On stochastic finite difference schemes

István Gyöngy

Received: 12 December 2013 / Published online: 16 September 2014
© Springer Science+Business Media New York 2014

Abstract Finite difference schemes in the spatial variable for degenerate stochastic parabolic PDEs are investigated. Sharp estimates on the rate of L_p and almost sure convergence of the finite difference approximations are presented and results on Richardson extrapolation are established for stochastic parabolic schemes under smoothness assumptions.

Keywords Cauchy problem · Finite differences · Extrapolation to the limit · Richardson's method

Mathematics Subject Classification 65M15 · 35J70 · 35K65

1 Introduction

We consider finite difference schemes to stochastic partial differential equations (SPDEs). The stochastic PDEs we are interested in are linear second order stochastic parabolic equations on the whole $\mathbb{R}^d$ in the spatial variable. They may degenerate and become first order stochastic or deterministic PDEs. The finite difference schemes which we investigate are spatial discretizations of such SPDEs. One can view them as (possibly degenerate) infinite systems of stochastic differential equations, whose components describe the time evolution of approximate values at the grid points of the solutions to SPDEs. Adapting the approach of [1] we view stochastic finite difference schemes, like in [2] and [3], as stochastic equations for random fields on the whole $\mathbb{R}^d$ not only on grids.

I. Gyöngy (✉)
School of Mathematics and Maxwell Institute, University of Edinburgh,
King's Buildings, Edinburgh EH9 3JZ, UK
e-mail: I.Gyongy@ed.ac.uk

Our aim is to investigate the rate of convergence in the supremum norm of the finite difference approximations. We show that under the *stochastic parabolicity* condition, if the coefficients and the data are sufficiently smooth, then the solutions to the finite difference schemes admit power series expansions in h , the mesh-size of the grid. The coefficients in these power series are random fields, independent of h , and for any $p > 0$ the p -th moments of the sup norm of the remainder term is estimated by a power of h . This is Theorem 2.2. Hence for $h \rightarrow 0$ we get the convergence (and the sharp rate) of the solutions of the finite difference schemes to a random field which is the solution to the corresponding SPDE. Moreover, by Richardson extrapolation we get that the rate of convergence can be accelerated to any high order if one takes appropriate mixtures of approximations corresponding to different grid sizes. In Theorem 2.4 we obtain convergence estimates for any (high) p -th moments of the sup norm of the approximation error. Hence in Theorem 2.5 we get almost sure rate of convergence of the finite difference approximations and of the accelerated approximations in sup norm.

Theorems 2.2 and 2.4 are generalisations of the main results, Theorems 2.3 and 2.5 of [3], which besides the conditions on the smoothness of the coefficients and of the data require also a *smooth factorisation* condition to be satisfied. Namely, in [3] it is assumed that the matrix $\tilde{a} = \tilde{a}(\omega, t, x)$ from the ‘stochastic parabolicity condition’ can be written as the product of a sufficiently smooth matrix and its transpose. This assumption, however, may not hold even if $\tilde{a}$ is an infinitely differentiable nonnegative matrix, and therefore it strongly restricts the applicability of the results of [3].

Theorem 2.4 follows easily from Theorem 2.2. To prove Theorem 2.2 we adapt an approach from [2]. A key ingredient of this approach is that we consider the finite difference scheme as a stochastic equation given on the whole $\mathbb{R}^d$ not just on the grid. Thus we can deal with generalised solutions to the finite difference scheme, with values in Sobolev spaces. This gives the possibility of estimating the finite difference approximations in Sobolev norms first, and hence in discrete Sobolev norms and also in sup norm, by the help of embedding theorems. Moreover, by differentiating formally the equation for the finite difference approximation with respect to the mesh parameter h , say k times, we can identify the coefficients from the first up to the k -th order terms in the power series expansion of the finite difference approximation, as the solution of a system of SPDEs. Hence by proving a suitable existence and uniqueness theorem for this system, we can propose a power series expansion as a Taylor polynomial of degree k and a remainder term, say r_h , which we define as the difference of the finite difference approximation and the Taylor polynomial. In order to estimate r_h independently of h , the key observation is that r_h satisfies a stochastic finite difference equation with the same finite difference operators as the original one, with zero initial condition and with appropriate free terms, so that general estimates in Sobolev norms we obtained for this equation can be applied.

To adapt the above approach in the present paper the main challenge is to estimate the L_2 -norm of the spatial derivatives of the solutions to stochastic finite difference schemes, without assuming the smooth factorisation condition of [3]. This requires a different method than that used in [3]. Instead of considering only spatial derivatives of solutions to stochastic finite difference schemes, we consider spatial derivatives mixed with finite differences, and to estimate the L_2 -norm of such expressions we extend the

method of [4] to stochastic finite difference schemes. This leads to somewhat tedious calculations with spatial derivatives and finite differences of any order. To organise the computations we treat spatial derivatives and finite differences in a unified way by introducing similar notation for them and using a general Leibniz formula involving derivatives and finite differences simultaneously. We note that the idea of considering suitable expressions with mixtures of spatial derivatives and finite differences is used also in [5] to estimate the supremum norm of solutions to deterministic finite difference schemes. The approach in [5], however, is different from the approach in the present paper. In [5] it is based on a maximum principle of monotone finite difference schemes considered there, and that cannot be applied to the stochastic finite difference schemes we deal with in this paper.

The method of finite differences is one of the basic methods of solving partial differential equations numerically. The rate of convergence of various finite difference schemes for elliptic and parabolic PDEs have been studied extensively in the literature when the equations are non degenerate, but there are only a few publications dedicated to the numerical analysis of finite difference schemes for degenerate equations. Sharp rate of convergence in *sup* norm is obtained in [6] for fully discretized degenerate elliptic and parabolic PDEs. The finite difference schemes investigated in [6] are *monotone schemes*. In [5] for a large class of monotone finite difference schemes (in the spatial variables) power series expansions are obtained and Richardson extrapolation is used to get accelerated schemes for degenerate elliptic and parabolic PDEs. We note that the finite difference schemes we study in the present paper are not necessarily monotone, and our main theorems extend some of the results of [5] to non-monotone finite difference schemes for degenerate parabolic PDEs.

About a century ago Richardson had the idea that the order of accuracy of an approximation method, which depends on a parameter, can be dramatically improved if the approximation calculated by the method admits a power series expansion in the parameter. One need only take appropriate linear combinations of approximations corresponding to different parameter values to eliminate the lower order terms in the power series to get approximations with accuracy of higher order. Richardson used this idea to solve numerically some PDEs by finite difference methods (see [7] and [8]). He called his method a *deferred approach to the limit*. It is often called *Richardson extrapolation* in numerical analysis. Richardson extrapolation is applied by Romberg to the trapezium rule to obtain high order approximations of definite integrals (see [9]). Since then Richardson extrapolation has been applied to a wide range of numerical approximations. (See, for example, the textbook [10], the monograph [11] and the survey papers [12] and [13]). To show that it is applicable to an approximation method to solve numerically a problem under certain conditions one has to show the existence of a suitable power series expansion, which can be quite difficult in some situations. A series expansion for the weak convergence of Euler approximations of SDEs was obtained in [14] and then in a more general setting in [15]. The applicability of Richardson extrapolation to accelerate the convergence of finite difference schemes in the spatial variable for stochastic PDEs was shown in [2], under the strong stochastic parabolicity condition, and then it was shown in [3] for degenerate SPDEs under the smooth factorisation condition mentioned above. The results from [2] and [3] have

been generalised to finite difference schemes in temporal and spatial variables in [16] and [17].

The paper is organised as follows. In the next section we formulate our main result, Theorem 2.2 on power series expansions for stochastic finite difference schemes. Hence our results on the L_p and almost sure rate of convergence of the accelerated schemes, Theorems 2.4 and 2.5, follow easily. In Sect. 2 we formulate also an existence and uniqueness theorem, Theorem 2.1, for stochastic parabolic (possibly degenerate) equations. Though this result is known from [18] and [19], we give a new proof of it in Sect. 4, based on Theorems 4.4 and 4.5 about stochastic finite difference schemes proved in the same section. In Sect. 3 we present the technical tools to prove our key estimate, Theorem 4.4 on stochastic finite difference schemes. In Sect. 5 we prove an existence and uniqueness theorem, Theorem 5.1 for a system of SPDEs by the help of Theorem 2.1. Theorem 5.1 plays a crucial role in identifying the coefficients of the expansion in the main theorem. We prove the main theorem in Sect. 6 adapting a method from [2].

We conclude by introducing some notation. We fix a complete probability space $(\Omega, \mathcal{F}, P)$ and an increasing family $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ of σ -algebras $\mathcal{F}_t \subset \mathcal{F}$ throughout the paper. We assume that $\mathbb{F}$ is right continuous and that $\mathcal{F}_0$ contains all P -zero sets. The σ -algebra of the predictable subsets of $\Omega \times [0, \infty)$ is denoted by $\mathcal{P}$, and the σ -algebra of the Borel subsets of $\mathbb{R}^d$ is denoted by $\mathcal{B}(\mathbb{R}^d)$. The notation $|v|$ means the Euclidean norm of v for vectors $v \in \mathbb{R}^d$, and it stands for the Hilbert-Schmidt norm of v for matrices $v \in \mathbb{R}^{d \times m}$. For a vector $v \in \mathbb{R}^d$ the derivative of a function φ in the direction of v is denoted by $\partial_v \varphi$. For the standard basis $e_1, \dots, e_d$ in $\mathbb{R}^d$ we use the notation

$$D_\alpha = D_i = \frac{\partial}{\partial x^i} = \partial_{e_i} \quad \text{for } \alpha = i \in \{1, \dots, d\},$$

and we use $D_\alpha = \partial_{e_\alpha}$ for the identity operator when $\alpha = 0$. For a positive integer k the notation $|D^k \varphi|$ is used for the Euclidean norm of the vector whose components are (in some ordering) the partial derivatives of φ of order k , and $|D^k \varphi|$ means $|\varphi|$ for $k = 0$. For every integer $m \geq 0$ we use the notation $H^m = W_2^m$ for the Sobolev space defined as the completion in the norm $|\varphi|_m$ of $C_0^\infty = C_0^\infty(\mathbb{R}^d)$, the smooth functions φ with compact support on $\mathbb{R}^d$, where

$$|\varphi|_m^2 = |\varphi|_{H^m}^2 = \sum_{j=0}^k \int_{\mathbb{R}^d} |D^j \varphi(x)|^2 dx.$$

The Sobolev spaces $H^m(l_2) = W_2^m(l_2)$ of functions with values in $l_2 = \{(c_n)_{n=1}^\infty \in \mathbb{R}^\infty : \sum_j |c_j|^2 < \infty\}$ is defined analogously, and the notation $|\varphi|_m = |\varphi|_{H^m(l_2)}$ is used for the norm of φ in $H^m(l_2)$. The inner product of functions ψ and φ in $H^0 = L_2(\mathbb{R}^d)$ is denoted by (ψ, φ) . For a Hilbert space H , an H -valued process $(u_t)_{t \in [0, T]}$ is said to be a weakly continuous process, if almost surely $(u_t, \varphi)_H$ is continuous in $t \in [0, T]$ for all $\varphi \in H$, where $(\psi, \varphi)_H$ denotes the inner product of ψ and φ in H .

The summation convention with respect to repeated indices with values in discrete sets is used throughout the paper, unless it is otherwise indicated at some expressions.

2 Formulation of the main results

We consider the stochastic PDE

$$\begin{aligned} du_t(x) = & \left(a_t^{\alpha\beta}(x) D_\alpha D_\beta u_t(x) + f_t(x) \right) dt \\ & + \left(b_t^{\alpha,r}(x) D_\alpha u_t(x) + g_t^r(x) \right) dw_t^r, \end{aligned} \quad (2.1)$$

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

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

where $(w^r)_{r=1}^\infty$ is a sequence of independent Wiener martingales with respect to $\mathbb{F}$. The coefficients, $a^{\alpha\beta} = a^{\beta\alpha}$ and $b^\alpha = (b^{\alpha,r})_{r=1}^\infty$, are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$ -measurable bounded functions on $\Omega \times H_T$, with values in $\mathbb{R}$ and in l_2 respectively, for every $\alpha, \beta \in \{0, 1, \dots, d\}$. The free terms, $f = f_t(\cdot)$ and $g = (g_t^r(\cdot))_{r=1}^\infty$ are H^1 -valued and $H^2(l_2)$ -valued adapted processes for $t \geq 0$, respectively, such that almost surely

$$\int_0^T |f_t|_{H^1}^2 dt < \infty, \quad \int_0^T |g_t|_{H^2(l_2)}^2 dt < \infty.$$

For a $\mathcal{F}_0$ -measurable H^1 -valued initial value ψ the solution to (2.1)–(2.2) is defined as follows.

Definition 2.1 An H^1 -valued adapted weakly continuous process $u = (u_t)_{t \in [0, T]}$ is a (generalised) solution to (2.1)–(2.2) on a stochastic interval $[0, \tau]$ for a stopping time $\tau \leq T$, if almost surely

$$\begin{aligned} (u_t, \varphi) = & (\psi, \varphi) + \int_0^t \left\{ (\partial_{e_\alpha} u_s, \partial_{-e_\beta} (a_s^{\alpha\beta} \varphi)) + (f_s, \varphi) \right\} ds \\ & + \int_0^t (b_s^{\alpha,r} \partial_{e_\alpha} u_s + g_s^r, \varphi) dw_s^r, \end{aligned}$$

for $t \in [0, \tau]$ and $\varphi \in C_0^\infty$, where the summation convention is in force with respect to the repeated indices $\alpha, \beta \in \{0, 1, \dots, d\}$ and $r \in \{0, 1, \dots\}$.

We approximate Eq. (2.1) with finite difference schemes in the spatial variable. To describe these schemes let $\Lambda_1 \subset \mathbb{R}^d$ be a finite set, containing the zero vector, and set $\Lambda_0 = \Lambda_1 \setminus \{0\}$. For $h \in \mathbb{R} \setminus \{0\}$ define the grid

$$\mathbb{G}_h = \{h(\lambda_1 + \dots + \lambda_n) : \lambda_i \in \Lambda_1 \cup \{-\Lambda_1\}, n = 1, 2, \dots\},$$

and the finite difference operators $\delta_{h,\lambda}, \delta_\lambda^h$ by

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

for $\lambda \in \Lambda_0 \cup \{-\Lambda_0\}$. For notational convenience we define $\delta_{h,\lambda}$ and δ_λ^h to be the identity operator for $\lambda = 0$.

In order to satisfy the natural requirement that the grid $\mathbb{G}_h$ should contain finitely many points in every bounded set, we make the following assumption on Λ_0 .

Assumption 2.1 If a subset $\Lambda' \subset \Lambda_0$ is linearly dependent over $\mathbb{R}$, then Λ' is linearly dependent also over $\mathbb{Q}$, the set of rational numbers.

For $h \in (-1, 1) \setminus \{0\}$ we consider the equation

$$du_t^h(x) = \left(L_t^h u_t^h(x) + f_t(x) \right) dt + \left(M_t^{h,r} u_t^h(x) + g_t^h(x) \right) dw_t^r \quad (2.3)$$

for $t \in [0, T]$ and $x \in \mathbb{G}_h$, with initial condition

$$u_0^h(x) = \psi(x), \text{ for } x \in \mathbb{G}_h, \quad (2.4)$$

where

$$L_t^h = \sum_{\lambda, \mu \in \Lambda_1} \alpha_t^{\lambda, \mu} \delta_\lambda^h \delta_\mu^h + \sum_{\gamma \in \Lambda_0} \left(p_t^\gamma \delta_{h,\gamma} - q_t^\gamma \delta_{-h,\gamma} \right), \quad (2.5)$$

$$M_t^{h,r} = \sum_{\lambda \in \Lambda_1} b_t^{\lambda, r} \delta_\lambda^h, \quad r = 1, 2, \dots, \quad (2.6)$$

with real-valued $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$ -measurable bounded functions, $\alpha^{\lambda, \mu} = \alpha^{\mu, \lambda}$, p^γ, q^γ , and l_2 -valued $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$ -measurable bounded functions $b^\lambda = (b^{\lambda, r})_{r=1}^\infty$ on $\Omega \times H_T$, for all $\lambda, \mu \in \Lambda_1$ and $\gamma \in \Lambda_0$.

