in\Lambda_0}z_\lambda^2.$$

Introduce

$$\mathbb{G}_h=\{\lambda_1h+...+\lambda_nh:n=1,2,...,\lambda_i\in\Lambda\cup(-\Lambda)\}$$

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. Observe that, under Assumption 2.3 (i), equation (2.2) is an ordinary Itô equation with Lipschitz continuous coefficients for $l_2(\mathbb{G}_h)$-valued processes. 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 2.3 (i) holds then equation (2.2) 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.).

For equation (2.2) to be consistent with (2.1) we impose the following.

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

$$\sum_{\lambda,\mu\in\Lambda_0} \mathfrak{a}_t^{\lambda\mu}\lambda^i\mu^j = a_t^{ij}, \quad \sum_{\lambda\in\Lambda_0} \mathfrak{a}_t^{\lambda 0}\lambda^i + \sum_{\mu\in\Lambda_0} \mathfrak{a}_t^{0\mu}\mu^i = a_t^{i0} + a_t^{0i}, \quad \mathfrak{a}_t^{00} = a_t^{00},$$

$$\sum_{\lambda\in\Lambda_0} \mathfrak{b}_t^{\lambda\rho}\lambda^i = b_t^{i\rho}, \quad \mathfrak{b}_t^{0\rho} = b_t^{0\rho}.$$

Remark 2.4. Clearly, if

$$a_t^{ij} = \sum_{\lambda,\mu\in\Lambda_0} \mathfrak{a}_t^{\lambda\mu}\lambda^i\mu^j, \quad i, j = 1, ..., d$$

is an invertible matrix for some ω, t, x, then Λ_0 spans the whole $\mathbb{R}^d$. On the other hand, if Λ_0 spans $\mathbb{R}^d$, then clearly a constant $\kappa' > 0$ exists such that

$$\sum_{\lambda\in\Lambda_0} |\sum_i z^i\lambda^i|^2 \geq \kappa'|z|^2, \quad \text{for all } z = (z^1, ..., z^d) \in \mathbb{R}^d,$$

and therefore Assumptions 2.3 (ii) and 2.4 imply Assumption 2.1 (ii). It is not hard to see that Assumptions 2.1 (ii) and 2.4 do not imply Assumption 2.3 (ii), in general, unless Λ_0 is a basis in $\mathbb{R}^d$.

There are several ways to construct appropriate $\mathfrak{a}$ and $\mathfrak{b}$.

Example 2.1. *The most natural, albeit sometimes not optimal, way to choose* $\mathfrak{a}$ *and* $\mathfrak{b}$ *is to set* $\Lambda = \{e_0, e_1, ..., e_d\}$, *where* $e_0 = 0$ *and* e_i *is the* i*th basis vector in* $\mathbb{R}^d$ *and let*

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

Thus, in (2.2) *the first order derivatives in* (2.1) *are approximated by usual finite differences and*

$$\sum_{\lambda,\mu\in\Lambda_0} \mathfrak{a}_t^{\lambda\mu}\delta_{h,\lambda}\delta_{-h,\mu}u = -a_t^{ij}\delta_{h,e_i}\delta_{h,-e_j}u, \tag{2.3}$$

which is a standard finite-difference approximation of $a_t^{ij}D_iD_ju$. *Also notice that*

$$\sum_{\lambda,\mu\in\Lambda_0} \mathfrak{a}_t^{\lambda\mu}z_\lambda z_\mu = a_t^{ij}z_{e_i}z_{e_j}, \quad \sum_{\lambda\in\Lambda_0} \mathfrak{b}_t^{\lambda\rho}z_\lambda = b_t^{i\rho}z_{e_i}.$$

It follows that $\mathfrak{a}$ *and* $\mathfrak{b}$ *satisfy the above assumptions as long as* a *and* b *do.*

Example 2.2. *The second choice is to use symmetric finite differences to approximate the first-order derivatives. Namely, we take* $\Lambda_0 = \{\pm e_1, ..., \pm e_d\}$ *and*

$$\mathfrak{a}_t^{0,\pm e_i} = \mathfrak{a}_t^{\pm e_i,0} = \pm(1/4)(a_t^{0i} + a_t^{i0}), \quad \mathfrak{b}_t^{\pm e_i,\rho} = \pm(1/2)b_t^{i,\rho},$$

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

so that, for instance,

$$\sum_{\lambda\in\Lambda_0} \mathfrak{b}_t^{\lambda\rho}\delta_{h,\lambda}u(x) = \sum_{i=1}^d b_t^{i\rho}\frac{u(x + he_i) - u(x - he_i)}{2h}.$$

For $\lambda,\mu\in\Lambda_0$ we define $\mathfrak{a}_t^{\lambda\mu}$ by

$$\mathfrak{a}_t^{\pm e_i,\pm e_j}=\tfrac{1}{2}a_t^{ij},\quad \mathfrak{a}_t^{\pm e_i,\mp e_j}=0.$$

Then Assumption 2.4 is satisfied and formula (2.3) *holds again ($a^{ij}=a^{ji}$). If Assumption 2.1 (ii) is satisfied, then for any numbers z_λ*

$$\sum_{\lambda,\mu\in\Lambda_0}(2\mathfrak{a}_t^{\lambda\mu}-\mathfrak{b}_t^{\lambda\rho}\mathfrak{b}_t^{\mu\rho})z_\lambda z_\mu=\sum_{i,j=1}^d a_t^{ij}z_{e_i}z_{e_j}+\sum_{i,j=1}^d a_t^{ij}z_{-e_i}z_{-e_j}$$

$$-(1/4)\sum_\rho\big|\sum_{i=1}^d b_t^{i\rho}z_{e_i}-\sum_{i=1}^d b_t^{i\rho}z_{-e_i}\big|^2\geq\sum_{i,j=1}^d a_t^{ij}z_{e_i}z_{e_j}-(1/2)\sum_\rho\big|\sum_{i=1}^d b_t^{i\rho}z_{e_i}\big|^2$$

$$+\sum_{i,j=1}^d a_t^{ij}z_{-e_i}z_{-e_j}-(1/2)\sum_\rho\big|\sum_{i=1}^d b_t^{i\rho}z_{-e_i}\big|^2$$

$$\geq\kappa\sum_{i=1}^d z_{e_i}^2+\kappa\sum_{i=1}^d z_{-e_i}^2=\kappa\sum_{\lambda\in\Lambda_0}z_\lambda^2,$$

so that Assumption 2.3 (ii) is also satisfied. By comparing Theorems 2.4 and 2.5 and also definitions (2.6) *and* (2.9) *for approximations $\bar{u}^h$ and $\tilde{u}^h$ below, notice that the above choice of $\mathfrak{a}$ and $\mathfrak{b}$ is better than that of the previous example, in the sense that for $\tilde{u}^h$ we have fewer terms to calculate than for $\bar{u}^h$ to get the same order of accuracy of the approximations.*