Equation (2.3) is an infinite system of stochastic differential equations. We look for its solutions in the space of adapted stochastic processes with values in $l_{h,2}$, the space of real functions ϕ on $\mathbb{G}_h$ with the norm $|\phi|_{l_{h,2}}$ defined by

$$|\phi|_{l_{h,2}}^2 := \sum_{x \in \mathbb{G}_h} |\varphi(x)|^2 |h|^d < \infty.$$

We note that this norm together with other discrete Hilbert–Sobolev norms are often used in numerical analysis of finite difference schemes for PDEs. In the context of finite difference approximations of stochastic PDEs they were used, for example, in [20].

Remark 2.1 Due to the boundedness of $\mathfrak{a}^{\lambda\mu}$ and $\mathfrak{b}^\lambda$ for $\lambda, \mu \in \Lambda_1$, it is easy to see that for each $h \neq 0$ there is a constant C such that for all $\phi \in l_{h,2}$

$$\left| L^h \phi \right|_{l_{h,2}}^2 \leq C \left| \phi \right|_{l_{h,2}}^2, \quad \sum_{r=1}^{\infty} \left| M^{h,r} \phi \right|_{l_{h,2}}^2 \leq C \left| \phi \right|_{l_{h,2}}^2.$$

Thus by standard results on SDEs in Hilbert spaces the initial value problem (2.3)–(2.4) admits a unique solution $(u_t^h)_{t \in [0,T]}$, provided almost surely

$$\left| \psi \right|_{l_{h,2}}^2 + \int_0^T \left| f_t \right|_{l_{h,2}}^2 + \left| g_t \right|_{l_{h,2}}^2 dt < \infty \text{ (a.s.)}, \quad (2.7)$$

where $\left| g_t \right|_{l_{h,2}}^2 = \sum_{r=1}^{\infty} \left| g_t^r \right|_{l_{h,2}}^2$.

Clearly, for $h \rightarrow 0$

$$\delta_{h,\lambda} \varphi(x) \rightarrow \partial_\lambda \varphi(x), \quad \delta_\lambda^h \varphi(x) \rightarrow \partial_\lambda \varphi(x)$$

for smooth functions φ on $\mathbb{R}^d$. Thus in order L^h and $M^{h,r}$ approximate the differential operators

$$\mathcal{L} = \mathfrak{a}^{\alpha\beta} D_\alpha D_\beta, \text{ and } \mathcal{M}^r = \mathfrak{b}^{\alpha,r} D_\alpha,$$

respectively, we make the following assumption.

Assumption 2.2 For every $i, j = 1, \dots, d$ we have

$$\begin{aligned} a^{ij} &= \sum_{\lambda, \mu \in \Lambda_0} \mathfrak{a}^{\lambda\mu} \lambda^i \mu^j, \quad a^{0i} + a^{0i} = \sum_{\lambda \in \Lambda_0} \left(\mathfrak{a}^{0\lambda} + \mathfrak{a}^{\lambda 0} + \mathfrak{p}^\lambda - \mathfrak{q}^\lambda \right) \lambda^i, \\ b^i &= \sum_{\lambda \in \Lambda_0} \mathfrak{b}^\lambda \lambda^i, \quad a^{00} = \mathfrak{a}^{00}, \quad b^0 = \mathfrak{b}^0. \end{aligned}$$

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

$$\begin{aligned} \mathfrak{a}^{e_\alpha e_\beta} &= a^{\alpha\beta}, \quad \mathfrak{b}^{e_\alpha} = b^\alpha, \quad \alpha, \beta = 0, 1, \dots, d, \\ \mathfrak{q}^{e_\gamma} &= \mathfrak{p}^{e_\gamma} = 0 \quad \gamma = 1, \dots, d. \end{aligned}$$

Example 2.2 Take $\Lambda = \{e_0, e_1, \dots, e_d\}$ with $e_0 = 0$ as before, and define

$$\begin{aligned} \mathfrak{a}^{e_\alpha e_\beta} &= a^{\alpha\beta}, \quad \mathfrak{a}^{e_\alpha 0} = \mathfrak{a}^{0e_\alpha} = 0, \quad \alpha, \beta = 1, \dots, d, \\ \mathfrak{p}^{e_\gamma} &= \frac{1}{2} a^{0\gamma} + \theta^\gamma, \quad \mathfrak{q}^{e_\gamma} = -\frac{1}{2} a^{\gamma 0} + \theta^\gamma, \quad \gamma = 1, \dots, d, \\ \mathfrak{a}^{00} &= a^{00}, \quad \mathfrak{b}^{e_\alpha} = b^\alpha, \quad \alpha = 0, 1, \dots, d, \end{aligned}$$

where $\theta^1, \dots, \theta^d$ are any constants such that $|a^{0,\gamma}| \leq 2\theta^\gamma$, $|a^{\gamma 0}| \leq 2\theta^\gamma$ for each $\gamma \in \{1, \dots, d\}$.

To formulate our main results we make further assumptions. Let $m \geq 0$ be an integer, and let $K \geq 0$ be a constant. Set $|\Lambda_0|^2 = \sum_{\lambda \in \Lambda_0} |\lambda|^2$ and

$$\mathcal{K}_l^2(t) = \int_0^t (|f_l|^2 + |g_{l+1}|^2) dt \quad (2.8)$$

for $t \geq 0$ and integers $l \geq 0$.

Assumption 2.3 The functions $a^{\lambda\mu}$, b^λ , Db^λ and their partial derivatives in $x \in \mathbb{R}^d$ up to order m are continuous in x and they are bounded in magnitude by K for all λ and $\mu \in \Lambda_1$. Moreover, for $\lambda, \mu \in \Lambda_0$ the partial derivatives in x of $a^{\lambda\mu}$ up to order $\max(m, 2)$, of b^λ up to order $\max(m+1, 2)$, and of $a^{\lambda 0}$, p^λ , q^λ up to order $\max(m, 1)$ are continuous in x and in magnitude are bounded by K .

Assumption 2.4 The initial value ψ is an H^m -valued $\mathcal{F}_0$ -measurable random variable, $(f_t)_{t \geq 0}$ is an H^m -valued predictable process and $(g_t)_{t \geq 0}$ is an $H^{m+1}(l_2)$ -valued predictable process, such that almost surely $\mathcal{K}_m^2(T) < \infty$.

Assumption 2.5 For $P \otimes dt \otimes dx$ -almost all $(\omega, t, x) \in \Omega \times H_T$ we have

- (i) $\sum_{\lambda, \mu \in \Lambda_0} (a^{\lambda\mu} - \frac{1}{2} b^{\lambda,r} b^{\mu,r}) z_\lambda z_\mu \geq 0$ for all $z_\lambda \in \mathbb{R}$, $\lambda \in \Lambda_0$;
- (ii) $p^\lambda \geq 0$, $q^\lambda \geq 0$ for all $\lambda \in \Lambda_0$.

Remark 2.2 By virtue of Sobolev's theorem on embedding H^m into C_b , the space of bounded and continuous functions on $\mathbb{R}^d$, if Assumption 2.4 holds with $m > d/2$, then there exist functions $\bar{\psi}$, $\bar{f}_t$ and $\bar{g}_t$ such that for every $\omega \in \Omega$ and $t \in [0, T]$ they are continuous in $x \in \mathbb{R}^d$, and $\psi(x) = \bar{\psi}(x)$, $f_t(x) = \bar{f}_t(x)$ and $g_t(x) = \bar{g}_t(x)$ for dx -almost every $x \in \mathbb{R}^d$. Therefore when Assumption 2.4 holds with $m > d/2$, then without loss of generality we can, and will always assume that ψ , f_t and g_t are continuous functions in x for all $\omega \in \Omega$ and $t \in [0, T]$. Moreover, in this case (2.7) also holds, due to a result on embedding H^m into $l_{h,2}$, see Lemma 4.1 below. Consequently, if Assumption 2.4 holds with $m > d/2$, then (2.3)–(2.4) admits a unique solution $(u^h)_{t \in [0, T]}$.

An existence and uniqueness theorem for (2.1)–(2.2) reads as follows.

Theorem 2.1 Let Assumptions 2.2, 2.3, 2.4 and 2.5 hold with $m \geq 1$. Then (2.1)–(2.2) has a unique solution $(u_t)_{t \in [0, T]}$. Moreover, u is an H^m -valued weakly continuous process, it is strongly continuous as an H^{m-1} -valued process, and for every $p > 0$ and stopping time $\tau \leq T$ we have

$$E \sup_{t \leq \tau} |u_t|_l^p \leq N \left(E |\psi|_l^p + E \mathcal{K}_l^p(\tau) \right) \quad (2.9)$$

for every integer $l \in [0, m]$, where N is a constant depending only on K , T , m , d , p and $|\Lambda_0|$.

This theorem is a special case of Theorem 3.1 from [19], which improves Theorem 3.1 from [18]. As a by-product of our results on finite difference approximations, we give a new proof of it in Section 4.

Our aim is to establish an expansion of u^h in the form

$$u_t^h(x) = \sum_{j=0}^k \frac{h^j}{j!} u_t^{(j)}(x) + h^{k+1} r_t^h(x) \quad (2.10)$$

for integers $k \geq 0$, where $u^{(0)}$ is the solution of (2.1)–(2.2), $u^{(1)}, \dots, u^{(k)}$ are random fields on H_T , independent of h , and r^h is a random field on H_T such that

$$E \sup_{t \leq T} \sup_{x \in \mathbb{G}_h} |r_t^h(x)|^p + E \sup_{t \leq T} |r_t^h|_{l_{h,2}}^p \leq N \left(E |\psi|_{\mathfrak{m}}^p + EK_{\mathfrak{m}}^p(T) \right) \quad (2.11)$$

holds for $h \in (0, 1)$, for any $p > 0$, for sufficiently large $\mathfrak{m}$, with a constant N independent of h .

Theorem 2.2 *Let $k \geq 0$ be an integer, and let Assumptions 2.1 through 2.5 hold with*

$$\mathfrak{m} > 2k + 3 + \frac{d}{2}.$$

Then there are continuous random fields $u^{(1)}, u^{(2)}, \dots, u^{(k)}$ on H_T , independent of h , such that for each $h \in (0, 1)$ (2.10) holds almost surely for $t \in [0, T]$ and $x \in \mathbb{G}_h$ with a continuous random field $(r_t^h(x))_{(t,x) \in H_T}$ satisfying (2.11) with $N = N(K, T, k, d, \Lambda_0)$. If k is odd and $\mathfrak{p}^\lambda = \mathfrak{q}^\lambda = 0$ for all $\lambda \in \Lambda_0$, then it is sufficient that Assumptions 2.1 through 2.5 hold only with

$$\mathfrak{m} > 2k + 2 + \frac{d}{2}$$

to have the same conclusion. Moreover, in this case (2.10) and (2.11) hold for each $h \in (-1, 1) \setminus \{0\}$, and $u^{(l)} = 0$ for all odd $l \leq k$.

This is our main theorem which is proved in Sect. 6. It clearly implies the following result.

Corollary 2.3 *Let $p > 0$. If Assumptions 2.1 through 2.5 hold with $\mathfrak{m} > 3 + (d/2)$, then for all $h \in (0, 1)$*

$$E \sup_{[0,T] \times \mathbb{G}_h} |u - u^h|^p + E \sup_{t \in [0,T]} |u_t - u_t^h|_{l_{h,2}}^p \leq Nh^p \left(E |\psi|_{\mathfrak{m}}^p + EK_{\mathfrak{m}}^p(T) \right).$$

If Assumptions 2.1 through 2.5 hold with $\mathfrak{m} > 4 + (d/2)$, and $\mathfrak{p}^\lambda = \mathfrak{q}^\lambda = 0$ for all $\lambda \in \Lambda_0$, then for all $h \in (0, 1)$

$$E \sup_{[0,T] \times \mathbb{G}_h} |u - u^h|^p + E \sup_{t \in [0,T]} |u_t - u_t^h|_{l_{h,2}}^p \leq Nh^{2p} \left(E |\psi|_{\mathfrak{m}}^p + EK_{\mathfrak{m}}^p(T) \right).$$

In these estimates N is a constant depending only on K, T, p, d and Λ_0 .

Remark 2.3 By an example given in [6] one can see that the rate of convergence stated in the above corollary in each of the cases is sharp. (See Remark 2.21 in [6]).

Remark 2.4 If $M^{h,r} = 0$ and $g^r = 0$ for all $r \geq 1$ in (2.3) then we get finite difference schemes (in the spatial variable) for parabolic (possibly degenerate) PDEs. Since these schemes are not necessarily monotone, Theorem 2.2 and Corollary 2.3 are new results also in the case of deterministic PDEs. They generalise the corresponding results, Theorem 2.3 in [5] and Theorems 2.16 and 2.18 in [6] on monotone schemes to a class of schemes which contains also non monotone finite difference schemes.

Now we formulate some implications of Theorem 2.2 on *Richardson extrapolation*. We set

$$(c_0, c_1, \dots, c_k) = (1, 0, \dots, 0)V^{-1}, \quad (\tilde{c}_0, \tilde{c}_1, \dots, \tilde{c}_{\tilde{k}}) = (1, 0, \dots, 0)\tilde{V}^{-1},$$

where $\tilde{k} = (k-1)/2$ for odd integer $k \geq 0$, V^{-1} is the inverse of the $(k+1) \times (k+1)$ Vandermonde matrix $V = (V^{ij})$ given by $V^{ij} = 2^{-(i-1)(j-1)}$, and $\tilde{V}^{-1}$ is the inverse of the $\tilde{k} \times \tilde{k}$ Vandermonde matrix $\tilde{V}$ given by $\tilde{V}^{ij} = 4^{-(i-1)(j-1)}$. Define

$$v^h = \sum_{i=0}^k c_i u^{h/2^i}$$

when Assumptions 2.2 through 2.5 hold with $m > 2k + 3 + d/2$ for some integer $k \geq 0$, and define also

$$\tilde{v}^h = \sum_{i=0}^{\tilde{k}} \tilde{c}_i u^{h/2^i}$$

when $k \geq 1$ is odd, $p^\lambda = q^\lambda = 0$ for all $\lambda \in \Lambda_0$, and Assumptions 2.2 through 2.5 hold with $m > 2k + 2 + d/2$.

Theorem 2.4 Let $p > 0$. Let Assumptions 2.1 through 2.5 hold with $m > 2k + 3 + d/2$ for some integer $k \geq 0$. Then for all $h \in (0, 1)$

$$\begin{aligned} & E \sup_{t \in [0, T]} \sup_{x \in \mathbb{G}_h} \left| u_t(x) - v_t^h(x) \right|^p + E \sup_{t \in [0, T]} \left| u_t - v_t^h \right|_{l_{h,2}}^p \\ & \leq N h^{p(k+1)} \left(E |\psi|_m^p + E \mathcal{K}_m^p(T) \right) \end{aligned} \quad (2.12)$$

holds with a constant $N = N(K, T, k, p, \Lambda_0)$. Let $k \geq 1$ be an odd number and let Assumptions 2.1 through 2.5 hold with $m > 2k + 2 + d/2$. Then estimate (2.12) holds for all $h \in (0, 1)$ with $\tilde{v}^h$ in place of v^h .

Proof By definition of the coefficients c_i and $\tilde{c}_i$, from Theorem 2.2 we have that for each h almost surely

$$u_t(x) - v_t^h(x) = h^{k+1} R_t^{(h)}(x), \quad u_t(x) - \tilde{v}_t^h(x) = h^{k+1} \tilde{R}_t^{(h)}(x)$$

for all $t \in [0, T]$ and $x \in \mathbb{G}_h$, where $R^{(h)} = \sum_{j=0}^k c_j 2^{-j} r^{h/2^j}$ and $\tilde{R}^{(h)} = \sum_{j=0}^{\tilde{k}} \tilde{c}_j 2^{-j} r^{h/2^j}$. Hence the theorem follows from Theorem 2.2 by virtue of estimate (2.11). $\square$

Example 2.3 Assume that we have $d = 2$, $m = 10$ and $p^\lambda = q^\lambda = 0$ for every $\lambda \in \Lambda_0$. Then

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

satisfies

$$E \sup_{t \leq T} \sup_{x \in \mathbb{G}_h} \left| u_t(x) - \tilde{v}_t^h(x) \right|^2 \leq Nh^8 \left(E|\psi|_m^2 + EK_m^2(T) \right).$$

Remark 2.5 If in addition to the conditions of Theorem 2.4 $\psi \geq 0$, $f_t \geq 0$ and $g_t^r = 0$ are also satisfied for $t \in [0, T]$, $x \in \mathbb{R}^d$ and $r \geq 1$, then u , the solution of (2.1)–(2.2), is nonnegative. Such situation arises, for example in the case of the Zakai equation in nonlinear filtering, where $f = 0$, $g = 0$ and ψ is the conditional density of the initial value of the signal, given the initial value of the observation. Notice, however that even in such cases, in general, v^h and $\tilde{v}^h$ can take negative values. If we want our approximations to be also non-negative then we may take, for example, $(v^h)^+$ and $(\tilde{v}^h)^+$ in place of v^h and $\tilde{v}^h$, respectively. Since $|z - w^+| \leq |z - w|$ for any $z \in [0, \infty)$ and $w \in \mathbb{R}$, for $u_t(x) \geq 0$ we have

$$\left| u_t(x) - (v_t^h(x))^+ \right| \leq \left| u_t(x) - v_t^h(x) \right|, \quad \left| u_t(x) - (\tilde{v}_t^h(x))^+ \right| \leq \left| u_t(x) - \tilde{v}_t^h(x) \right|.$$

Consequently, if u is a nonnegative random field, then Theorem 2.4 holds also with $(v^h)^+$ and $(\tilde{v}^h)^+$ in place of v^h and $\tilde{v}^h$, respectively.

Theorem 2.4 implies the following results on almost sure rate of convergence.

Theorem 2.5 Let $(h_n)_{n=1}^\infty$ be a nonnegative sequence from l_q for some $q \geq 1$. Let Assumptions 2.1 through 2.5 hold with $m > 2k + 3 + d/2$ for some integer $k \geq 0$. Then for each $\varepsilon > 0$ there is a finite random variable ξ_ε such that almost surely

$$\sup_{t \in [0, T]} \sup_{x \in \mathbb{G}_h} \left| u_t(x) - v_t^h(x) \right| + \sup_{t \in [0, T]} \left| u_t - v_t^h \right|_{l_{h,2}} \leq \xi_\varepsilon h^{k+1-\varepsilon} \quad (2.13)$$

holds for $h = h_n$, for integers $n \geq 1$. Let $k \geq 1$ be an odd number such that Assumptions 2.2 through 2.5 hold with $m > 2k + 2 + d/2$. Then for every $\varepsilon > 0$ there is a finite random variable ξ_ε such that almost surely (2.13) holds with $\tilde{v}^h$ in place of v^h for $h = h_n$ for all $n \geq 1$.

Proof We prove only the statement concerning v^h , since the assertion for $\tilde{v}^h$ can be proved in the same way. Set $\Omega_r = \{\omega \in \Omega : |\psi(\omega)|_m \leq r\}$ for integers $r \geq 1$. Then $\Omega_r \in \mathcal{F}_0$, and $P(\cup_{r=1}^\infty \Omega_r) = 1$. If for each r we have (2.13) on Ω_r , with some almost

surely finite random variable $\xi_{\varepsilon r}$ instead of ξ_ε , then clearly we have (2.13) almost surely with some finite random variable ξ_ε . Thus we may additionally assume that $|\psi|_m$ is bounded by a constant. Define

$$\tau_r = \inf \left\{ t \geq 0 : \mathcal{K}_m(t) \geq r \right\} \text{ for integers } r \geq 1.$$

Then τ_r is a stopping time for each r , and due to Assumption 2.4, $\tau_r \rightarrow \infty$ as $r \rightarrow \infty$. Thus if for each r we have a finite random variable $\xi_{\varepsilon r}$ such that (2.13) holds almost surely with $u_{t \wedge \tau_r}$, $v_{t \wedge \tau_r}^h$ and $\xi_{\varepsilon r}$ in place of u_t , v_t^h and ξ_ε , respectively for every $h = h_n$, then there is a finite random variable ξ_ε such that (2.13) holds almost surely for all $h = h_n$, $n \geq 1$. Therefore in addition to the assumptions of the theorem we may assume that $|\psi|_m + \mathcal{K}_m(T)$ is bounded by a constant, say c . Let η_h denote the left-hand side of inequality (2.13), and set $\zeta_n = h_n^{-k-1+\varepsilon} \eta_{h_n}$. Then under the additional assumption that $|\Psi|_m + \mathcal{K}_m(T) \leq c$, by Theorem 2.4 we obtain

$$E \zeta_n^p \leq N c^p h_n^{\varepsilon p} \quad \text{for all } n \geq 1 \text{ and } p > 0,$$

where the constants N and c are independent of n . Taking here p so large that $p\varepsilon \geq q$, we get

$$E \sum_{n=1}^{\infty} \zeta_n^p = \sum_{n=1}^{\infty} E \zeta_n^p \leq N c^p \sum_{n=1}^{\infty} h_n^{\varepsilon p} < \infty.$$

Hence for

$$\xi_\varepsilon := \left(\sum_{n=1}^{\infty} \zeta_n^p \right)^{1/p}$$

we obtain that almost surely $\xi_\varepsilon < \infty$ and $\eta_{h_n} \leq \xi_\varepsilon h_n^{k+1-\varepsilon}$ for all n , which completes the proof of the theorem. $\square$

Example 2.4 Consider the degenerate parabolic SPDE

$$du_t = 2D^2 u_t dt + 2Du_t dw_t \quad t \in (0, 1], x \in \mathbb{R},$$

with initial condition $u_0(x) = \cos x$, and approximate it by the finite difference equation

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

The unique bounded solution of the SPDE problem is

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

and the unique bounded solution to the finite difference equation (with initial condition $u_0^h(x) = \cos x$) 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

$$\begin{aligned} u_1(0) &\approx -0.4161468365, \\ u_1^h(0) &\approx -0.4131150562, \quad u_1^{h/2}(0) \approx -0.415389039, \\ \tilde{u}_1^h(0) &= \frac{4}{3}u_1^{h/2}(0) - \frac{1}{3}u_1^h(0) \approx 0.4161470333. \end{aligned}$$

Such level of accuracy by $\tilde{u}_1^h(0)$ is achieved with $\tilde{h} = 0.0008$, which is more than 60 times smaller than $h/2$.

Note that since $\cos x$ is not square integrable over $\mathbb{R}$, the above example does not fit into our setting, but using suitable Sobolev spaces we can extend our setting so that this example can be included.

3 Preliminaries

We introduce some notation. For $\lambda \in \Lambda_1 \cup \{-\Lambda_1\}$ and for $h \neq 0$ we use $T_{h,\lambda}$ and T_λ^h for the operators defined by

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

for functions φ given on $\mathbb{R}^d$, where I denotes the identity operator. For the sake of uniformity of notation we take a set of unit vectors, $\Lambda_2 = \{\ell_1, \dots, \ell_d\}$, which is disjoint from $\Lambda_0 \cup \{-\Lambda_0\}$, and use the notation

$$\delta_{h,\ell_i} = \delta_{\ell_i}^h = D_i = \frac{\partial}{\partial x^i}$$

for the partial derivative in the direction of the i -th basis vector e_i for $i = 1, \dots, d$ and $h \neq 0$. We introduce this new notation for D_i in order to present in a unified way calculations with expressions involving partial derivatives and finite differences. We will use also the notation I_λ^h for T_λ^h when $\lambda \in \Lambda_1$, and for the identity when $\lambda \in \Lambda_2$. Set $T_{h,\lambda} = I$ for $\lambda \in \Lambda_2$. It is easy to see that

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

$$= v\delta_{h,\lambda}u + u\delta_{h,\lambda}v + h_\lambda(\delta_{h,\lambda}u)(\delta_{h,\lambda}v) \tag{3.2}$$

for all $h \neq 0$ and

$$\lambda \in \Lambda := \Lambda_0 \cup \Lambda_2,$$

where $h_\lambda = h$ if $\lambda \in \Lambda_0$ and $h_\lambda = 0$ if $\lambda \in \Lambda_2$. Hence we get

$$\delta_\lambda^h(uv) = (\delta_\lambda^h u) I_\lambda^h v + (I_\lambda^h u) \delta_\lambda^h v \text{ for } \lambda \in \Lambda. \quad (3.3)$$

Indeed, this is the classical Leibniz rule when $\lambda \in \Lambda_2$, and for $\lambda \in \Lambda_0$ by virtue of (3.1) we have

$$\begin{aligned} \delta_\lambda^h(uv) &= (\delta_\lambda^h u) v + \frac{1}{2} \left\{ (\delta_{h,\lambda} v) T_{h,\lambda} u + (\delta_{-h,\lambda} v) T_{h,-\lambda} u \right\} \\ &= (\delta_\lambda^h u) v + (\delta_\lambda^h v) T_{-h,\lambda} u + \frac{1}{2} (T_{h,\lambda} u - T_{-h,\lambda} u) \delta_{h,\lambda} v \\ &= (\delta_\lambda^h u) v + (\delta_\lambda^h v) T_{-h,\lambda} u + (\delta_\lambda^h u) (T_{h,\lambda} v - v) \\ &= (\delta_\lambda^h u) T_{h,\lambda} v + (\delta_\lambda^h v) T_{-h,\lambda} u. \end{aligned}$$

Since $\delta_\lambda^h = \delta_\lambda^{-h}$, symmetrizing the last equality in h we get (3.3).

To extend this Leibniz rule to higher order finite differences and derivatives we introduce further notation.

Let Λ^n denote the set of sequences $\lambda_1, \dots, \lambda_n$ of length n of elements of Λ . We use the notation $|\alpha| := n$ for the length of $\alpha \in \Lambda^n$. We introduce a ‘sequence’ of ‘length zero’, which we denote by ϵ . It is considered a sub-sequence of any $\lambda \in \Lambda^n$ for each $n \geq 1$. For $\mu \in \Lambda^m$ and $\lambda \in \Lambda^n$, $m \leq n$ we write $\mu \leq \lambda$ if μ is a sub-sequence of λ , and $\lambda \setminus \mu$ denotes the sequence we obtain from λ by removing μ from it. In particular, $\lambda \setminus \lambda = \epsilon$, $\lambda \setminus \epsilon = \lambda$. For $\lambda = \lambda_1, \dots, \lambda_n \in \Lambda^n$ we set

$$\delta_\lambda = \delta_\lambda^h = \delta_{\lambda_1}^h, \dots, \delta_{\lambda_n}^h, \quad I_\lambda = I_\lambda^h = I_{\lambda_1}^h, \dots, I_{\lambda_n}^h,$$

and for ϵ we set

$$\delta_\epsilon = I_\epsilon = I.$$

Now a generalisation of (3.1) reads as follows.

Lemma 3.1 *Let $\lambda \in \Lambda^n$ for $n \geq 1$. Then*

$$\delta_\lambda(uv) = \sum_{\mu \leq \lambda} (\delta_\mu I_{\lambda \setminus \mu} u) (\delta_{\lambda \setminus \mu} I_\mu v),$$

where the summation is taken over all sub-sequences of λ , including ϵ .

Proof The lemma can be proved by a straightforward induction on n , the length of the multi-sequence λ . $\square$

Notice that for $\lambda \in \Lambda$

$$I_\lambda = I + \frac{h^2}{2} \Delta_\lambda^h = h P_\lambda + I, \quad (3.4)$$

where

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

$$P_\lambda := \frac{1}{2}(\delta_{h,\lambda} - \delta_{-h,\lambda})$$

for $\lambda \in \Lambda_1$, and $\Delta_\lambda^h = P_\lambda = 0$ for $\lambda \in \Lambda_2$ and $h \neq 0$. Hence for $\alpha = \alpha_1\alpha_2, \dots, \alpha_m \in \Lambda^m$, $m \geq 1$ we get

$$I_\alpha = I + h^2\mathcal{O}_\alpha = I + h\mathcal{P}_\alpha, \quad (3.5)$$

where

$$\mathcal{O}_\alpha = \left(\Delta_{\alpha_1} I_{\alpha_2\alpha_3, \dots, \alpha_m} + \Delta_{\alpha_2} I_{\alpha_3, \dots, \alpha_m} + \dots + \Delta_{\alpha_{m-1}} I_{\alpha_m} + \Delta_{\alpha_m} \right) / 2,$$

$$\mathcal{P}_\alpha = P_{\alpha_1} I_{\alpha_2\alpha_3, \dots, \alpha_m} + P_{\alpha_2} I_{\alpha_3, \dots, \alpha_m} + \dots + P_{\alpha_{m-1}} I_{\alpha_m} + P_{\alpha_m}.$$

Set

$$R_\lambda = (T_{h,\lambda} - T_{h,-\lambda})/2 \text{ for } \lambda \in \Lambda,$$

and notice that for any functions a and u on $\mathbb{R}^d$ we have

$$I_\mu(au) = aI_\mu u + h(P_\mu a)I_\mu u + h(\delta_\mu a)(R_\mu u) \quad (3.6)$$

$$= (I_\mu a)(I_\mu u) + (R_\mu a)(R_\mu u) \quad (3.7)$$

for all $\mu \in \Lambda$ and $h \neq 0$. Indeed, if $\mu \in \Lambda_2$, then $P_\mu = R_\mu = 0$, and these equalities hold by definition. If $\mu \in \Lambda_0$ then

$$I_\mu(au) = aI_\mu u + \frac{h}{2} \{ (\delta_{h,\mu} a) T_{h,\mu} u + (\delta_{h,-\mu} a) T_{h,-\mu} u \}$$

$$= aI_\mu u + h(P_\mu a)(T_{h,-\mu} u) + (R_\mu a)(R_\mu u),$$

and symmetrizing the right-hand side of the last equality in μ and $-\mu$, we obtain (3.6). Hence we get equality (3.7) by using (3.4).

We will often make use of the fact that for $\lambda \in \Lambda^k$

$$\left| \delta_\lambda v \right|_0 \leq N |v|_k, \text{ for } v \in H^k,$$

where N is a constant depending only on d, k and $|\Lambda|$.

More precisely, the following lemma holds.

Lemma 3.2 *Let k be a positive integer. Then for $\lambda = \lambda_1\lambda_2, \dots, \lambda_k \in \Lambda^k$ and $h \neq 0$*

$$\left| \delta_{h,\lambda} v \right|_0 \leq \prod_{i=1}^k |\lambda_i| \left| D^k v \right|_0, \quad \left| \delta_\lambda^h v \right|_0 \leq \prod_{i=1}^k |\lambda_i| \left| D^k v \right|_0$$

for all $v \in H^k$.