Our results revolve about the possibility to prove the existence of random processes $u_t^{(j)}(x)$, $t\in[0,T],x\in\mathbb{R}^d$, $j=0,\dots,k$, for some integer $k\geq 0$ such that they are independent of h, $u^{(0)}$ is the solution of (2.1) with initial value u_0 and almost surely we have

$$u_t^h(x)=\sum_{j=0}^k\frac{h^j}{j!}u_t^{(j)}(x)+R_t^h(x)\tag{2.4}$$

for $h\neq 0$ and for all $t\in[0,T]$ and $x\in\mathbb{G}_h$, where u_t^h is the solution to (2.2) with initial data u_0 and R^h is a continuous $l_2(\mathbb{G}_h)$-valued adapted process, such that

$$E\sup_{t\in[0,T]}\sup_{x\in\mathbb{G}_h}|R_t^h(x)|^2\leq Nh^{2(k+1)}\mathcal{K}_m^2\tag{2.5}$$

with a constant N independent of h.

Theorem 2.2. *Let Assumptions 2.1, 2.2, 2.3 and 2.4 hold with*

$$\mathfrak{m}=m>k+1+d/2,$$

where $k\geq 0$ is an integer. Then expansion (2.4) *and estimate* (2.5) *hold with a constant N depending only on Λ, d, m, K_0, ..., K_{m+1}, $A_0,...,A_m$, κ, and T.*

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 2.2, one can replace $\mathbb{G}_h$ in (2.5) with $\mathbb{R}^d$.

Remark 2.6. Let Λ_0 be a basis in $\mathbb{R}^d$ such that Assumption 2.4 holds. Then Assumption 2.1 (i) implies Assumption 2.3 (i), and Assumption 2.1 (ii) implies Assumption 2.3 (ii) with $\mathfrak{m} = m$. Thus if Assumptions 2.1 and 2.2 hold with

$$m > k + 1 + d/2,$$

then the conditions of Theorem 2.2 are satisfied.

Equality (2.4) 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 $n \geq 0$, where $\Lambda^0 = \{0\}$ and

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

Theorem 2.2 can be generalised as follows.

THEOREM 2.3. *Let the conditions of Theorem 2.2 hold with*

$$\mathfrak{m} = m > k + n + 1 + d/2$$

for some integers $k \geq 0$ and $n \geq 0$. Then expansion (2.4) *holds and for $\lambda = (\lambda_1, ..., \lambda_n) \in \Lambda^n$*

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

where N depends only on Λ, d, m, K_0, ..., K_{m+1}, $A_0, ..., A_m$, κ and T.

We prove Theorem 2.3 in Section 4 after some preliminaries presented in Section 3. To discuss the method of acceleration we fix an integer $k \geq 0$ and set

$$\bar{u}^h = \sum_{j=0}^{k} b_j u^{2^{-j}h}, \tag{2.6}$$

where, naturally, $u^{2^{-j}h}$ are the solutions to (2.2), with $2^{-j}h$ in place of h,

$$(b_0, b_1, ..., b_k) := (1, 0, 0, ..., 0)V^{-1} \tag{2.7}$$

and V^{-1} is the inverse of the Vandermonde matrix with entries

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

The following consequence of Theorem 2.2 is the first main result of the paper on accelerated convergence. Its generalisation is presented in Section 4.

THEOREM 2.4. *Under the assumptions of Theorem 2.2 we have*

$$E \sup_{t\leq T} \sup_{x\in\mathbb{G}_h} |\bar{u}^h_t(x) - u^{(0)}_t(x)|^2 \leq N |h|^{2(k+1)} \mathcal{K}^2_m, \tag{2.8}$$

where N depends only on Λ, d, m, K_0, ..., K_{m+1}, κ, $A_0, ..., A_m$, and T.

Proof. By Theorem 2.2

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

with $\bar{r}^{2^{-j}h} := h^{-j(k+1)} R^{2^{-j}h}$, which gives

$$\bar{u}^h = \sum_{j=0}^{k} b_j u^{2^{-j}h} = (\sum_{j=0}^{k} b_j) u^{(0)} + \sum_{j=0}^{k} \sum_{i=1}^{k} b_j \frac{h^i}{i!2^{ij}} u^{(i)} + \sum_{j=0}^{k} b_j \bar{r}^{2^{-j}h} h^{k+1}$$

$$= u^{(0)} + \sum_{i=1}^{k} \frac{h^i}{i!} u^{(i)} \sum_{j=0}^{k} \frac{b_j}{2^{ij}} + \sum_{j=0}^{k} b_j \bar{r}^{2^{-j}h} = u^{(0)} + \sum_{j=0}^{k} b_j \bar{r}^{2^{-j}h} h^{k+1},$$

since

$$\sum_{j=0}^{k} b_j = 1, \quad \sum_{j=0}^{k} b_j 2^{-ij} = 0, \quad i = 1, 2, ...k$$

by the definition of $(b_0, ..., b_k)$. This and (2.5) yield the result and the theorem is proved. □

Remark 2.7. Let the conditions of Theorem 2.2 hold with

$$\mathfrak{m} = m > k + 1 + n + d/2,$$

where k and n are nonnegative integers. Then (2.8) holds with $\delta_{h,\lambda} \bar{u}^h$ and $\delta_{h,\lambda} u^{(0)}$ in place of $\bar{u}^h$ and $u^{(0)}$, respectively, for $\lambda \in \Lambda^n$.

Proof. This follows from Theorem 2.3 in the same way as Theorem 2.4 follows from Theorem 2.2. □

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

Sometimes it suffices to combine fewer terms $u^{2^{-j}h}$ to get accuracy of order $k+1$. For integers $k \geq 0$ define

$$\tilde{u}^h = \sum_{j=0}^{\tilde{k}} \tilde{b}_j u^{2^{-j}h}, \tag{2.9}$$

where

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

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

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

THEOREM 2.5. *Let the conditions of Theorem 2.2 hold. Then in the situation of Example 2.2 we have*

$$E \sup_{t \leq T} \sup_{x \in \mathbb{G}_h} |\tilde{u}^h_t(x) - u^{(0)}_t(x)|^2 \leq N |h|^{2(k+1)} \mathcal{K}^2_m, \tag{2.10}$$

where N is a constant depending only on d, m, κ, $K_0, \dots, K_{m+1}$, $A_0, ..., A_m$, and T.

To prove this result we need only repeat the proof of Theorem 2.4 taking into account that in (2.4) we have $u_t^{(j)} = 0$ for odd $j \le k$ since $u_t^h = u_t^{-h}$ owing to the fact that in the case of Example 2.2 equation (2.2) does not change if we replace h with $-h$.