Proof The second inequality clearly follows from the first one. It is sufficient to prove the first inequality for smooth functions v with compact support. For $\lambda \in \Lambda$ let $\bar{\lambda}$ denote λ if $\lambda \in \Lambda_0$, and let $\bar{\lambda} = e_i$ if $\lambda = l_i \in \Lambda_2$. Then

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

for every $x \in \mathbb{R}^d$, where $\partial_{\bar{\lambda}}$ is the directional derivative along $\bar{\lambda}$, $h_\lambda = h$ for $\lambda \in \Lambda_0$ and $h_\lambda = 0$ for $\lambda \in \Lambda_2$. Hence for $\lambda \in \Lambda^k$ and smooth v we get by induction on k that

$$\delta_{h,\lambda} v(x) = \int_{[0,1]^k} \partial_{\bar{\lambda}_1}, \dots, \partial_{\bar{\lambda}_k} v(x + \theta_1 h_{\lambda_1} \lambda_1 + \dots + \theta_k h_{\lambda_k} \lambda_k) d\theta_1 d\theta_2, \dots, d\theta_k$$

for every $x \in \mathbb{R}^d$, which by Minkowski's inequality and by the shift invariance of the Lebesgue measure yields

$$|\delta_{h,\lambda} v|_0 \leq |\partial_{\bar{\lambda}_1}, \dots, \partial_{\bar{\lambda}_k} v|_0.$$

We can finish the proof by noting that

$$\left| \partial_{\bar{\lambda}_1}, \dots, \partial_{\bar{\lambda}_k} v(x) \right|^2 \leq \left| D^k v(x) \right|^2 \prod_{i=1}^k |\lambda_i|^2.$$

□

Lemma 3.3 *Let $\mathfrak{a}$ be a bounded function on $\mathbb{R}^d$. Assume that the first order partial derivatives of $\mathfrak{a}$ are functions, which together with $\mathfrak{a}$ are bounded in magnitude by a constant K . Then for all $h \neq 0$ and $u \in H^1$ we have*

$$\begin{aligned} (I_\mu u, \mathfrak{a} \delta_\lambda u) &= -\frac{1}{2}((\delta_\lambda \mathfrak{a}) I_\lambda I_\mu u, u) - \frac{1}{2}(I_\mu R_\lambda u, (P_\lambda \mathfrak{a}) u) \\ &\quad - \frac{1}{2}(R_\lambda u, (P_\mu \mathfrak{a}) I_\mu u) - \frac{1}{2}(R_\lambda u, (\delta_\mu \mathfrak{a}) R_\mu u), \end{aligned} \quad (3.8)$$

$$|(I_\mu u, \mathfrak{a} \delta_\lambda u)| \leq N |u|_0^2 \quad (3.9)$$

for $\lambda, \mu \in \Lambda$, where N is a constant depending only on $K, d, |\lambda|$ and $|\mu|$.

Proof Notice that

$$\delta_\lambda^* = -\delta_\lambda, \quad I_\lambda^* = I_\lambda \quad \text{for } \lambda \in \Lambda$$

for the adjoints δ_λ^* and I_λ^* of δ_λ and I_λ in $L_2(\mathbb{R}^d)$, respectively. Using this, the Leibniz rule (3.3), and taking into account that δ_λ and I_μ commute, we have

$$(I_\mu u, \mathfrak{a} \delta_\lambda u) = -(\delta_\lambda (\mathfrak{a} I_\mu u), u) = -((\delta_\lambda \mathfrak{a}) I_\lambda I_\mu u, u) - A$$

with

$$A = ((I_\lambda \mathbf{a}) \delta_\lambda I_\mu u, u) = (\delta_\lambda u, I_\mu ((I_\lambda \mathbf{a}) u)).$$

Using (3.4) we get

$$A = h(\delta_\lambda u, I_\mu ((P_\lambda \mathbf{a}) u)) + B = (I_\mu R_\lambda u, ((P_\lambda \mathbf{a}) u)) + B$$

with

$$B = (\delta_\lambda u, I_\mu (\mathbf{a} u)),$$

and using (3.6) we obtain

$$\begin{aligned} B &= (\delta_\lambda u, \mathbf{a} I_\mu u) + h(\delta_\lambda u, (P_\mu \mathbf{a}) I_\mu u) + h(\delta_\lambda u, (\delta_\mu \mathbf{a}) R_\mu u) \\ &= (\delta_\lambda u, \mathbf{a} I_\mu u) + (R_\lambda u, (P_\mu \mathbf{a}) I_\mu u) + (R_\lambda u, (\delta_\mu \mathbf{a}) R_\mu u). \end{aligned}$$

Hence

$$\begin{aligned} (I_\mu, \mathbf{a} \delta_\lambda u) &= -((\delta_\lambda \mathbf{a}) I_\lambda I_\mu u, u) - (I_\mu R_\lambda u, (P_\lambda \mathbf{a}) u) \\ &\quad - (I_\mu u, \mathbf{a} \delta_\lambda u) - (R_\lambda u, (P_\mu \mathbf{a}) I_\mu u) - (R_\lambda u, (\delta_\mu \mathbf{a}) R_\mu u), \end{aligned} \quad (3.10)$$

which gives (3.8). To prove (3.9) notice that

$$|P_\lambda \mathbf{a}| \leq |\lambda| |D\mathbf{a}|, \quad |R_\lambda \mathbf{a}| \leq |\lambda| |D\mathbf{a}|,$$

where $|D\mathbf{a}|$ is the length of the gradient of $\mathbf{a}$. Notice also that due to the shift invariance of the Lebesgue measure, the linear operators I_μ and R_λ are contractions on H^0 . Hence it is easy to see that (3.8) implies (3.9). $\square$

Corollary 3.4 *Let m be a nonnegative integer and let $b = (b^r(x))_{r=1}^\infty$ be an l_2 -valued function on $\mathbb{R}^d$ such that b and its partial derivatives up to order $\max(m, 1)$ are l_2 -valued functions which in magnitude are bounded by K . Then for all $h \neq 0$, $u \in H^m$ and each $\alpha \in \Lambda^m$*

$$\sum_{r=1}^{\infty} \left| (\delta_\alpha I_\mu u, \delta_\alpha (b^r \delta_\lambda u)) \right|^2 \leq N |u|_m^4 \quad (3.11)$$

for $\lambda \in \Lambda_0$, $\mu \in \Lambda$, and $\alpha \in \Lambda^m$, where N is a constant depending only on K , d , m and $|\Lambda|$.

Proof Consider the case $m = 0$. Then by definition δ_α is the identity, and by (3.8) we have

$$\sum_{r=1}^{\infty} \left| (I_\mu u, b^r \delta_\lambda u) \right|^2 \leq 4 \sum_{i=1}^4 A_i$$

with

$$\begin{aligned} A_1 &= \sum_{r=1}^{\infty} \left| (u \delta_{\lambda} b^r, I_{\lambda} I_{\mu} u) \right|^2 \leq |u|_0^2 \sum_{r=1}^{\infty} \left| u \delta_{\lambda} b^r \right|_0^2 \leq N |u|_0^4, \\ A_2 &= \sum_{r=1}^{\infty} \left| (I_{\mu} R_{\lambda} u, u P_{\lambda} b^r) \right|^2 \leq |u|_0^2 \sum_{r=1}^{\infty} \left| u P_{\lambda} b^r \right|_0^2 \leq N |u|_0^4, \\ A_3 &= \sum_{r=1}^{\infty} \left| (R_{\lambda} u, (P_{\mu} b^r) I_{\mu} u) \right|^2 \leq |u|_0^2 \sum_{r=1}^{\infty} \left| (I_{\mu} u) P_{\mu} b^r \right|_0^2 \leq N |u|_0^4, \\ A_4 &= \sum_{r=1}^{\infty} \left| (R_{\lambda} u, (\delta_{\mu} b^r) R_{\mu} u) \right|^2 \leq |u|_0^2 \sum_{r=1}^{\infty} \left| (R_{\mu} u) \delta_{\mu} b^r \right|_0^2 \leq N |u|_0^4, \end{aligned}$$

where N is a constant depending only on m, d, K and $|\Lambda|$. This proves (3.11) when $m = 0$. Assume now $m \geq 1$. Then

$$\sum_{r=1}^{\infty} \left| (\delta_{\alpha} I_{\mu} u, \delta_{\alpha} (b^r \delta_{\lambda} u)) \right|^2 \leq 2B_1 + 2B_2,$$

where

$$\begin{aligned} B_1 &= \sum_{r=1}^{\infty} \left| (I_{\mu} u_{\alpha}, (I_{\alpha} b^r) \delta_{\lambda} u_{\alpha}) \right|^2, \\ B_2 &= \sum_{r=1}^{\infty} \left| (I_{\mu} u_{\alpha}, \delta_{\alpha} (b^r \delta_{\lambda} u) - (I_{\alpha} b^r) \delta_{\lambda} u_{\alpha}) \right|^2. \end{aligned}$$

Using that (3.11) holds when $m = 0$, we have

$$B_1 \leq N |\delta_{\alpha} u|_0^4 \leq N' |u|_m^4$$

with a constant $N' = N'(m, d, K, |\Lambda|)$. Using Lemma 3.1 and Lemma 3.2 we get

$$B_2 \leq |u_{\alpha}|_0^2 \sum_{\rho=1}^{\infty} \left| \delta_{\alpha} (b^{\rho} \delta_{\lambda} u) - (I_{\alpha} b^{\rho}) \delta_{\lambda} u_{\alpha} \right|^2 \leq N |u|_m^4$$

with a constant $N = N(m, d, K, |\Lambda|)$, and the proof is complete. $\square$

Lemma 3.5 *Let $m \geq 0$ be an integer and let $\mathfrak{p}$ be a nonnegative function on $\mathbb{R}^d$ such that its partial derivatives up to order $\max(m, 1)$ are functions, which together with $\mathfrak{p}$ are bounded in magnitude by a constant K . Then for every $\alpha \in \Lambda^m$, $\lambda \in \Lambda_0$ and for all $h \in (0, \infty)$ we have*

$$\left(\delta_{\alpha} u, \delta_{\alpha} (\mathfrak{p} \delta_{h, \lambda} u) \right) \leq N |u|_m^2$$

for all $u \in H^m$, where N is a constant depending only on $K, d, |\Lambda|$ and m .

Proof Clearly

$$\delta_\alpha(\mathbf{p}\delta_{h,\lambda}u) = \delta_\alpha(\mathbf{p}\delta_{h,\lambda}u) - (I_\alpha\mathbf{p})\delta_\alpha\delta_{h,\lambda}u + (I_\alpha\mathbf{p})\delta_{h,\lambda}\delta_\alpha u,$$

and

$$\left(\delta_\alpha u, \delta_\alpha(\mathbf{p}\delta_{h,\lambda}u) - (I_\alpha\mathbf{p})\delta_\alpha\delta_{h,\lambda}u \right) \leq N|u|_m^2$$

with a constant $N = N(m, K, d, \Lambda)$. By the Leibniz rule (3.2) we have

$$\left(\delta_\alpha u \right) \delta_{h,\lambda} \delta_\alpha u = \frac{1}{2} \delta_{h,\lambda} \left(\delta_\alpha u \right)^2 - h_\lambda \left(\delta_{h,\lambda} \delta_\alpha u \right)^2.$$

Consequently,

$$\begin{aligned} \left(\delta_\alpha u, \delta_\alpha(\mathbf{p}\delta_{h,\lambda}u) \right) &\leq N|u|_m^2 + (I_\alpha\mathbf{p}, (\delta_\alpha u)\delta_{h,\lambda}\delta_\alpha u) \\ &= N|u|_m^2 + \frac{1}{2}(I_\alpha\mathbf{p}, \delta_{h,\lambda}(\delta_\alpha u)^2) - h_\lambda(I_\alpha\mathbf{p}, (\delta_{h,\lambda}\delta_\alpha u)^2). \end{aligned} \quad (3.12)$$

Due to $I_\alpha\mathbf{p} \geq 0$ and $h_\lambda \geq 0$ we have

$$h_\lambda(I_\alpha\mathbf{p}, (\delta_{h,\lambda}\delta_\alpha u)^2) \geq 0,$$

and by taking the adjoint $\delta_{h,\lambda}^* = \delta_{h,-\lambda}$ in L_2 , we have

$$\left(I_\alpha\mathbf{p}, \delta_{h,\lambda}(\delta_\alpha u)^2 \right) = \left(\delta_{h,-\lambda}I_\alpha\mathbf{p}, (\delta_\alpha u)^2 \right) \leq N|u|_m^2$$

with a constant $N = N(K, m, d, |\Lambda|)$. Thus the lemma follows from (3.12). $\square$

Lemma 3.6 *The following statements hold for all $h > 0$ and $\lambda, \mu \in \Lambda_0$.*

(i) *For $v, \varphi \in H^1$ we have*

$$|(\delta_\lambda^h v, \varphi) - (\partial_\lambda v, \varphi)| \leq \frac{h}{2} |\lambda|^2 |Dv|_0 |D\varphi|_0. \quad (3.13)$$

(ii) *Let $b = (b^r(x))_{r=1}^\infty$ be an l_2 -valued function on $\mathbb{R}^d$, with l_2 norm bounded by a constant K . Assume that its first order derivatives in x are l_2 -valued functions, which are in l_2 norm bounded by K . Then there is a constant $N = N(K, d)$ such that for $v, \varphi \in H^1$*

$$\sum_{r=1}^\infty |(b^r \delta_\lambda^h v, \varphi) - (b^r \partial_\lambda v, \varphi)|^2 \leq N h^2 |\lambda|^4 |Dv|_0^2 |\varphi|_1^2.$$

(iii) Let $a = a(x)$ be a real function on $\mathbb{R}^d$, bounded by a constant K , such that its derivatives up to second order are functions, bounded by K . Then for all $v \in H^1$ and $\varphi \in H^2$ we have

$$\begin{aligned} |(a\delta_\mu^h\delta_\lambda^h v, \varphi) - (\partial_\lambda v, \partial_{-\mu}(a\varphi))| &\leq \frac{h}{2}|\lambda|(|\lambda| + |\mu|)|Dv|_0|D^2(a\varphi)|_0 \\ &\leq Nh|Dv|_0|\varphi|_2 \end{aligned} \quad (3.14)$$

with a constant $N = N(K, d, |\lambda|, |\mu|)$.

Proof It is sufficient to prove the lemma for $v, \varphi \in C_0^\infty(\mathbb{R}^d)$. For such v and φ we have

$$(\delta_\lambda^h v - \partial_\lambda v, \varphi) = \frac{h}{4} \int_{-1}^1 \int_{-1}^1 \int_{\mathbb{R}^d} \theta_1 \partial_\lambda^2 v(x + \theta_1 \theta_2 h \lambda) \varphi(x) dx d\theta_1 d\theta_2.$$

Hence by integration by parts, using the Bunyakovsky-Cauchy-Schwarz inequality and the shift invariance of the Lebesgue measure, we obtain

$$|(\delta_\lambda^h v - \partial_\lambda v, \varphi)| \leq \frac{h}{4} \int_{-1}^1 \int_{-1}^1 |\theta_1| |\partial_\lambda v|_0 |\partial_\lambda \varphi|_0 d\theta_1 d\theta_2 = \frac{h}{2} |\partial_\lambda v|_0 |\partial_\lambda \varphi|_0,$$

which proves (3.13). Hence we can easily obtain (ii) by substituting $b^r \varphi$ in place of φ in (3.13), and then by taking the square and summing up the inequalities over $r \in \{1, 2, \dots\}$. To prove (iii) notice that

$$(a\delta_\mu^h\delta_\lambda^h v, \varphi) - (\partial_\lambda v, \partial_{-\mu}(a\varphi)) = A + B$$

with

$$\begin{aligned} A &= (v, \delta_\mu^h\delta_\lambda^h(a\varphi)) - (v, \partial_\mu\delta_\lambda^h(a\varphi)) \\ B &:= (v, \delta_\lambda^h\partial_\mu(a\varphi)) - (v, \partial_\lambda\partial_\mu(a\varphi)). \end{aligned}$$

By virtue of (i) we have

$$\begin{aligned} |A| &\leq \frac{h}{2} |\mu|^2 |Dv|_0 |D\delta_\lambda(a\varphi)|_0 \leq \frac{h}{2} |\mu|^2 |\lambda| |Dv|_0 |D^2(a\varphi)|_0, \\ |B| &\leq \frac{h}{2} |\lambda|^2 |Dv|_0 |DD_\mu(a\varphi)|_0 \leq \frac{h}{2} |\lambda|^2 |\mu| |Dv|_0 |D^2(a\varphi)|_0, \end{aligned}$$

and by adding up these estimates we get (3.14). $\square$

Next we formulate a generalisation of a well-known lemma from [4].

Lemma 3.7 Let $\sigma(x)$ be a $d \times m$ matrix for every $x \in \mathbb{R}^d$ such that

$$|\sigma(x) - \sigma(y)| \leq L|x - y| \quad \text{for all } x, y \in \mathbb{R}^d.$$

Let V be a symmetric $d \times d$ matrix. Then the following estimates hold.

(i) For every $\lambda \in \mathbb{R}^d \setminus \{0\}$

$$\left| \partial_\lambda (\sigma \sigma^*)^{ij} V^{ij} \right|^2 \leq 4L^2 |\lambda| (\sigma \sigma^*)^{ij} V^{ik} V^{jk} \quad \text{for } dx - \text{almost every } x \in \mathbb{R}^d;$$

(ii) For every $\lambda \in \Lambda_0$, $h \neq 0$, $\varepsilon > 0$ and $x \in \mathbb{R}^d$ we have

$$\left| \delta_{h,\lambda} (\sigma \sigma^*)^{ij} V^{ij} \right|^2 \leq 4(1 + \varepsilon) L^2 |\lambda| (\sigma \sigma^*)^{ij} V^{ik} V^{jk} + (1 + \varepsilon^{-1}) L^4 |\lambda|^4 h^2 |V|^2;$$

(iii) For every $\lambda \in \Lambda$, $h \neq 0$ and $x \in \mathbb{R}^d$ we have

$$\left| \delta_\lambda^h (\sigma \sigma^*)^{ij} V^{ij} \right|^2 \leq 5L^2 |\lambda| (\sigma \sigma^*)^{ij} V^{ik} V^{jk} + 5L^4 |\lambda|^4 h^2 |V|^2.$$

Proof Part (i) is well-known from [4], and (iii) obviously follows from (ii). To prove (ii) we use the discrete Leibniz rule to get

$$\delta_{h,\lambda} (\sigma \sigma^*) = \sigma_\lambda \sigma^* + \sigma \sigma_\lambda^* + h \sigma_\lambda \sigma_\lambda^*,$$

where $\sigma_\lambda := \delta_{h,\lambda} \sigma$. Hence using the Cauchy inequality and the simple inequality

$$(a + b)^2 \leq (1 + \varepsilon) a^2 + (1 + \varepsilon^{-1}) b^2,$$

we have

$$\begin{aligned} \left| \delta_{h,\lambda} (\sigma \sigma^*)^{ij} V^{ij} \right|^2 &= \left| 2\sigma_\lambda^{ik} \sigma^{jk} V^{ij} + h \sigma_\lambda^{ik} \sigma_\lambda^{jk} V^{ij} \right|^2 \\ &\leq 4(1 + \varepsilon) |\sigma_\lambda|^2 (\sigma \sigma^*)^{jl} V^{ji} V^{li} + (1 + \varepsilon^{-1}) h^2 |\sigma_\lambda|^4 |V|^2, \end{aligned}$$

and obtain (ii) by taking into account $|\sigma_\lambda| \leq L\lambda$. □

Finally we present a stochastic Gronwall lemma from [21], which improves Lemma 3.7 from [22].

Lemma 3.8 *Let $y = (y_t)_{t \in [0, T]}$ and $F = (F_t)_{t \in [0, T]}$ be adapted nonnegative stochastic processes and let $m = (m_t)_{t \in [0, T]}$ be a continuous local martingale such that almost surely*

$$dy_t \leq N(y_t + F_t) dt + dm_t \quad \text{for all } t \in [0, T],$$

where N is a constant. Assume that

$$d\langle m \rangle_t \leq N(y_t^2 + G_t y_t) dt$$

for a nonnegative stochastic process $G = (G_t)_{t \in [0, T]}$ and a constant N . Then for every $p > 0$

$$E \sup_{t \leq T} |y_t|^p \leq CE|y_0|^p + CE \left\{ \int_0^T (F_t + G_t) dt \right\}^p,$$

where C is a constant depending only on p , N and T .

For the proof we refer to [21].

4 Solvability of the finite difference scheme

In this section we study the finite difference scheme (2.3)–(2.4) in the whole $\mathbb{R}^d$ instead of $\mathbb{G}_h$, i.e., we consider

$$\begin{aligned} du_t(x) = & \left(L_t^h u_t(x) + f_t(x) \right) dt \\ & + \left(M_t^{h,r} u_t(x) + g_t^r(x) \right) dw_t^r \quad \text{for } (t, x) \in H_T, \end{aligned} \quad (4.1)$$

with initial condition

$$u_0(x) = \psi(x), \quad \text{for } x \in \mathbb{R}^d, \quad (4.2)$$

where L^h and $M^{h,r}$ are defined in (2.5) and (2.6).

Let, as before, m be a nonnegative integer, $K \geq 0$ be a constant, and make the following assumptions, which are somewhat weaker than Assumptions 2.3 and 2.4.

Assumption 4.1 The functions $a^{\lambda\mu}$, b^λ , p^γ and q^γ and their partial derivatives in $x \in \mathbb{R}^d$ up to order m are continuous in x and are bounded in magnitude by K , for all $\lambda, \mu \in \Lambda_1$ and $\gamma \in \Lambda_0$.

Assumption 4.2 The initial value ψ is an H^m -valued $\mathcal{F}_0$ -measurable random variable, and $f = (f_t)_{t \geq 0}$ and $g = (g_t)_{t \geq 0}$ are predictable processes with values in H^m and $H^m(l_2)$, respectively, such that almost surely

$$\int_0^T \left| f_t \right|_m^2 + \left| g_t \right|_m^2 dt < \infty. \quad (4.3)$$

Definition 4.1 An H^0 -valued continuous and adapted process $(\bar{u}_t^h)_{t \in [0, T]}$ is called a solution to (4.1)–(4.2) if almost surely

$$(\bar{u}_t^h, \varphi) = (\psi, \varphi) + \int_0^t \left(L_s^h \bar{u}_s^h + f_s, \varphi \right) ds + \int_0^t \left(M_s^{h,r} \bar{u}_s^h + g_s^r, \varphi \right) dw_s^r \quad (4.4)$$

for all $t \in [0, T]$ and $\varphi \in C_0^\infty(\mathbb{R}^d)$.

Note that due to Assumption 4.1 for each $h \neq 0$ we have a constant C such that

$$\left| L_t^h \phi \right|_l^2 + \sum_{r=1}^{\infty} \left| M_t^{h,r} \phi \right|_l^2 \leq C^2 |\phi|_l^2 \quad \text{for all } \phi \in H^l, \quad (4.5)$$

for all $t \in [0, T]$ and integer $l \in [0, m]$. Thus L_t^h and $M_t^h = (M_t^{h,r})_{r=1}^{\infty}$ are bounded linear operators from H^l to H^l and to $H^l(l_2)$, respectively, such that their operator norm is bounded by C for all $t \in [0, T]$ and $\omega \in \Omega$. Using this with $l = 0$, (4.3) with $m = 0$ and assuming that ψ is a $\mathcal{F}_0$ measurable random variable in H^0 , we can see that $\bar{u}$ is a solution to (4.1)–(4.2) if and only if it is the solution of the SDE

$$\bar{u}_t^h = \psi + \int_0^t (L_s^h \bar{u}_s^h + f_s) ds + \int_0^t (M_s^{h,r} \bar{u}_s^h + g_s^r) dw_s^r, \quad t \in [0, T], \quad (4.6)$$

in the Hilbert space H^0 , where the first integral is a Bochner integral and the second one is a stochastic integral in a Hilbert space. Thus by a well-known theorem on SDEs in Hilbert spaces, with Lipschitz continuous coefficients, problem (4.1)–(4.2) admits a unique solution $\bar{u}^h$. Due to $\psi \in H^m$, (4.3) and (4.5) with $l = m$, Eq. (4.6) admits a unique continuous H^m -valued solution by virtue of the same theorem on solvability of SDEs in Hilbert spaces with Lipschitz continuous coefficients. Consequently, $\bar{u}^h$ is a continuous H^m -valued process. If $m > d/2$ then by Sobolev's theorem on embedding H^m into $C_b(\mathbb{R}^d)$, the space of continuous and bounded functions on $\mathbb{R}^d$, there exists a linear operator $J : H^m \rightarrow C_b$ such that $J\varphi(x) = \varphi(x)$ for almost every $x \in \mathbb{R}^d$, and

$$\sup_{\mathbb{R}^d} |J\varphi| \leq N|\varphi|_m$$

for all $\varphi \in H^m$, where N is a constant depending only on d . One has also the following lemma on the embedding $H^m \subset l_2(\mathbb{G}_h)$.

Lemma 4.1 *Let Assumption 2.1 hold. Then for $\varphi \in H^m$, $m > d/2$ and $h \in (0, 1)$ we have*

$$\sum_{x \in \mathbb{G}_h} \left| J\varphi(x) \right|^2 h^d \leq N \left| \varphi \right|_l^2, \quad (4.7)$$

where N is a constant depending only on d and Λ_0 .

Proof This lemma is a straightforward consequence of Sobolev's theorem on embedding W_2^m functions on the unit ball B_1 of $\mathbb{R}^d$ into $C(B_1)$, the space of continuous functions on B_1 . (See, e.g., [2]). $\square$

If Assumption 2.1 holds and $m > d/2$ then by virtue of the above lemma the solution $(J\bar{u}_t^h)_{t \in [0, T]}$ of (4.1)–(4.2) restricted to $\mathbb{G}_h$ in x is a continuous $l_{h,2}$ -valued process. Thus we have the following proposition. Remember that when f and g are H^m and $H^m(l_2)$ -valued processes with $m > d/2$, and $\psi \in H^m$, then we always take their continuous modifications in x , i.e., we take $\hat{f} = Jf$, $\hat{g} = Jg$ and $\hat{\psi} = J\psi$ in place of f , g and ψ , respectively.

Proposition 4.2 *Let Assumptions 4.1 and 4.2 hold. Then (4.1)–(4.2) has a unique solution $\bar{u}^h = (\bar{u}_t^h)_{t \in [0, T]}$ for each $h \neq 0$ in the sense of Definition 4.1. Moreover, $\bar{u}^h$ is a continuous H^m -valued process. If $m > d/2$ then $\hat{u}^h := J\bar{u}^h$, the continuous modification in x of $\bar{u}^h$, restricted to $\mathbb{G}_h$ is u^h , the continuous $l_{h,2}$ -valued solution of (2.3)–(2.4).*

Proof Except of the last statement we have already proved this proposition. To prove the last statement assume $m > d/2$. Fix a point x of $\mathbb{G}_h$ and take a nonnegative smooth function φ with compact support in $\mathbb{R}^d$ whose integral over $\mathbb{R}^d$ is one. Define for each integer $n \geq 1$ the function φ_n by $\varphi_n(y) = n^d \varphi(n(y - x))$ for $y \in \mathbb{R}^d$. Then by virtue of Definition 4.1 we have almost surely

$$(\hat{u}_t^h, \varphi_n) = (\hat{\psi}, \varphi_n) + \int_0^t (L_t^h \hat{u}_t^h + \hat{f}_t, \varphi_n) dt + \int_0^t (M_t^{h,r} \hat{u}_t^h + \hat{g}_t^r, \varphi_n) dw_t^r$$

for all $t \in [0, T]$ and for all $n \geq 1$. Letting here $n \rightarrow \infty$ we obtain

$$\hat{u}_t^h(x) = \hat{\psi}(x) + \int_0^t (L_t^h \hat{u}_t^h(x) + \hat{f}_t(x)) dt + \int_0^t (M_t^{h,r} \hat{u}_t^h(x) + \hat{g}_t^r(x)) dw_t^r$$

almost surely for each $t \in [0, T]$, where almost surely the right-hand side is continuous in $t \in [0, T]$. Since $\hat{u}_t^h(x)$ is also continuous in t , we have this equality almost surely for all $t \in [0, T]$ and all $x \in \mathbb{G}_h$. Moreover, we know that $(\hat{u}^h)_{t \in [0, T]}$ is a continuous and adapted l_2 -valued process, and that the l_2 -valued solution u^h of (2.3)–(2.4) is unique. Hence $\hat{u}^h = u^h$. $\square$

Our aim now is to obtain an estimate for the solution $\bar{u}^h$ to (4.1)–(4.2) independently of h . To this end first for $u \in H^m$, $f \in H^m$ and $g = (g^r)_{r=1}^\infty \in H^{m+1}(l_2)$ we set

$$\begin{aligned} \mathcal{Q}_m^h(u, f, g, t) &= \int_{\mathbb{R}^d} \sum_{\alpha \in \Lambda^m} 2(\delta_\alpha^h u) \delta_\alpha^h (L_t^h u(x) + f(x)) dx \\ &\quad + \int_{\mathbb{R}^d} \sum_{\alpha \in \Lambda^m} \sum_r |\delta_\alpha^h (M_t^{h,r} u(x) + g^r(x))|^2 dx, \end{aligned} \quad (4.8)$$

and prove the following lemma.

Lemma 4.3 *Let Assumptions 2.3 and 2.5 hold. Then for each $h > 0$ for $P \otimes dt$ -almost all $(\omega, t) \in \Omega \times [0, T]$ we have*

$$\mathcal{Q}_m^h(u, f, g, t) \leq N \left(|u|_m^2 + |f|_m^2 + |g|_{m+1}^2 \right) \quad (4.9)$$

for all $u, f \in H^m$ and $g \in H^{m+1}(l_2)$, where N is a constant depending only on d, m, K and $|\Lambda|$. If in addition to the above assumptions $\mathfrak{p}^\lambda = \mathfrak{q}^\lambda = 0$ for all $\lambda \in \Lambda_0$, then (4.9) holds for each $h \in \mathbb{R} \setminus \{0\}$ for $P \otimes dt$ -almost every (ω, t) .

Proof For real functions v and w on $\mathbb{R}^d$ we write $v \sim w$ if their integrals over $\mathbb{R}^d$ are equal. We write $v \leq w$ if $v = w + F$ for a function F on $\mathbb{R}^d$ such that the integral of F over $\mathbb{R}^d$ can be estimated by the right-hand side of (4.9). We use the notation $v_\lambda = \delta_\lambda v$ for functions v on $\mathbb{R}^d$ and for $\lambda \in \Lambda^k$, $k \geq 1$. To simplify the notation we often write δ_α , Δ_λ and I_α in place of Δ_α^h , δ_λ^h and I_α^h , respectively. Moreover, we often use the convention that if there is no parenthesis to indicate the order of operations in a formula then operators act only on the first function written after them. For example, $\alpha \delta_\mu u I_\lambda f u$ means $\alpha(\delta_\mu u)(I_\lambda f)u$.