Remark 2.8. Notice that without acceleration, i.e., when $k = 1$ in the above theorem, the mean square norm of the supremum in t and x of the error of the finite difference approximations in Example 2.2 is proportional to h^2. This is a sharp result see, e.g., Remark 2.21 in [3] on finite difference approximations for deterministic parabolic PDEs.

EXAMPLE 2.3. *Assume that in the situation of Example 2.2 we have $d = 2$ and $m = 7$. Then*

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

satisfies

$$E \sup_{t \le T} \sup_{x \in \mathbb{G}_h} |u_t^{(0)}(x) - \tilde{u}_t^h(x)| \le N h^4.$$

EXAMPLE 2.4. *Take $d = 1$ and consider the following SPDE:*

$$du_t = 3D^2 u_t \, dt + 2Du_t \, dw_t$$

with initial data $u_0(x) = \cos x$, where w_t is a one-dimensional Wiener process. Then a unique bounded solution is $u_t(x) = e^{-t} \cos(x + 2w_t)$. Example 2.2 suggests the following version of (2.2):

$$du_t^h(x) = 3 \frac{u_t^h(x+h) - 2u_t^h(x) + u_t^h(x-h)}{h^2} \, dt + \frac{u_t^h(x+h) - u_t^h(x-h)}{h} \, dw_t,$$

the unique bounded solution of which with initial condition $\cos x$ is given by

$$u_t^h(x) = e^{-c_h t} \cos(x + 2\phi_h w_t), \quad h^2 c_h = 12 \sin^2 \frac{h}{2} - 2 \sin^2 h, \quad \phi_h = \frac{\sin h}{h}.$$

For $t = 1$, $h = 0.1$, and $w_t = 0$ we have

$$u_1(0) \approx 0.3678794412, \quad u_1^h(0) \approx 0.366352748, \quad u_1^{h/2}(0) \approx 0.3674966179,$$

$$\tilde{u}_1^h(0) = \tfrac{4}{3} u_1^{h/2}(0) - \tfrac{1}{3} u_1^h(0) \approx 0.3678779079.$$

It is instructive to observe that such a level of accuracy is achieved for $u_1^{\tilde{h}}(0)$ with $\tilde{h} = 0.00316$, which is more than 15 times smaller than $h/2$.

Actually, this example does not quite fit into our scheme because u_0 is not square summable over $\mathbb{R}$. In connection with this we reiterate that the main goal of the present article was to introduce a method and not to prove the most general results. Without much trouble our approach can be extended to a class of SPDEs with growing data by the help of weighted Sobolev spaces (see [10]), and then the above example can be included formally.

3. Auxiliary facts. The following fact is easily obtained by Young's inequality owing to Assumption 2.3.

LEMMA 3.1. *Let Assumption 2.3 hold. Then for all $\varphi \in L_2$ we have*

$$\mathbb{Q}_t(\varphi) := \int_{\mathbb{R}^d} \left[2\varphi(x)L_t^h\varphi(x) + \sum_{\rho=1}^{\infty} |M_t^{h,\rho}\varphi(x)|^2\right] dx$$

$$\le -\frac{\kappa}{2} \sum_{\lambda\in\Lambda_0} \|\delta_{h,\lambda}\varphi\|_{L_2}^2 + N\|\varphi\|_{L_2}^2,$$

where N depends only on κ, A_0, A_1, and the number of vectors in Λ.

Proof. First observe that for $\mu \in \Lambda_0$ the conjugate operator in L_2 to $\delta_{-h,\mu}$ is $\delta_{h,\mu}$. Notice also that

$$\delta_{h,\mu}(\phi\psi) = \phi\delta_{h,\mu}\psi + (T_{h,\mu}\psi)\delta_{h,\mu}\phi,$$

where $T_{h,\mu}\psi(x) = \psi(x + h\mu)$. Thus by simple calculations we get

$$\mathbb{Q}_t(\varphi) = \sum_{i=1}^{4} \mathbb{Q}_t^{(i)}(\varphi)$$

with

$$\mathbb{Q}_t^{(1)}(\varphi) = -\int_{\mathbb{R}^d} \sum_{\lambda,\mu\in\Lambda_0} ((2\mathfrak{a}_t^{\lambda\mu} - \mathfrak{b}_t^{\lambda\rho}\mathfrak{b}_t^{\mu\rho})(\delta_{h,\lambda}\varphi)\delta_{h,\mu}\varphi)(x)\, dx$$

$$\mathbb{Q}_t^{(2)}(\varphi) = 2\int_{\mathbb{R}^d} \sum_{\lambda,\mu\in\Lambda_0} ((T_{h,\mu}\varphi)(\delta_{h,\lambda}\varphi)\delta_{h,\mu}\mathfrak{a}_t^{\lambda\mu})(x)\, dx$$

$$\mathbb{Q}_t^{(3)}(\varphi) = \int_{\mathbb{R}^d} 2\mathfrak{a}_t^{00}\varphi^2(x) + 2\varphi(x) \sum_{\lambda\in\Lambda_0} (\mathfrak{a}_t^{\lambda 0}\delta_{h,\lambda}\varphi + \mathfrak{a}_t^{0\lambda}\delta_{-h,\lambda}\varphi)(x)\, dx$$

$$\mathbb{Q}_t^{(4)}(\varphi) = \int_{\mathbb{R}^d} \mathfrak{b}_t^{00}\varphi^2(x) + 2\sum_{\lambda\in\Lambda_0} \mathfrak{b}_t^{\lambda\rho}\mathfrak{b}_t^{0\rho}\varphi\delta_{h,\lambda}\varphi(x)\, dx.$$

Due to Assumption 2.3 (ii)

$$\mathbb{Q}_t^{(1)}(\varphi) \le -\kappa \sum_{\lambda\in\Lambda_0} \|\delta_{h,\lambda}\varphi\|_{L_2}^2.$$

By Assumption 2.3(i), Young's inequality and the shift invariance of Lebesgue measure

$$\mathbb{Q}_t^{(i)}(\varphi) \le \frac{\kappa}{6} \sum_{\lambda\in\Lambda_0} \|\delta_{h,\lambda}\varphi\|_{L_2}^2 + N\|\varphi\|^2, \quad i = 2, 3, 4,$$

with a constant N depending only on the number of elements of Λ, κ, A_0 and, for $i = 2$ also on A_1. We finish the proof by summing up these estimates. ∎

Recall the notation $\mathbb{W}_2^m(T) = L_2(\Omega\times[0,T],\mathcal{P},W_2^m)$. Remember that W_2^m denotes the Hilbert-Sobolev space of real-valued and also that of l_2-valued functions on $\mathbb{R}^d$. Thus $\mathbb{W}_2^m(T)$ denotes the Hilbert space of predictable functions $\phi=\phi_t$ on $\Omega\times[0,T]$ with values in the W_2^m space of real-valued functions, and $\mathbb{W}_2^m(T)=\mathbb{W}_2^m(T,l_2)$ denotes the Hilbert space of functions $g=(g^\rho)_{\rho=1}^\infty$ with values in the W_2^m space of l_2-valued functions on $\mathbb{R}^d$, with norm defined by

$$\|\phi\|^2_{\mathbb{W}_2^m(T)} = E\int_0^T \|\phi_t\|^2_{W_2^m}\,dt<\infty \quad\text{and}$$

$$\|g\|^2_{\mathbb{W}_2^m(T)} = E\int_0^T \sum_{\rho=1}^\infty \|g^\rho_t\|^2_{W_2^m}\,dt<\infty,$$

respectively.

Theorem 3.2. *Let Assumption 2.3 (i) hold. Let $f^\mu\in\mathbb{W}_2^{\mathfrak{m}}(T)$, $\mu\in\Lambda$, and $(g^\rho)_{\rho=1}^\infty\in\mathbb{W}_2^{\mathfrak{m}}(T)$ be some functions. Then for each $h\neq 0$ there exists a unique continuous L_2-valued solution u_t^h of*

$$du_t^h = (\mathfrak{a}_t^{\lambda\mu}\delta_{h,\lambda}\delta_{-h,\mu}u_t^h+\delta_{-h,\mu}f_t^\mu)\,dt+(\mathfrak{b}_t^{\lambda\rho}\delta_{h,\lambda}u_t^h+g_t^\rho)\,dw_t^\rho \tag{3.1}$$

for any $W_2^{\mathfrak{m}}$-valued $\mathcal{F}_0$-measurable initial condition u_0. This solution is a $W_2^{\mathfrak{m}}$-valued continuous process. Moreover, if Assumption 2.3 (ii) is also satisfied, then

$$E\sup_{t\le T}\|u_t^h\|^2_{W_2^{\mathfrak{m}}}+E\int_0^T\sum_{\lambda\in\Lambda}\|\delta_{h,\lambda}u_t^h\|^2_{W_2^{\mathfrak{m}}}\,dt$$

$$\le NE\int_0^T\big(\sum_{\mu\in\Lambda}\|f_t^\mu\|^2_{W_2^{\mathfrak{m}}}+\|g_t\|^2_{W_2^{\mathfrak{m}}}\big)\,dt+NE\|u_0\|^2_{W_2^{\mathfrak{m}}}, \tag{3.2}$$

where N depends only on d, $\mathfrak{m}$, Λ, κ, $A_0,\dots,A_{\bar{\mathfrak{m}}}$, and T.

Proof. The first assertion is a simple consequence of the fact that (2.2) is an ordinary Itô equation with Lipschitz continuous coefficients for L_2-valued processes. Similarly, (2.2) has a unique $W_2^{\mathfrak{m}}$-valued solution and, since $W_2^{\mathfrak{m}}\subset L_2$, this proves that the L_2-valued solution is actually $W_2^{\mathfrak{m}}$-valued. Moreover, we can easily get estimate (3.2) with a constant N which depends on h. In particular we have that the solution is in $\mathbb{W}_2^{\mathfrak{m}}(T)$.

The proof of estimate (3.2) with N independent of h is rather standard but still contains a point which usually does not appear. This concerns the treatment of $\tilde{\mathfrak{a}}^{\lambda\mu}\delta_{h,\lambda}\delta_{-h,\mu}u_t^h$ after (3.8) without assuming that 2 derivatives of $\mathfrak{a}$ are bounded.

By Itô's formula for L_2-valued processes we find

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

$$+2(u_t^h,\mathfrak{b}^{\lambda\rho}\delta_{h,\lambda}u_t^h+g_t^\rho)\,dw_t^\rho. \tag{3.3}$$

We use Lemma 3.1, the inequalities like $|ab|\le\varepsilon a^2+\varepsilon^{-1}b^2$, and Assumption 2.3 (i) to conclude that

$$E\|u_t^h\|^2_{L_2}+\frac{\kappa}{2}E\int_0^t\sum_{\lambda\in\Lambda_0}\|\delta_{h,\lambda}u_s^h\|^2_{L_2}\,ds\le E\|u_0\|^2_{L_2}$$

$$+NE\int_0^t \big(\|u^h_s\|^2_{L_2}+\sum_{\lambda\in\Lambda}\|f^\lambda_s\|^2_{L_2}+\|g_s\|^2_{L_2}\big)\,ds<\infty. \tag{3.4}$$

By Gronwall's lemma we can eliminated the first term in the integral on the right in (3.4) and get that

$$E\int_0^t \sum_{\lambda\in\Lambda}\|\delta_{h,\lambda}u^h_s\|^2_{L_2}\,ds\le NE\|u_0\|^2_{L_2}$$

$$+NE\int_0^t \big(\sum_{\lambda\in\Lambda}\|f^\lambda_s\|^2_{L_2}+\|g_s\|^2_{L_2}\big)\,ds. \tag{3.5}$$

After that we come back to (3.3) and use Davis's inequality to derive that

$$E\sup_{t\le T}\|u^h_t\|^2_{L_2}\le NE\|u_0\|^2_{L_2}$$

$$+NE\int_0^T \big(\sum_{\lambda\in\Lambda}\|f^\lambda_t\|^2_{L_2}+\|g_t\|^2_{L_2}\big)\,dt+N_1J, \tag{3.6}$$

where

$$J=E\Big(\int_0^T\sum_{\rho=1}^\infty\Big(\int_{\mathbb{R}^d}|u^h_t(\mathfrak{b}^{\lambda\rho}\delta_{h,\lambda}u^h_t+g^\rho_t)|\,dx\Big)^2\,dt\Big)^{1/2}$$

$$\le E\Big(\int_0^T\|u^h_t\|^2_{L_2}\|\mathfrak{b}^\lambda\delta_{h,\lambda}u^h_t+g_t\|^2_{L_2}\,dt\Big)^{1/2}$$

$$\le E\sup_{t\le T}\|u^h_t\|_{L_2}\Big(\int_0^T\|\mathfrak{b}^\lambda\delta_{h,\lambda}u^h_t+g_t\|^2_{L_2}\,dt\Big)^{1/2}$$

$$\le(2N_1)^{-1}E\sup_{t\le T}\|u^h_t\|^2_{L_2}+NE\int_0^T\big(\sum_{\lambda\in\Lambda}\|\delta_{h,\lambda}u^h_t\|^2_{L_2}+\|g_t\|^2_{L_2}\big)\,dt.$$

This and (3.5) allow us to drop the last term in (3.6) which again combined with (3.5) yields

$$E\sup_{t\le T}\|u^h_t\|^2_{L_2}+E\int_0^T\sum_{\lambda\in\Lambda}\|\delta_{h,\lambda}u^h_t\|^2_{L_2}\,dt$$

$$\le NE\|u_0\|^2_{L_2}+NE\int_0^T\big(\sum_{\lambda\in\Lambda}\|f^\lambda_t\|^2_{L_2}+\|g_t\|^2_{L_2}\big)\,dt. \tag{3.7}$$