Assume first that $p^\lambda = q^\lambda = 0$ for all $\lambda \in \Lambda_0$. Consider the case $m = 0$. Notice that for $\lambda, \mu \in \Lambda_0$ we have

$$u \alpha^{\lambda\mu} \delta_\lambda \delta_\mu u \sim -I_\lambda u \alpha^{\lambda\mu}_\lambda u_\mu - u_\lambda I_\lambda \alpha^{\lambda\mu} u_\mu.$$

By Lemma 3.3 we get $-I_\lambda u \alpha^{\lambda\mu}_\lambda u_\mu \leq 0$, and using $I_\lambda = I + h^2 \Delta_\lambda / 2$ we have

$$-u_\lambda I_\lambda \alpha^{\lambda\mu} u_\mu = -u_\lambda \alpha^{\lambda\mu} u_\mu - \frac{1}{2} T_\lambda u \left(\Delta_\lambda \alpha^{\lambda\mu} \right) T_\mu u \leq -u_\lambda \alpha^{\lambda\mu} u_\mu.$$

Using Lemma 3.3 for $\lambda \in \Lambda_0$ we have $u \alpha^{\lambda 0} u_\lambda \leq 0$. Hence we have

$$2u \sum_{\lambda, \mu \in \Lambda_1} \alpha^{\lambda\mu} \delta_\lambda \delta_\mu u \leq -2 \sum_{\lambda, \mu \in \Lambda_0} u_\lambda \alpha^{\lambda\mu} u_\mu. \quad (4.10)$$

Notice that by Lemma 3.3 we have $b^{\lambda r} u_\lambda b^{0r} u \leq 0$, and

$$b^{\lambda r} u_\lambda g^r \sim -b^{\lambda r}_\lambda \left(I_\lambda g^r \right) u - \left(I_\lambda b^{\lambda r} \right) g^{\lambda r}_\lambda u \leq 0.$$

Hence

$$\sum_r \left| M^{h,r} u + g^r \right|^2 \leq \sum_{\lambda, \mu \in \Lambda_0} \sum_r b^{\lambda r} b^{\mu r} u_\lambda u_\mu,$$

which together with (4.10) gives

$$2u \left(L^h u + f \right) + \sum_r \left| M^{h,r} u + g^r \right|^2 \leq - \sum_{\lambda, \mu \in \Lambda_0} \left(2\alpha^{\lambda\mu} - \sum_r b^{\lambda r} b^{\mu r} \right) u_\lambda u_\mu \leq 0$$

by virtue of Assumption 2.5. This proves Lemma 4.3 for $m = 0$.

Consider now the case $m \geq 1$. Let $\alpha \in \Lambda^m$. Then by Lemmas 3.1 and 3.3 for $\lambda \in \Lambda_0$ we have

$$u_\alpha \delta_\alpha \left(\alpha^{\lambda 0} u_\lambda \right) \leq u_\alpha \left(I_\alpha \alpha^{\lambda 0} \right) \delta_\lambda u_\alpha \leq 0.$$

Let $\lambda, \mu \in \Lambda_0$. Then using Lemma 3.1 we have

$$\begin{aligned} u_\alpha \delta_\alpha \left(\mathfrak{a}^{\lambda\mu} \delta_\lambda \delta_\mu u \right) &\leq u_\alpha \sum_{\gamma \leq \alpha, |\gamma|=1} \left(I_{\alpha \setminus \gamma} \mathfrak{a}_\gamma^{\lambda\mu} \right) I_\gamma \delta_{\alpha \setminus \gamma} \delta_\lambda \delta_\mu u \\ &\quad + u_\alpha \left(I_\alpha \mathfrak{a}^{\lambda\mu} \right) \delta_\lambda \delta_\mu u_\alpha. \end{aligned}$$

By virtue of (3.5) we get

$$\begin{aligned} \left(I_\alpha \mathfrak{a}^{\lambda\mu} \right) \delta_\lambda \delta_\mu u_\alpha &= \mathfrak{a}^{\lambda\mu} \delta_\lambda \delta_\mu u_\alpha + h^2 \left(\mathcal{O}_\alpha \mathfrak{a}^{\lambda\mu} \right) \delta_\lambda \delta_\mu u_\alpha \\ &= \mathfrak{a}^{\lambda\mu} \delta_\lambda \delta_\mu u_\alpha + \left(\mathcal{O}_\alpha \mathfrak{a}^{\lambda\mu} \right) R_\lambda R_\mu u_\alpha. \end{aligned}$$

Thus, using also Lemma 3.3, we obtain

$$\begin{aligned} u_\alpha \left(I_\alpha \mathfrak{a}^{\lambda\mu} \right) \delta_\lambda \delta_\mu u_\alpha &\leq u_\alpha \mathfrak{a}^{\lambda\mu} \delta_\lambda \delta_\mu u_\alpha \sim -u_{\alpha\lambda} \left(I_\lambda \mathfrak{a}^{\lambda\mu} \right) u_{\alpha\mu} - (I_\lambda u_\alpha) \mathfrak{a}_\lambda^{\lambda\mu} \delta_\mu u_\alpha \\ &\leq -u_{\alpha\lambda} \mathfrak{a}^{\lambda\mu} u_{\alpha\mu}. \end{aligned}$$

For $|\gamma| = 1$ and $|\beta| \leq m - 1$ due to (3.5) and (3.6) we have

$$\begin{aligned} u_\alpha \left(I_\beta \mathfrak{a}_\gamma^{\lambda\mu} \right) I_\gamma \delta_\beta \delta_\lambda \delta_\mu u &= u_\alpha \mathfrak{a}_\gamma^{\lambda\mu} I_\gamma \delta_\lambda \delta_\mu u_\beta + h u_\alpha \left(\mathcal{P}_\beta \mathfrak{a}_\gamma^{\lambda\mu} \right) I_\gamma \delta_\lambda \delta_\mu u_\beta \\ &= u_\alpha \mathfrak{a}_\gamma^{\lambda\mu} I_\gamma \delta_\lambda \delta_\mu u_\beta + u_\alpha \left(\mathcal{P}_\beta \mathfrak{a}_\gamma^{\lambda\mu} \right) I_\gamma R_\lambda \delta_\mu u_\beta \\ &\leq u_\alpha \mathfrak{a}_\gamma^{\lambda\mu} I_\gamma \delta_\lambda \delta_\mu u_\beta \sim \left(I_\gamma (u_\alpha \mathfrak{a}_\gamma^{\lambda\mu}) \right) \delta_\lambda \delta_\mu u_\beta \\ &= (I_\gamma u_\alpha) \mathfrak{a}_\gamma^{\lambda\mu} \delta_\lambda \delta_\mu u_\beta + \left(I_\gamma u_\alpha \right) \left(P_\gamma \mathfrak{a}_\gamma^{\lambda\mu} \right) R_\lambda \delta_\mu u_\beta \\ &\quad + \left(R_\gamma u_\alpha \right) \left(\delta_\gamma \mathfrak{a}_\gamma^{\lambda\mu} \right) R_\lambda \delta_\mu u_\beta \\ &\leq \left(I_\gamma u_\alpha \right) \mathfrak{a}_\gamma^{\lambda\mu} \delta_\lambda \delta_\mu u_\beta. \end{aligned}$$

Hence

$$A := 2\delta_\alpha^h u \delta_\alpha^h (L^h u + f) \leq A_1 + A_2 \quad (4.11)$$

with

$$A_1 = -2u_{\alpha\lambda} \mathfrak{a}^{\lambda\mu} u_{\alpha\mu}, \quad A_2 = 2 \sum_{\gamma \leq \alpha, |\gamma|=1} \left(I_\gamma u_\alpha \right) \mathfrak{a}_\gamma^{\lambda\mu} \delta_\lambda \delta_\mu u_{\alpha \setminus \gamma},$$

where the summation convention is used with respect to repeated λ, μ from Λ_0 . By Lemma 3.5 for $h > 0$ we have

$$\left(\delta_\alpha u, \delta_\alpha (\mathfrak{p}^\lambda \delta_{h,\lambda} u - \mathfrak{q}^\lambda \delta_{-h,\lambda} u) \right) \leq 0. \quad (4.12)$$

Clearly

$$\begin{aligned} |\delta_\alpha(\mathfrak{b}^\mu \delta_\mu u + g)|_{l_2}^2 &= |\delta_\alpha(\mathfrak{b}^\lambda \delta_\lambda u)|_{l_2}^2 + |\delta_\alpha(g + \mathfrak{b}^0 u)|_{l_2}^2 \\ &\quad + 2(\delta_\alpha(\mathfrak{b}^\lambda \delta_\lambda u), \delta_\alpha(g + \mathfrak{b}^0 u))_{l_2}, \end{aligned}$$

where repeated indices μ mean summation over $\mu \in \Lambda_1$ and repeated indices λ mean summation over $\lambda \in \Lambda_0$. It is easy to see that

$$\begin{aligned} |\delta_\alpha(g + \mathfrak{b}^0 u)|_{l_2}^2 &\leq 0, \\ (\delta_\alpha(\mathfrak{b}^\lambda \delta_\lambda u), \delta_\alpha(g + \mathfrak{b}^0 u))_{l_2} &\leq (\delta_\alpha(\mathfrak{b}^\lambda \delta_\lambda u), \delta_\alpha(\mathfrak{b}^0 u))_{l_2}. \end{aligned}$$

Using the Lemmas 3.1 and 3.3 we get

$$\left(\delta_\alpha(\mathfrak{b}^\lambda \delta_\lambda u), \delta_\alpha(\mathfrak{b}^0 u) \right)_{l_2} \leq \left((I_\alpha \mathfrak{b}^\lambda) I_\alpha (\delta_\lambda \mathfrak{b}^0) u_\alpha, u_\alpha \right)_{l_2} \leq 0.$$

Hence

$$B := |\delta_\alpha(\mathfrak{b}^\lambda \delta_\lambda u + g)|_{l_2}^2 \leq |\delta_\alpha v(\mathfrak{b}^\lambda \delta_\lambda u)|_{l_2}^2 = B_0 + B_1 + B_2 \quad (4.13)$$

with

$$\begin{aligned} B_0 &= |\delta_\alpha(\mathfrak{b}^\lambda \delta_\lambda u) - (I_\alpha \mathfrak{b}^\lambda) \delta_\lambda u_\alpha|_{l_2}^2, \quad B_1 = |(I_\alpha \mathfrak{b}^\lambda) \delta_\lambda u_\alpha|_{l_2}^2, \\ B_2 &= 2(\delta_\alpha(\mathfrak{b}^\lambda \delta_\lambda u) - (I_\alpha \mathfrak{b}^\lambda) \delta_\lambda u_\alpha, (I_\alpha \mathfrak{b}^\lambda) \delta_\lambda u_\alpha)_{l_2}. \end{aligned}$$

It is easy to notice that

$$B_0 \leq 0, \quad (4.14)$$

$$\begin{aligned} B_1 &\leq \left(I_\alpha \mathfrak{b}^{\lambda r} \right) \left(I_\alpha \mathfrak{b}^{\mu r} \right) \delta_\lambda u_\alpha \delta_\mu u_\alpha, \\ B_2 &\leq 2 \sum_\gamma \left(I_\beta \mathfrak{b}_\gamma^{\lambda r} \right) \left(I_\alpha \mathfrak{b}^{\mu r} \right) \left(I_\gamma \delta_\lambda u_\beta \right) \delta_\mu u_\alpha, \end{aligned} \quad (4.15)$$

where $\sum_\gamma$ in B_2 denotes summation over all sub-sequences γ of α which have length 1, and β denotes $\alpha \setminus \gamma$. Here, and later on in the proof, λ and μ are from Λ_0 . By (3.5) we have

$$\begin{aligned} \left(I_\alpha \mathfrak{b}^{\lambda r} \right) \left(I_\alpha \mathfrak{b}^{\mu r} \right) u_{\alpha\lambda} u_{\alpha\mu} &= \left(\mathfrak{b}^{\lambda r} + h^2 \mathcal{O}_\alpha \mathfrak{b}^{\lambda r} \right) \left(\mathfrak{b}^{\mu r} + h^2 \mathcal{O}_\alpha \mathfrak{b}^{\mu r} \right) u_{\alpha\lambda} u_{\alpha\mu} \\ &= \mathfrak{b}^{\lambda r} \mathfrak{b}^{\mu r} u_{\alpha\lambda} u_{\alpha\mu} + \left(\mathfrak{b}^{\lambda r} \mathcal{O}_\alpha \mathfrak{b}^{\mu r} + \mathfrak{b}^{\mu r} \mathcal{O}_\alpha \mathfrak{b}^{\lambda r} \right) \left(R_\lambda u_\alpha \right) \left(R_\mu u_\alpha \right) \\ &\quad + \left(\mathcal{P}_\alpha \mathfrak{b}^{\mu r} \right) \left(\mathcal{P}_\alpha \mathfrak{b}^{\lambda r} \right) \left(R_\lambda u_\alpha \right) \left(R_\mu u_\alpha \right) \leq \mathfrak{b}^{\lambda r} \mathfrak{b}^{\mu r} u_{\alpha\lambda} u_{\alpha\mu}. \end{aligned}$$

Consequently,

$$B_1 \leq \sum_{\lambda, \mu \in \Lambda_0} \mathfrak{b}^{\lambda r} \mathfrak{b}^{\mu r} u_{\alpha\lambda} u_{\alpha\mu}. \quad (4.16)$$

By virtue of (3.5) we have

$$\begin{aligned}
 (I_\beta \mathfrak{b}_\gamma^{\lambda r})(I_\alpha \mathfrak{b}^{\mu r})(I_\gamma u_{\beta\lambda})u_{\alpha\mu} &= ((I + h\mathcal{P}_\beta)\mathfrak{b}_\gamma^{\lambda r})((I + h\mathcal{P}_\alpha)\mathfrak{b}^{\mu r})(I_\gamma u_{\beta\lambda})u_{\alpha\mu} \\
 &= \mathfrak{b}_\gamma^{\lambda r}\mathfrak{b}^{\mu r}(I_\gamma u_{\beta\lambda})u_{\alpha\mu} \\
 &\quad + (\mathfrak{b}_\gamma^{\lambda r}(\mathcal{P}_\alpha \mathfrak{b}^{\mu r}) + (\mathcal{P}_\beta \mathfrak{b}_\gamma^{\lambda r})\mathfrak{b}^{\mu r})(I_\gamma u_{\beta\lambda})R_\mu u_\alpha \\
 &\quad + (\mathcal{P}_\beta \mathfrak{b}_\gamma^{\lambda r})(\mathcal{P}_\alpha \mathfrak{b}^{\mu r})(I_\gamma R_\lambda u_\beta)R_\mu u_\alpha \\
 &\leq \mathfrak{b}_\gamma^{\lambda\rho}\mathfrak{b}^{\mu r}(I_\gamma u_{\beta\lambda})u_{\alpha\mu}
 \end{aligned}$$

for $\lambda, \mu \in \Lambda_0$. Thus from (4.15) we obtain

$$B_2 \leq 2 \sum_\gamma \sum_{\lambda, \mu \in \Lambda_0} \mathfrak{b}_\gamma^{\lambda\rho}\mathfrak{b}^{\mu r}(I_\gamma u_{\beta\lambda})u_{\alpha\mu}.$$

Hence using that by (3.4)

$$2\mathfrak{b}_\gamma^{\lambda r}\mathfrak{b}^{\mu r} = 2\mathfrak{b}_\gamma^{\lambda r}I_\gamma \mathfrak{b}^{\mu r} - 2h\mathfrak{b}_\gamma^{\lambda r}P_\gamma \mathfrak{b}^{\mu r} = \delta_\gamma(\mathfrak{b}^{\lambda r}\mathfrak{b}^{\mu r}) - 2h\mathfrak{b}_\gamma^{\lambda r}P_\gamma \mathfrak{b}^{\mu r},$$

we get

$$\begin{aligned}
 B_2 &\leq \sum_\gamma \sum_{\lambda, \mu \in \Lambda_0} \delta_\gamma(\mathfrak{b}^{\lambda r}\mathfrak{b}^{\mu r})(I_\gamma u_{\beta\lambda})u_{\alpha\mu} \\
 &\quad - 2 \sum_\gamma \sum_{\lambda, \mu \in \Lambda_0} \mathfrak{b}_\gamma^{\lambda r}(P_\gamma \mathfrak{b}^{\mu r})(I_\gamma u_{\beta\lambda})R_\mu u_\alpha \\
 &\leq \sum_\gamma \sum_{\lambda, \mu \in \Lambda_0} (\delta_\mu u_\alpha)(\delta_\gamma(\mathfrak{b}^{\lambda r}\mathfrak{b}^{\mu r}))I_\gamma u_{\beta\lambda}. \tag{4.17}
 \end{aligned}$$

Taking the adjoint $\delta_\mu^* = -\delta_\mu$, using the Leibniz rule (3.3) and then using (3.4), we have

$$\begin{aligned}
 (\delta_\mu u_\alpha)(\delta_\gamma(\mathfrak{b}^{\lambda r}\mathfrak{b}^{\mu r}))I_\gamma u_{\beta\lambda} &\sim -u_\alpha(I_\mu \delta_\gamma(\mathfrak{b}^{\lambda r}\mathfrak{b}^{\mu r}))I_\gamma u_{\beta\lambda\mu} \\
 &\quad - u_\alpha(\delta_\mu \delta_\gamma(\mathfrak{b}^{\lambda r}\mathfrak{b}^{\mu r}))I_\gamma u_{\beta\lambda} \\
 &\leq -u_\alpha(I_\mu \delta_\gamma(\mathfrak{b}^{\lambda r}\mathfrak{b}^{\mu r}))I_\gamma u_{\beta\lambda\mu} \\
 &= -u_\alpha(\delta_\gamma(\mathfrak{b}^{\lambda r}\mathfrak{b}^{\mu r}))I_\gamma u_{\beta\lambda\mu} \\
 &\quad - u_\alpha(P_\mu \delta_\gamma(\mathfrak{b}^{\lambda r}\mathfrak{b}^{\mu r}))I_\gamma R_\mu u_{\beta\lambda} \\
 &\leq -u_\alpha(\delta_\gamma(\mathfrak{b}^{\lambda r}\mathfrak{b}^{\mu r}))I_\gamma u_{\beta\lambda\mu}. \tag{4.18}
 \end{aligned}$$

Taking the adjoint $I_\gamma^* = I_\gamma$ and using (3.6) we get

$$\begin{aligned} u_\alpha \left(\delta_\gamma \left(\mathfrak{b}^{\lambda r} \mathfrak{b}^{\mu r} \right) \right) I_\gamma u_{\beta \lambda \mu} &= \left(I_\gamma u_\alpha \right) \left(\delta_\gamma \left(\mathfrak{b}^{\lambda r} \mathfrak{b}^{\mu r} \right) \right) u_{\beta \lambda \mu} \\ &\quad + \left(I_\gamma u_\alpha \right) \left(P_\gamma \delta_\gamma \left(\mathfrak{b}^{\lambda r} \mathfrak{b}^{\mu r} \right) \right) R_\mu u_{\beta \lambda} \\ &\quad + \left(I_\gamma u_\alpha \right) \left(\delta_\gamma^2 \left(\mathfrak{b}^{\lambda r} \mathfrak{b}^{\mu r} \right) \right) R_\mu u_{\beta \lambda} \\ &\leq \left(I_\gamma u_\alpha \right) \left(\delta_\gamma \left(\mathfrak{b}^{\lambda r} \mathfrak{b}^{\mu r} \right) \right) u_{\beta \lambda \mu} \end{aligned}$$

Using this, from (4.18) we get

$$\left(\delta_\mu u_\alpha \right) \left(\delta_\gamma \left(\mathfrak{b}^{\lambda r} \mathfrak{b}^{\mu r} \right) \right) I_\gamma u_{\beta \lambda} \leq - \left(I_\gamma u_\alpha \right) \left(\delta_\gamma \left(\mathfrak{b}^{\lambda r} \mathfrak{b}^{\mu r} \right) \right) u_{\beta \lambda \mu}.$$

Thus from (4.17) we have

$$B_2 \leq - \sum_{\gamma} \sum_{\lambda, \mu \in \Lambda_0} \left(I_\gamma u_\alpha \right) \left(\delta_\gamma \left(\mathfrak{b}^{\lambda r} \mathfrak{b}^{\mu r} \right) \right) u_{\beta \lambda \mu}. \quad (4.19)$$

Set

$$\hat{\mathfrak{a}}^{\lambda \mu} = \mathfrak{a}^{\lambda \mu} - \frac{1}{2} \mathfrak{b}^{\lambda r} \mathfrak{b}^{\mu r} \quad \text{for } \lambda, \mu \in \Lambda_0.$$

Then from (4.16) and (4.19) we obtain

$$A_1 + B_1 \leq -2 \sum_{\lambda, \mu \in \Lambda_0} \hat{\mathfrak{a}}^{\lambda \mu} u_{\alpha \lambda} u_{\alpha \mu}, \quad (4.20)$$

$$A_2 + B_2 \leq 2 \sum_{\lambda, \mu \in \Lambda_0} \sum_{\gamma} \left(I_\gamma u_\alpha \right) \hat{\mathfrak{a}}_\gamma^{\lambda \mu} u_{\beta \lambda \mu}. \quad (4.21)$$

By the simple inequality $2yz \leq \varepsilon y^2 + \varepsilon^{-1} z^2$ and by using Lemma 3.7 with $V = (V^{\lambda \mu}) = (u_{\beta \lambda \mu})$, we have

$$\begin{aligned} 2 \sum_{\lambda, \mu \in \Lambda_0} I_\gamma u_\alpha \hat{\mathfrak{a}}_\gamma^{\lambda \mu} u_{\beta \lambda \mu} &\leq \varepsilon \left| \sum_{\lambda, \mu \in \Lambda_0} \hat{\mathfrak{a}}_\gamma^{\lambda \mu} u_{\beta \lambda \mu} \right|^2 + \varepsilon^{-1} |I_\gamma u_\alpha|^2 \\ &\leq \varepsilon N \sum_{\lambda, \mu, \eta \in \Lambda_0} \hat{\mathfrak{a}}^{\lambda \mu} \left(\delta_\lambda u_{\beta \eta} \right) \left(\delta_\mu u_{\beta \eta} \right) \\ &\quad + \varepsilon N \sum_{\lambda, \mu \in \Lambda_0} |R_\mu u_{\beta \lambda}|^2 + \varepsilon^{-1} |I_\gamma u_\alpha|^2 \\ &\leq \varepsilon N \sum_{\lambda, \mu, \eta \in \Lambda_0} \hat{\mathfrak{a}}^{\lambda \mu} u_{\beta \eta \lambda} u_{\beta \eta \mu} \end{aligned}$$

for each $\varepsilon > 0$, where N is a constant depending only on d, K, m and $|\Lambda|$. Choosing here ε sufficiently small, by virtue of (4.11), (4.13), (4.14), (4.20) and (4.21) we have for each $h \neq 0$

$$\sum_{\alpha \in \Lambda^m} 2(\delta_\alpha u) \delta_\alpha (L^h u + f) + \sum_{\alpha \in \Lambda^m} \sum_r |\delta_\alpha (M^{h,r} u + g^r)|^2 \leq 0,$$

which proves the lemma when $p^\lambda = q^\lambda = 0$ for all $\lambda \in \Lambda_0$. If this additional condition is not satisfied, then the only change in the proof is that there is an additional term,

$$A_0 = 2 \left(\delta_\alpha u, \delta_\alpha \left(p^\lambda \delta_{h,\lambda} u - q^\lambda \delta_{-h,\lambda} u \right) \right),$$

in the right-hand side of (4.11), where the summation convention with respect to repeated $\lambda \in \Lambda_0$ is in force. By Lemma 3.5 for each $h > 0$

$$A_0 \leq 0,$$

which completes the proof of the lemma. $\square$

Our key estimate is formulated in the following theorem, where $\mathcal{K}_m(T)$ is defined by (2.8).

Theorem 4.4 *Let Assumptions 2.3 and 2.4 hold with $m \geq 0$. Then problem (2.3)–(2.4) has a unique L_2 -valued solution $(\bar{u}_t^h)_{t \in [0, T]}$, which is a continuous H^m -valued process. If Assumption 2.5 also holds then for every $p > 0$ there is a constant $N = N(T, d, K, p, m, |\Lambda_0|)$ such that*

$$E \sup_{t \leq T} |\bar{u}_t^h|_m^p \leq N \left(E |\psi|_m^p + E \mathcal{K}_m^p(T) \right) \quad (4.22)$$

for all $h > 0$. If in addition to the above assumptions $p^\lambda = q^\lambda = 0$ for all $\lambda \in \Lambda_0$, then (4.22) holds for all $h \in \mathbb{R} \setminus \{0\}$.

Proof By virtue of Proposition 4.2 we need only prove the statements on the estimate (4.22). To this end we introduce the Hilbert norm $|v|_{\mathcal{H}^m}$ on H^m by

$$|v|_{\mathcal{H}^m}^2 = \sum_{|\alpha| \leq m} |\delta_\alpha v|_0^2,$$

and apply Itô's formula to $|u_t|_{\mathcal{H}^m}^2$ to obtain

$$\begin{aligned} d|\bar{u}_t^h|_{\mathcal{H}^m}^2 &= \left\{ 2(\bar{u}_t, L_t^h \bar{u}_t^h + f_t)_{\mathcal{H}^m} + \sum_r |M_t^{h,r} \bar{u}_t^h + g_t^r|_{\mathcal{H}^m}^2 \right\} dt + dm_t^h \\ &= Q_m^h(\bar{u}_t^h, f_t, g_t) dt + dm_t^h, \end{aligned}$$

where

$$dm_t^h = 2\left(\bar{u}_t^h, M_t^{h,r} \bar{u}_t + g_t^r\right)_{\mathcal{H}^m} dw_t^r,$$

and Q_m^h is defined by (4.8). Since $|\varphi|_m \leq |\varphi|_{\mathcal{H}^m}$ for all $\varphi \in H^m$, by virtue of Lemma 4.3 we get

$$d|\bar{u}_t^h|_{\mathcal{H}^m}^2 \leq N\left(|\bar{u}_t^h|_{\mathcal{H}^m}^2 + |f_t|_m^2 + |g_t^r|_m^2\right) dt + dm_t^h$$

for each $h > 0$, provided Assumptions 2.3, 2.4 and 2.5 hold with $m \geq 0$, and we get this for all $h \neq 0$ if in addition to these assumptions, $p^\lambda = q^\lambda = 0$ also holds for every $\lambda \in \Lambda_0$. Clearly,

$$|(\bar{u}^h, M^{h,r} \bar{u} + g^r)_{\mathcal{H}^m}| \leq |(\bar{u}^h, M^{h,r})_{\mathcal{H}^m}| + |(\bar{u}^h, g^r)_{\mathcal{H}^m}|,$$

and by Lemma 3.2

$$\sum_r |(\bar{u}^h, g^r)_{\mathcal{H}^m}|^2 \leq |\bar{u}^h|_{\mathcal{H}^m}^2 \sum_r |g^r|_{\mathcal{H}^m}^2 \leq N|\bar{u}^h|_{\mathcal{H}^m}^2 |g|_m^2$$

with a constant $N = N(m, d, \Lambda)$. By Corollary 3.4

$$\sum_r |(\bar{u}^h, M^{h,r})_{\mathcal{H}^m}|^2 \leq N|\bar{u}^h|_m^2 \leq N|\bar{u}^h|_{\mathcal{H}^m}^2.$$

with a constant $N = N(m, d, K, |\Lambda_0|)$. Thus for $y_t^h = |\bar{u}_t^h|_{\mathcal{H}^m}^2$ for each $h > 0$ we have that almost surely

$$y_t^h \leq N \int_0^t \left(y_s^h + F_s\right) ds + m_t^h \quad \text{for all } t \in [0, T], \quad (4.23)$$

where $F_s = |f_s|_m^2 + |g_s|_m^2$ and m^h is a continuous local martingale for every $h \neq 0$, such that by virtue of Corollary 3.4 we have for every $h \neq 0$

$$\begin{aligned} d\langle m^h \rangle_t &= 4 \sum_{r=1}^{\infty} |(\bar{u}_t^h, M_t^{h,r} \bar{u}_t^h + g_t^r)_{\mathcal{H}^m}|^2 dt \\ &\leq N \left(|y_t^h|^2 + y_t^h |g_t|_m^2 \right) dt \end{aligned} \quad (4.24)$$

with a constant $N = N(K, m, d, |\Lambda_0|)$. Moreover, (4.23) holds for each $h \neq 0$ if $p^\lambda = q^\lambda = 0$ for all $\lambda \in \Lambda_0$. Hence we finish the proof by applying Lemma 3.8 to (4.23) and using the obvious inequality $|u^h|_m^2 \leq y^h$. $\square$

For $p > 0$ let $\mathbb{H}_p^l$ denote the space of H^l -valued adapted processes $(v_t)_{t \in [0, T]}$ such that

$$|v|_{\mathbb{H}_p^l}^p = E \int_0^T |v_t|_l^p dt < \infty.$$

Clearly, $\mathbb{H}_p^l$ is a Banach space for $p \geq 1$, and it is a Hilbert space for $p = 2$. Our aim now is to show that if $E\mathcal{K}_1^p(T) < \infty$ for some $p \geq 2$, then Theorem 4.4 implies the weak convergence of $\bar{u}^h$ in $\mathbb{H}_p^1$ to some $\bar{u}$, which is the unique solution of

$$du_t = \left(\alpha_t^{\lambda\mu} \partial_\lambda \partial_\mu u_t + f_t \right) dt + \left(b_t^{\lambda r} \partial_\lambda u_t + g_t^r \right) dw_t^r, \quad u_0 = \psi, \quad (4.25)$$

on $[0, T]$, where the summation convention is used over repeated $\lambda, \mu \in \Lambda_1$ and $r \in \{0, 1, \dots\}$. The notion of solution is understood in the sense of Definition 2.1. Thus, saying that a process $\bar{u}$ is a solution to (4.25) on $[0, T]$ means that $\bar{u}$ has a weakly continuous H^1 -valued process from its equivalence class, say u , such that almost surely

$$\begin{aligned} (u_t, \varphi) = & (\psi, \varphi) + \int_0^t \{(\partial_\lambda u_s, \partial_{-\mu}(\alpha_s^{\lambda\mu} \varphi)) + (f_s, \varphi)\} ds \\ & + \int_0^t (b_s^{\lambda r} \partial_\lambda u_s + g_s^r, \varphi) dw_s^r \end{aligned} \quad (4.26)$$

for all $t \in [0, T]$ and $\varphi \in C_0^\infty(\mathbb{R}^d)$.

Theorem 4.5 *Let Assumptions 2.3, 2.4 and 2.5 hold with $m \geq 1$. Let $p \geq 2$ and assume that $E|\psi|_m^p < \infty$ and $E\mathcal{K}_m^p(T) < \infty$. Then $\bar{u}^h$ converges weakly in $\mathbb{H}_p^l$, for every $l \in [0, m]$, to some $u = (u_t)_{t \in [0, T]}$ as $n \rightarrow \infty$. The process u is the unique solution of (4.25) on $[0, T]$. Moreover, u is an H^m -valued weakly continuous process, it is strongly continuous as an H^{m-1} -valued process, and*

$$E \sup_{t \in [0, T]} |u|_l^p \leq N \left(E|\psi|_l^p + E\mathcal{K}_l^p(T) \right) \quad (4.27)$$

for $l \in \{0, 1, \dots, m\}$, with a constant $N = N(K, d, p, T, m, |\Lambda_0|)$.