This proves the theorem if $\mathfrak{m}=0$. If $\mathfrak{m}\ge1$, we differentiate (3.1) with respect to x^i, and introduce the notation $\tilde\phi$ for the derivative of a function ϕ in x^i. Then we obtain

$$d\tilde u^h_t=(\mathfrak{a}^{\lambda\mu}\delta_{h,\lambda}\delta_{-h,\mu}\tilde u^h_t+\delta_{-h,\mu}\hat f^\mu_t)\,dt+(\mathfrak{b}^{\lambda\rho}_t\delta_{h,\lambda}\tilde u_t+\hat g^\rho_t)\,dw^\rho_t, \tag{3.8}$$

where

$$\hat{f}^\mu_t = \tilde{f}^\mu_t, \quad \mu \neq 0, \quad \hat{f}^0_t = \tilde{f}^0_t + \tilde{\mathfrak{a}}^{\lambda\mu}\delta_{h,\lambda}\delta_{-h,\mu}u^h_t, \quad \hat{g}^\rho_t = \tilde{g}^\rho_t + \tilde{\mathfrak{b}}^{\lambda\rho}_t\delta_{h,\lambda}u^h_t.$$

We proceed with (3.8) as above with (3.1) with one exception that for $\mu \in \Lambda_0$ we use the inequality (cf. Remark 3.1)

$$E\int_0^t\int_{\mathbb{R}^d}|\tilde{u}^h_s\delta_{h,\lambda}\delta_{-h,\mu}u^h_s|\,dxds \leq \int_0^t E\|\tilde{u}^h_s\|_{L_2}\|\delta_{h,\lambda}\partial_\mu u^h_s\|_{L_2}\,ds$$

$$\leq \varepsilon\int_0^t E\|D\delta_{h,\lambda}u^h_s\|^2_{L_2}\,ds + N\varepsilon^{-1}\int_0^t E\|\tilde{u}^h_s\|^2_{L_2}\,ds,$$

where $\partial_\mu = \mu^i D_i$ and $\varepsilon > 0$ is arbitrary and N depends only on $|\mu|$ (cf. Remark 3.1 below). Then we come to the following counterpart of (3.4)

$$E\|\tilde{u}^h_t\|^2_{L_2} + E\int_0^t\sum_{\lambda\in\Lambda}\|\delta_{h,\lambda}\tilde{u}^h_s\|^2_{L_2}\,ds$$

$$\leq NE\|\tilde{u}_0\|^2_{L_2} + (2d)^{-1}E\int_0^t\sum_{\lambda\in\Lambda}\|D\delta_{h,\lambda}u^h_s\|^2_{L_2}\,ds$$

$$+NE\int_0^t\big(\|\tilde{u}^h_s\|^2_{L_2} + \sum_{\lambda\in\Lambda}\|f^\lambda_s\|^2_{W^1_2} + \|g_s\|^2_{W^1_2}\big)\,ds. \tag{3.9}$$

Recall that here $\tilde{u}^h_t$ is the derivative of u^h_t with respect to x^i. By writing (3.9) for all $i = 1, ..., d$ and summing them up we see that the term with the factor $(2d)^{-1}$ is estimated by other terms on the right-hand side of (3.9) and, hence, can be dropped. After that the already familiar procedure yields

$$E\sup_{t\leq T}\|Du^h_t\|^2_{L_2} + E\int_0^T\sum_{\lambda\in\Lambda}\|D\delta_{h,\lambda}u^h_t\|^2_{L_2}\,dt$$

$$\leq NE\|u_0\|^2_{W^1_2} + NE\int_0^T\big(\sum_{\lambda\in\Lambda}\|f^\lambda_t\|^2_{W^1_2} + \|g_t\|^2_{W^1_2}\big)\,dt, \tag{3.10}$$

which along with (3.7) proves (3.2) with 1 in place of $\mathfrak{m}$.

Once this step is done the rest is routine. Assume that $\mathfrak{m} \geq 2$ and (3.2) is true with n in place of $\mathfrak{m}$ for an integer $n \in [1, \mathfrak{m}-1]$. Then we differentiate (3.1) $n+1$ times and now use the notation $\tilde{\phi}$ for certain $n+1$-th order derivative of ϕ with respect to x. Then we will obtain (3.8) with slightly modified $\hat{f}^0$ and $\hat{g}^\rho$. Namely, the $\hat{f}^0$ will be the sum of $\tilde{f}^0$ and the linear combination with constant coefficients of certain i-th derivatives of $\mathfrak{a}^{\lambda\mu}_t$ times certain $n+1-i$-th derivatives of $\delta_{h,\lambda}\delta_{-h,\mu}u^h_t$. Here i should be restricted to $[1, n+1]$. As above, the L_2-norms of the $n+1-i$-th derivatives of $\delta_{h,\lambda}\delta_{-h,\mu}u^h_t$ are dominated by the L_2-norms of the $n+2-i$-th derivatives of $\delta_{h,\lambda}u^h_t$ which are less than the W^{n+1}_2-norm of $\delta_{h,\lambda}u^h_t$ an estimate of which is contained in

(3.2) with n in place of l. Similar changes should be made in $\hat{g}^{\rho}$. After that we obtain the corresponding counterpart of (3.10) which yields (3.2) with $n+1$ in place of $\mathfrak{m}$. This obviously brings the proof of the theorem to an end. □

LEMMA 3.3. *Let $n \geq 0$ be an integer, let $\phi \in W_2^{n+1}$, $\psi \in W_2^{n+2}$, and $\lambda, \mu \in \Lambda_0$. Set*

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

Then we have

$$\frac{\partial^n}{(\partial h)^n} \delta_{h,\lambda} \phi(x) = \int_0^1 \theta^n \partial_\lambda^{n+1} \phi(x + h\theta\lambda)\, d\theta, \tag{3.11}$$

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

$$= \int_0^1 \int_0^1 (\theta_1 \partial_\lambda - \theta_2 \partial_\mu)^n \partial_{\lambda\mu} \psi(x + h(\theta_1 \lambda - \theta_2 \mu))\, d\theta_1 d\theta_2, \tag{3.12}$$

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

$$\big\| \delta_{h,\lambda} \phi - \sum_{i=0}^n \frac{h^i}{(i+1)!} \partial_\lambda^{i+1} \phi \big\|_{W_2^l} \leq \frac{|h|^{n+1}}{(n+2)!} \| \partial_\lambda^{n+2} \phi \|_{W_2^l} \tag{3.13}$$

$$\big\| \delta_{h,\lambda} \delta_{-h,\mu} \psi - \sum_{i=0}^n h^i \sum_{r=0}^i A_{i,r} \partial_\lambda^{r+1} \partial_\mu^{i-r+1} \psi \big\|_{W_2^l} \leq N |h|^{n+1} \| \psi \|_{W_2^{l+n+3}}, \tag{3.14}$$

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

$$A_{i,r} = \frac{(-1)^{i-r}}{(r+1)!(i-r+1)!}. \tag{3.15}$$

Proof. Clearly, it suffices to prove the lemma for $\phi, \psi \in C_0^\infty(\mathbb{R}^d)$. For $n = 0$ formula (3.11) is obtained by applying the Newton-Leibnitz formula to $\phi(x + \theta h \lambda)$ as a function of $\theta \in [0,1]$. Applying it one more time derives (3.12) from (3.11) for $n = 0$. After that for $n \geq 1$ one obtains (3.11) and (3.12) by differentiating both parts of these equations written with $n = 1$.

Next by Taylor's formula for smooth $f(h)$ we have

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

By applying this to

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

as a function of h we see that

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

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

Now to prove (3.13) it only remains to use that by Minkowski's inequality the W_2^l-norm of the last term 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, by observing that the value at $h=0$ of the right-hand side of (3.12) is

$$n!\sum_{r=0}^{n}A_{n,r}\partial_\lambda^{r+1}\partial_\mu^{n-r+1}\psi(x)$$

we see that the left-hand side of (3.14) is the W_2^l-norm of

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

This yields (3.14) in an obvious way.□

Remark 3.1. Formula (3.11) 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

$$\|\varphi\|_{l,r,h}^2=\sum_{\lambda_1,\dots,\lambda_r\in\Lambda}\|\delta_{h,\lambda_1}\cdot\ldots\cdot\delta_{h,\lambda_r}\varphi\|_{W_2^l}^2. \tag{3.16}$$

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

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

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

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

$$+(n+1)^{-1}\sum_{\mu\in\Lambda_0}\mathfrak{a}_t^{0\mu}\partial_\mu^{n+1},$$

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

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

where $A_{n,r}$ are defined by (3.15).

Remark 3.2. Formally, for $n\geq 1$ 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 3.3. Owing to Assumption 2.4 we have

$$\mathcal{L}_t^{(0)}=\mathcal{L}_t,\quad \mathcal{M}_t^{(0)\rho}=\mathcal{M}_t^\rho. \tag{3.17}$$

Also observe that in light of Lemma 3.3, under Assumption 2.3, for $\phi\in W_2^{n+2+l}$ and $\psi\in W_2^{n+3+l}$ we have

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

$$\|\mathcal{R}_t^{h(n)}\phi\|_{W_2^l}\leq N|h|^{n+1}\|\phi\|_{W_2^{l+n+2}}, \tag{3.18}$$

where N denotes constants depending only on n, d, l, $A_0,\dots,A_l$, and Λ.

Let $k\in[1,m]$ be an integer. The functions $u_t^{(1)},...,u_t^{(k)}$ we need in (2.4) will be obtained as the result of embedding of certain functions $v^{(i)}$ taking values in certain Sobolev spaces. Define $v_t^{(0)}$ as the solution of (2.1) from Theorem 2.1 and for finding $v^{(1)}$,...,$v^{(k)}$ introduce the following system of stochastic PDEs:

$$\begin{aligned}dv_t^{(n)}=&\big(\mathcal{L}_tv_t^{(n)}+\sum_{l=1}^{n}C_n^l\mathcal{L}_t^{(l)}v_t^{(n-l)}\big)\,dt\\&+\big(\mathcal{M}_t^\rho v_t^{(n)}+\sum_{l=1}^{n}C_n^l\mathcal{M}_t^{(l)\rho}v_t^{(n-l)}\big)\,dw_t^\rho,\quad n=1,...,k,\end{aligned} \tag{3.19}$$

where $C_n^l=n(n-1)\cdot...\cdot(n-l+1)/l!$ is the binomial coefficient.

THEOREM 3.4. *Let Assumptions 2.1, 2.2, and 2.3 (i) hold,* $\mathfrak{m}=m$, *and let* $1\leq k\leq m$. *Then there exists a unique set* $v_t^{(1)},...,v_t^{(k)}$ *of solutions of* (3.19) *with initial condition* $v_0^{(1)}=...=v_0^{(k)}=0$ *and such that* $v^{(n)}\in\mathbb{W}_2^{m+2-n}(T)$, $n=1,...,k$. *Furthermore, with probability one* $v_t^{(n)}$ *are continuous* W_2^{m+1-n}*-valued functions and there exists a constant* N *depending only on* T, d,κ, Λ, m, *and* K_0, ..., K_{m+1}, A_0,...,A_m *such that for* $n=1,...,k$

$$E\sup_{t\leq T}\|v_t^{(n)}\|^2_{W_2^{m+1-n}}+E\int_0^T\|v_t^{(n)}\|^2_{W_2^{m+2-n}}\,dt\leq N\mathcal{K}_m^2. \tag{3.20}$$

Proof. Notice that for each $n = 1, \dots, k$ the equation for $v_t^{(n)}$ does not involve the unknown functions $v_t^{(l)}$ with indices $l > n$. Therefore we can prove the solvability of (3.19) and the stated properties of $v_t^{(n)}$ recursively on n.

Denote

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

and first let $n = 1$. By Theorem 2.1 we have $v^{(0)} \in \mathbb{W}_2^{m+2}(T)$, which owing to Assumption 2.3(i) yields that $S^{(1)} \in \mathbb{W}_2^{m-1}(T)$ and $R^{(1)} = (R^{(1)\rho}) \in \mathbb{W}_2^m(T)$(here we need the assumption that $\mathfrak{m} = m$). Hence, it follows again by Theorem 2.1 that there exists a unique $v^{(1)} \in \mathbb{W}_2^{m+1}(T)$ satisfying (3.19) with zero initial condition. Furthermore, $v_t^{(1)}$ is a continuous $\mathbb{W}_2^m$-valued function (a.s.) and (3.20) 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. Observe that for $i = 1, \dots, n$

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

$$= N \|v^{(n-i)}\|_{\mathbb{W}_2^{m+2-(n-i)}(T)}, \tag{3.21}$$

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

$$= N \|v^{(n-i)}\|^2_{\mathbb{W}_2^{m+2-(n-i)}(T)}. \tag{3.22}$$

It follows by the induction hypothesis that $S^{(n)} \in \mathbb{W}_2^{m-n}(T)$ and $R^{(n)} \in \mathbb{W}_2^{m-n+1}(T)$. By applying Theorem 2.1 we see that there exists a unique $v^{(n)} \in \mathbb{W}_2^{m-n+2}(T)$ satisfying (3.19) with zero initial condition. This theorem also yields the continuity property of $v_t^{(n)}$ and an estimate, that combined with (3.21) and (3.22) and the induction hypothesis yields (3.20). This proves the existence. Uniqueness is obtained by inspecting the above proof in which each $v^{(n)}$ was found uniquely. ☐

LEMMA 3.5. *Let Assumptions 2.1, 2.2, and 2.3 (i) hold and* $\mathfrak{m} = m$. *Let* $l, k \geq 0$ *be integers such that* $l + k + 1 = m$, *and let* $v^{(0)}, ..., v^{(k)}$ *be the functions from Theorem 3.4. Set*

$$r_t^h = v_t^h - v_t^{(0)} - \sum_{1 \leq j \leq k} \frac{h^j}{j!} v_t^{(j)}, \tag{3.23}$$

where v^h *is the unique* L_2*-valued solution of* (3.1) *with initial condition* u_0, $f^0 = f$ *and* $f^\mu = 0$ *for* $\mu \in \Lambda_0$. *Then* $r_0^h = 0$, $r^h \in \mathbb{W}_2^{m-k}(T)$, *and*

$$dr_t^h = (L_t^h r_t^h + F_t^h)\, dt + (M_t^{h,\rho} r_t^h + G_t^{h,\rho})\, dw_t^\rho, \tag{3.24}$$

where

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

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

Proof. Due to Assumptions 2.2 and 2.3(i) we have $v^h \in \mathbb{W}_2^m(T)$, and owing to Assumptions 2.1 and 2.2, by Theorem 2.1 we have $v^{(0)} \in \mathbb{W}^{m+2}(T)$. Hence clearly $r^h \in \mathbb{W}^m(T)$ when $k=0$, and $r^h \in \mathbb{W}_2^{m-k}(T)$ follows from Theorem 3.4 when $k \geq 1$. A direct computation shows that (3.24) 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 an empty set 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. Finally, the last assertion of the lemma immediately follows from Remark 3.3 and Theorem 3.4 (see however the proof of Theorem 4.1). □

4. Proof of Theorem 2.3 . In this section we suppose that $\mathfrak{m} = m$. We start with a result which, as will be seen later, is more general than Theorem 2.3 .

THEOREM 4.1. *Let $m = l + k + 1$ for some integers $l, k \geq 0$, and let Assumptions 2.1, 2.2, 2.3, and 2.4 hold. Then for r_t^k, defined as in Lemma 3.5, we have*

$$E \sup_{t \leq T} \|r_t^h\|_{W_2^l}^2 + E \int_0^T \sum_{\lambda \in \Lambda} \|\delta_{h,\lambda} r_t^h\|_{W_2^l}^2 \, dt \leq N |h|^{2(k+1)} \mathcal{K}_m^2, \tag{4.1}$$

where N depends only on T, d, κ, Λ, m and $K_0, \dots, K_{m+1}$, $A_0, \dots$, A_m. Moreover,in the situation of Example 2.2 we have $v^{(j)} = 0$ in (3.23) for odd $j \leq k$.

Proof. By Lemma 3.5 we have $F^h \in \mathbb{W}_2^l(T)$ and $G^{h,\cdot} \in \mathbb{W}_2^{l+1}(T)$, which by Lemma 3.5 and Theorem 3.2 yields that the left-hand side of (4.1) is dominated by

$$NE \int_0^T (\|F_t^h\|_{W_2^l}^2 + \|G_t^h\|_{W_2^l}^2) \, dt. \tag{4.2}$$

To estimate (4.2) we observe that for $j \leq k$ by Remark 3.3 we have

$$\|\mathcal{O}_t^{h(k-j)} u_t^{(j)}\|_{W_2^l} \leq N |h|^{k-j+1} \|u_t^{(j)}\|_{W_2^{l+k-j+3}} = N |h|^{k-j+1} \|u_t^{(j)}\|_{W_2^{m+2-j}}.$$

Upon combining this result with Theorem 3.4 we see that

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

Similarly one can estimate the remaining part of (4.2) thus proving estimate (4.1). Finally, observe that in Example 2.2 we have $v^h = v^{-h}$ due to the uniqueness of the L_2-valued solution for equation (2.2) with initial condition u_0. Hence (4.1) yields $v^{(j)} = 0$ for odd $j \leq k$. ☐

By Sobolev's 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

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

for all $\varphi \in W_2^l$, where N is a constant depending only on d and l. 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 (2.2).

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

$$\sum_{x \in \mathbb{G}_h} |I\varphi(x)|^2 |h|^d \leq N \|\varphi\|_{W_2^l}^2, \tag{4.3}$$

where N is a constant depending only on d and l.

Proof. By Sobolev's embedding of W_2^l into C_b, for $z \in \mathbb{R}^d$ and smooth φ we have

$$|\varphi(z)|^2 \leq \sup_{x \in B_1(0)} \varphi^2(z + hx) \leq N \sum_{|\alpha| \leq l} h^{2|\alpha|} \int_{B_1(0)} |(D^\alpha \varphi)(z + hx)|^2 \, dx$$

$$= N \sum_{|\alpha| \leq l} |h|^{2|\alpha| - d} \int_{B_h(z)} |(D^\alpha \varphi)(x)|^2 \, dx \leq N |h|^{-d} \sum_{|\alpha| \leq l} \int_{B_h(z)} |(D^\alpha \varphi)(x)|^2 \, dx,$$

with a constant $N = N(d,l)$, where $B_r(z) = \{x \in \mathbb{R}^d : |x-z| < r\}$. Thus

$$\sum_{z\in\mathbb{G}_h} |\varphi(z)|^2|h|^d \le N \sum_{|\alpha|\le l}\sum_{z\in\mathbb{G}_h}\int_{B_h(z)} |(D^\alpha\varphi)(x)|^2\,dx$$

that yields (4.3). □

Set $R^h_t = Ir^h_t$. 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 $(\lambda_1,\ldots,\lambda_n)\in\Lambda^n$, $n\ge 1$. Then we have the following corollary of Theorem 4.1

Corollary 4.3. *Let the assumptions of Theorem 4.1 hold with $l > n + d/2$ for some integer $n \ge 0$. Then for $\lambda\in\Lambda^n$ we have*

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

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

with a constant N depending only on Λ, d, m, K_0, ..., K_{m+1}, A_0,..., A_m, κ, and T.

Proof. Set $j = n - l$. Then $j > d/2$ and using Sobolev's theorem on embedding W^j_2 into C_b and taking into account Remark 3.1, from Theorem 4.1 we get

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

$$\le C_2 E\sup_{t\in[0,T]}\|R^h_t\|^2_{W^l_2} \le Nh^{2(k+1)}\mathcal{K}^2_m,$$

where C_1 and C_2 are constants depending only on m and d, and N is a constant depending only on m, d, T, κ, Λ and K_i for $i\le m+1$. Similarly, by Lemma 4.2 and Remark 3.1

$$E\sup_{t\in[0,T]}\sum_{x\in\mathbb{G}_h}|\delta_{h,\lambda}R^h_t(x)|^2|h|^d \le C_1 E\sup_{t\in[0,T]}\|\delta_{h,\lambda}R^h_t\|^2_{W^j_2}$$

$$\le C_2 E\sup_{t\in[0,T]}\|R^h_t\|^2_{W^l_2} \le Nh^{2(k+1)}\mathcal{K}^2_m.$$

□

Now we show that Theorem 2.3 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 (2.2) with initial condition u_0, the processes $v^{(0)}$,...,$v^{(k)}$ are given by Theorem 3.4, and I is the embedding operator from W^l_2 into C_b. By virtue of Theorem 3.2, v^h is a continuous W^l_2-valued process, and by Theorem 3.4 $v^{(j)}$, $j = 1,2,...,k$, are W^{n+1-k}_2-valued continuous processes. Since $l > d/2$ and $n+1-k > d/2$, the processes $\hat u^h$ and $u^{(j)}$ are well-defined and clearly (3.23) implies (2.4) with $\hat u^h$ in place of u^h. To show that Corollary 4.3 yields Theorem 2.3 we need only show that almost surely

$$\hat u^h_t(x) = u^h_t(x) \quad \text{for all } t\in[0,T] \tag{4.4}$$

for each $x \in \mathbb{G}_h$, where u^h is the unique $\mathcal{F}_t$-adapted l_2-valued continuous solution of (2.2). 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 (2.2), for each ε almost surely

$$\int_{\mathbb{R}^d} \hat{u}^h_t(y)\varphi_\varepsilon(y)\,dy = \int_{\mathbb{R}^d} \hat{u}(y)\varphi_\varepsilon(y)\,dy + \int_0^t \int_{\mathbb{R}^d} (L^h_s \hat{u}^h_s(y) + f_s(y))\varphi_\varepsilon(y)\,dy\,ds$$

$$+ \int_0^t \int_{\mathbb{R}^d} (M^{h,\rho}_s \hat{u}^h_s(y) + g^\rho_s(y))\varphi_\varepsilon(y)\,dy\,dw^\rho_s$$

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

$$\hat{u}^h_t(x) = u_0(x) + \int_0^t \left[L^h_s \hat{u}^h_s(x) + f_s(x)\right] ds + \int_0^t \left[M^{h,\rho}_s \hat{u}^h_s(x) + g^\rho_s(x)\right] dw^\rho_s$$

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 (2.2) for any l_2-valued $\mathcal{F}_0$-measurable initial condition, we have (4.4), that finishes the proof Theorem 2.3.

Theorem 2.3 yields the following generalisation of Theorem 2.4.

Theorem 4.4. *Let the conditions of Theorem 2.3 hold with $n = 0$. Then*

$$E \sup_{t \le T} \sup_{x \in \mathbb{G}_h} |\bar{u}^h_t(x) - u^{(0)}_t(x)|^2$$

$$+ E \sup_{t \le T} \sum_{x \in \mathbb{G}_h} |\bar{u}^h_t(x) - u^{(0)}_t(x)|^2 |h|^d \le N |h|^{2(k+1)} \mathcal{K}^2_m, \tag{4.5}$$

where $\bar{u}^h$ is defined by (2.9) *and N depends only on Λ, d, m, K_0, ..., K_{m+1}, A_1,...,A_m, κ, and T. In the situation of Example 2.2 estimate* (4.5) *holds also for $\tilde{u}^h$, defined by* (2.9)*, in place of $\bar{u}$.*

REFERENCES

[1] V. Bally, and D. Talay, *The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function*, Probab. Theory Related Fields 104 (1996), no. 1, pp. 43–60.

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

[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, pp. 435–490.

[5] I. Gyöngy and N. V. Krylov, *An accelerated splitting-up method for parabolic equations*, SIAM Journal on Mathematical Analysis, 37 (2006), pp. 1070–1097.

[6] I. Gyöngy and N. V. Krylov, *Expansion of solutions of parametrized equations and acceleration of numerical methods*, Illinois Journal of Mathematics, **50** (2006), pp. 473-514. Special Volume in Memory of Joseph Doob (1910-2004).

[7] I. Gyöngy and N. V. Krylov, *First derivative estimates for finite difference schemes*, Math. Comp., 78 (2009), pp. 2019–2046

[8] I. Gyöngy and N. V. Krylov, *Higher order derivative estimates for finite-difference schemes*, Methods and Applications of Analysis, 16 (2009), pp. 187–216.

[9] I. Gyöngy and N.V. Krylov, *Accelerated finite difference schemes for second order degenerate elliptic and parabolic problems in the whole space*, submitted.

[10] I. Gyöngy and N.V. Krylov, *Stochastic partial differential equations with unbounded coefficients and applications III.*, Stochastics and Stochastics Rep. 40 (1992), pp. 77–115.

[11] P. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, Springer–Verlag, Berlin, 1992.

[12] P. Kloeden, E. Platen and N. Hoffmann, *Extrapolation methods for the weak approximation to Itô diffusions* , SIAM J. Numer. Analysis 32 (1995), pp. 1519–1534.

[13] N. V. Krylov, *An analytic approach to SPDEs*, pp. 185–242 in Stochastic Partial Differential Equations: Six Perspectives, Mathematical Surveys and Monographs, Vol. 64, AMS, Providence, RI, 1999.

[14] N. V. Krylov and B.L. Rosovskii, *Stochastic evolution equations*, J. Soviet Mathematics, 16 (1981), pp. 1233–1277.

[15] P. Malliavin and A. Thalmaier, *Numerical error for SDE: asymptotic expansion and hyperdistributions*, C. R. Math. Acad. Sci. Paris 336 (2003), no. 10, pp. 851–856.

[16] G. I. Marchuk, *Methods of numerical mathematics*, Third edition, “Nauka”, Moscow, 1989 (in Russian).

[17] G. I. Marchuk and V. V. Shaidurov, *Difference methods and their extrapolations*, New York Berlin Heidelberg, Springer 1983.

[18] 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), pp. 307–357.

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

[20] B. L. Rozovskii, *Stochastic evolution systems. Linear theory and applications to nonlinear filtering*, Kluwer Academic Publishers Group, Dordrecht, 1990.

[21] D. Talay and L. Tubaro, *Extension of the global error for numerical schemes solving stochastic differential equations*, Stochastic Analysis and Applications, 8 (1990), pp. 483–509.

[22] H. Yoo, *An analytic approach to stochastic partial differential equations and its applications*, Thesis, University of Minnesota 1998.

[23] H. Yoo, *Semi-discretization of stochastic partial differential equations on R^1 by a finite-difference method*, Math. Comp. 69 (2000), pp. 653–666.