Proof If $u^{(1)}$ and $u^{(2)}$ are solutions to of (4.25) on $[0, \tau]$ for a stopping time τ , then using Itô's formula for $y_t = |u_{t \wedge \tau}^{(1)} - u_{t \wedge \tau}^{(2)}|_0^2$, due to Assumptions 2.3 through 2.5, by simple calculations one obtains that almost surely

$$y_t \leq N \int_0^t y_s ds + m_t \quad \text{for all } t \in [0, T],$$

for a continuous local martingale m starting from 0. Hence the uniqueness of the solution to (4.25) on $[0, \tau]$ follows immediately for stopping times τ , in particular, for $\tau = T$. Thus to show that for $0 < h_n \rightarrow 0$ we have $\bar{u}^h \rightarrow u$ weakly in $\mathbb{H}_p^l$ for

$l = 1, \dots, m$, where u is the solution to (4.25), we need only prove that for every sequence $0 < h_n \rightarrow 0$ there is a subsequence h'_n such that $\bar{u}^{h_n} \rightarrow u$ converges weakly in $\mathbb{H}_p^l$, for $l = 1, \dots, m$, to the solution u of (4.25) as $h = h'_n \rightarrow 0$. To this end let us consider a sequence $0 < h_n \rightarrow 0$. Then, since by virtue of Theorem 4.4 we have

$$\sup_n |\bar{u}^{h_n}|_{\mathbb{H}_p^l} < \infty, \quad l = 0, 1, \dots, m, \quad (4.28)$$

there is a subsequence of h_n , for simplicity we denote it also by h_n , such that

$$\bar{u}^{h_n} \rightarrow u \text{ weakly in } \mathbb{H}^l \text{ for } l = 0, 1, \dots, m,$$

for some $u \in \mathbb{H}^m$. Fix a function $\varphi \in C_0^\infty(\mathbb{R}^d)$ and an adapted real valued process $(\eta_t)_{t \in [0, T]}$ which is bounded by some constant c . Define the functionals Φ , Φ_h , Ψ and Ψ_h over $\mathbb{H}_p^1$ by

$$\begin{aligned} \Phi(v) &= E \int_0^T \eta_t \int_0^t \sum_{\lambda, \mu \in \Lambda_1} \left(\partial_\lambda v_s, \partial_{-\mu} (\mathfrak{a}_s^{\lambda, \mu} \varphi) \right) ds dt, \\ \Phi_h(v) &= E \int_0^T \eta_t \int_0^t \sum_{\lambda, \mu \in \Lambda_1} \left(\mathfrak{a}_s^{\lambda, \mu} \delta_\lambda^h \delta_\mu^h v_s, \varphi \right) ds dt, \\ \Psi(v) &= E \int_0^T \eta_t \int_0^t \sum_{\lambda \in \Lambda_1} \left(\mathfrak{b}_s^{\lambda, r} \partial_\lambda v_s, \varphi \right) dw_s^r dt \\ \Psi_h(v) &= E \int_0^T \eta_t \int_0^t \sum_{\lambda \in \Lambda_1} \left(\mathfrak{b}_s^{\lambda, r} \delta_\lambda^h v_s, \varphi \right) dw_s^r dt \end{aligned}$$

for $v \in \mathbb{H}_p^1$ and $h > 0$, where ∂_λ is the identity for $\lambda = 0$. By the Bunyakovsky-Cauchy-Schwarz and the Davis-Burkholder-Gundy inequalities we have

$$\begin{aligned} \Phi(v) &\leq cTE \int_0^T \sum_{\lambda, \mu \in \Lambda_1} |\partial_\lambda v_s|_0 |\partial_\mu (\mathfrak{a}_s^{\lambda, \mu} \varphi)|_0 ds \leq N |v|_{\mathbb{H}_p^1} |\varphi|_1 \\ \Psi(v) &\leq cTE \sup_{t \leq T} \left| \int_0^t \sum_{\lambda \in \Lambda_1} (\mathfrak{b}_s^{\lambda, r} \partial_\lambda v, \varphi) dw_s^r \right| \\ &\leq 3cT \left\{ E \int_0^T \sum_{r=1}^\infty \left| \sum_{\lambda \in \Lambda_1} |\mathfrak{b}_t^{\lambda, r} \partial_\lambda v_s|_0 |\varphi|_0|^2 ds \right\}^{1/2} \\ &\leq N |v|_{\mathbb{H}_p^1} |\varphi|_1 \end{aligned}$$

with a constant $N = N(c, K, T, d, p, |\Lambda_0|)$. Consequently, Φ and Ψ are continuous linear functionals over $\mathbb{H}_p^1$, and therefore

$$\lim_{n \rightarrow \infty} \Phi(\bar{u}^{h_n}) = \Phi(\bar{u}), \quad \lim_{n \rightarrow \infty} \Psi(\bar{u}^{h_n}) = \Psi(\bar{u}). \quad (4.29)$$

Using the Bunyakovsky–Cauchy–Schwarz and the Davis–Burkholder–Gundy inequalities by Lemma 3.6 we get

$$\left| \Phi_h(\bar{u}^h) - \Phi(\bar{u}^h) \right| + \left| \Psi_h(\bar{u}^h) - \Psi(\bar{u}^h) \right| \leq Nh \left| u^h \right|_{\mathbb{H}_p^1} \left| \varphi \right|_2 \quad (4.30)$$

with a constant $N = N(c, K, T, d, p, |\Lambda|_0)$. From (4.29) and (4.30), taking into account (4.28) we obtain

$$\lim_{n \rightarrow \infty} \Phi_{h_n}(\bar{u}^{h_n}) = \Phi(\bar{u}), \quad \lim_{n \rightarrow \infty} \Psi_{h_n}(\bar{u}^{h_n}) = \Psi(\bar{u}). \quad (4.31)$$

Let us now multiply both sides of Eq. (4.4) by η_t and integrate over $\Omega \times [0, T]$ against the measure $P \times dt$ to get

$$\begin{aligned} E \int_0^T \eta_t \left(\bar{u}_t^h, \varphi \right) dt &= E \int_0^T \eta_t \left(\psi, \varphi \right) dt \\ &\quad + E \int_0^T \eta_t \int_0^t \left(L_s^h \bar{u}_s^h + f_s, \varphi \right) ds dt \\ &\quad + E \int_0^T \eta_t \int_0^t \left(M_s^{h,r} \bar{u}_s^h + g_s^r, \varphi \right) dw_s^r dt. \end{aligned}$$

Taking here the limit $h = h_n \rightarrow 0$, by virtue of (4.31) we get

$$\begin{aligned} E \int_0^T \eta_t \left(\bar{u}_t, \varphi \right) dt &= E \int_0^T \eta_t \left(\psi, \varphi \right) dt \\ &\quad + E \int_0^T \eta_t \int_0^t \left(\partial_\lambda \bar{u}_s, \partial_{-\mu} (a_s^{\lambda\mu} \varphi) \right) + \left(f_s, \varphi \right) ds dt \\ &\quad + E \int_0^T \eta_t \int_0^t \left(b_s^{\lambda r} \partial_\lambda \bar{u}_s + g_s^r, \varphi \right) dw_s^r dt \end{aligned}$$

for every bounded predictable process $(\eta_t)_{t \in [0, T]}$. Hence (4.26) holds for $P \times dt$ almost every $(\omega, t) \in \Omega \times [0, T]$ for each $\varphi \in C_0^\infty(\mathbb{R}^d)$, which by virtue of the theorem on Itô's formula from [23] or [24] implies that in the equivalence class of u there is an H^0 -valued continuous process, denoted also by u , such that almost surely (4.26) holds for all $t \in [0, T]$ and $\varphi \in C_0^\infty(\mathbb{R}^d)$. Note that

$$v_t^* := a_t^{\lambda\mu} \partial_\lambda \partial_\mu u_t + f_t, \quad t \in [0, T]$$

is an H^{m-2} -valued process and

$$m_t := \psi + \int_0^t \left(b^{\lambda r} \partial_\lambda u_s + g_s^r \right) dw_s^r \quad t \in [0, T],$$

is an H^{m-1} -valued continuous local martingale, such that

$$u_t = \int_0^t v_s^* ds + m_t$$

for $P \times dt$ almost every $(\omega, t) \in \Omega \times [0, T]$, and almost surely

$$\int_0^T \|u_t\|_m^2 dt < \infty, \quad \int_0^T \|u_t^*\|_{m-2}^2 dt < \infty.$$

Hence by considering the continuous and dense embedding $H^{m-2} \hookrightarrow H^{m-1}$ and the identification of H^{m-2} with its adjoint by the help of the inner product in H^{m-1} , using the theorem on Itô's formula from [23] or [24] again we see that the equivalence class of $u \in \mathbb{H}_P^m$ contains an H^{m-1} -valued continuous process. It remains to prove that this H^{m-1} -valued continuous process, which we keep denoting also by u , is almost surely an H^m -valued weakly continuous process, and that (4.27) holds. To this end we follow an argument from [18]. By the Banach-Saks theorem there is a sequence v^n of finite convex combinations of elements of $(u^{h_n})_{n=1}^\infty$ such that $v^n \rightarrow u$ strongly in $\mathbb{H}^m$ as $n \rightarrow \infty$. Hence for a subsequence $n_k \rightarrow \infty$ we have

$$v^n \rightarrow u \text{ strongly in } H^m, \text{ for } P \times dt \text{ almost all } (\omega, t) \in \Omega \times [0, T].$$

Thus there is a dense set $\{t_i\}_{i=1}^\infty$ in $[0, T]$ such that for all multi-indices $|\alpha| \leq m$ for each set $\{\phi_j\}_{j=1}^\infty \subset C_0^\infty(\mathbb{R}^d)$ we have almost surely

$$\begin{aligned} \left(u_{t_i}, (-1)^{|\gamma|} D^\alpha \phi_j \right) &= \left(D^\alpha u_{t_i}, \phi_j \right) \\ &= \lim_{n \rightarrow \infty} \left(D^\alpha v_{t_i}^n, \phi_j \right) \leq \lim_{n \rightarrow \infty} \|u_{t_i}^{h_n}\|_m \|\phi_j\|_0 \end{aligned} \quad (4.32)$$

for all integers $i, j \geq 1$. Since u is an H^{m-1} valued continuous process, substituting $D^\alpha \phi_j$ for ϕ in (4.26) and integrating by parts, we can see that almost surely

$$\left(D^\alpha u_t, \phi_j \right) \text{ is continuous in } t \in [0, T] \quad (4.33)$$

for all multi-indices $|\alpha| \leq m$ and for all $j \geq 1$. Taking $\{\phi_j\}_{j=1}^\infty$ dense in the unit ball of H^0 , hence we have that almost surely

$$\begin{aligned} \sup_{t \in [0, T]} \|D^\alpha u_t\|_0 &= \sup_{t \in [0, T]} \sup_{j \geq 1} \left(D^\alpha u_t, \phi_j \right) = \sup_{i \geq 1} \sup_{j \geq 1} \left(D^\alpha u_{t_i}, \phi_j \right) \\ &\leq \sup_{i \geq 1} \lim_{n \rightarrow \infty} \|u_{t_i}^{h_n}\|_m \leq \lim_{n \rightarrow \infty} \sup_{t \in [0, T]} \|u_t^{h_n}\|_m. \end{aligned} \quad (4.34)$$

Hence by Fatou's lemma

$$E \sup_{t \in [0, T]} \left| D^\alpha u_t \right|_0^p \leq \liminf_{n \rightarrow \infty} E \sup_{t \in [0, T]} \left| u_t^{h_n} \right|_m^p,$$

which together with (4.22) yield

$$E \sup_{t \in [0, T]} \left| u_t \right|_m^p \leq N \left(E \left| \psi \right|_m^p + E \mathcal{K}_m^p(T) \right) < \infty.$$

Using this, from (4.33) we get that almost surely u , as an H^m -valued process, is weakly continuous in $t \in [0, T]$, which completes the proof. $\square$

Proof of Theorem 2.1 Note first that due to Assumption 2.2, Eqs. (2.1) and (4.25) are equivalent. Set

$$\mathcal{R}_m^2(t) = \left| \psi \right|_m^2 + \mathcal{K}_m^2(t) = \left| \psi \right|_m^2 + \int_0^t \left| f_s \right|_m^2 + \left| g_s \right|_{m+1}^2 ds, \quad t \in [0, T], \quad (4.35)$$

and define

$$\tau_n = \inf \left\{ t \in [0, T] : \mathcal{R}_m(t) \geq n \right\},$$

$$\psi^{(n)} = \psi \mathbf{1}_{\tau_n > 0}, \quad f_t^{(n)} = f_t \mathbf{1}_{t \leq \tau_n}, \quad g_t^{(n)} = g_t \mathbf{1}_{t \leq \tau_n}.$$

Then $\{\tau_n\}_{n=1}^\infty$ is a nondecreasing sequence of stopping times such that $P(\tau_n \geq T) \rightarrow 1$ for $n \rightarrow \infty$, and the process $\mathcal{R}_m$ with $\psi^{(n)}$, $f^{(n)}$ and $g^{(n)}$ in place of ψ , f and g , respectively, is bounded by n . Then by virtue of Theorem 4.5 for each n the Cauchy problem (4.25) with $\mathbf{1}_A \psi^{(n)}$, $\mathbf{1}_A \mathbf{1}_{(0, \tau]} f^{(n)}$ and $\mathbf{1}_A \mathbf{1}_{(0, \tau]} g^{(n)}$ in place of ψ , f and g , respectively, admits a solution $u^{(n)}$ on $[0, T]$, and because of the uniqueness of the solution on any stochastic interval, for $m \geq n$ we have $u_t^{(n)} = u_t^{(m)}$ almost surely for $t \in [0, \tau_n]$. Hence almost surely $u_t = \lim_{n \rightarrow \infty} u_t^{(n)}$ exists for $t \in [0, T]$ and it is the solution to (4.25). Moreover, for each stopping time $\tau \leq T$ and set $A \in \mathcal{F}_0$ problem (4.25) admits a unique solution v on $[0, T]$ and $v_t = \mathbf{1}_A u_t^{(n)}$. Thus by (4.27) for every $p \geq 2$ we have

$$E(\mathbf{1}_A \sup_{t \leq \tau \wedge \tau_n} |u_t|_l^p) \leq N E(\mathbf{1}_A \mathcal{R}_l^p(\tau_n \wedge \tau))$$

for every stopping time $\tau \leq T$ and $A \in \mathcal{F}_0$, where N is a constant depending only on K, T, p, d, m , and $|\Lambda_0|$. Hence by Lemma 3.2 from [25] for each $p > 0$ we get

$$E \sup_{t \leq \tau \wedge \tau_n} |u_t|_l^p \leq N E \mathcal{R}_l^p(\tau_n \wedge \tau)$$

for every stopping time $\tau \leq T$ with a constant $N = N(K, T, p, d, m, |\Lambda_0|)$. Letting here $n \rightarrow \infty$ we get (2.9) for any $p > 0$ and stopping time $\tau \leq T$. $\square$

5 The coefficients of the expansion

To determine the coefficients $u^{(1)}, \dots, u^{(k)}$ in the expansion (2.10) we assume that Assumptions 2.2 through 2.5 hold and consider the system of stochastic PDEs

$$\begin{aligned} dv_t^{(n)} = & \left(\mathcal{L}_t v_t^{(n)} + \sum_{l=1}^n \binom{n}{l} \mathcal{L}_t^{(l)} v_t^{(n-l)} \right) dt \\ & + \left(\mathcal{M}_t^r v_t^{(n)} + \sum_{l=1}^n \binom{n}{l} \mathcal{M}_t^{(l)r} v_t^{(n-l)} \right) dw_t^r, \quad n = 1, \dots, k, \end{aligned} \quad (5.1)$$

with initial condition $v_0^{(n)} = 0$ for $v^{(n)}$, $n = 1, \dots, k$, where $v^{(0)} = u$ is the solution of (2.1)–(2.2),

$$\mathcal{L}_t^{(0)} = \sum_{\lambda, \mu \in \Lambda} \alpha_t^{\lambda\mu} \partial_\lambda \partial_\mu + \sum_{\lambda \in \Lambda_0} \left(p_t^\lambda - q_t^\lambda \right) \partial_\lambda, \quad \mathcal{M}_t^{(0)r} = \sum_{\lambda \in \Lambda} b_t^{\lambda r} \partial_\lambda \quad (5.2)$$

and

$$\begin{aligned} \mathcal{L}_t^{(n)} = & \sum_{\lambda, \mu \in \Lambda_0} \alpha_t^{\lambda\mu} \sum_{j=0}^n A_{nj} \partial_\lambda^{j+1} \partial_\mu^{n-j+1} + (n+1)^{-1} \sum_{\lambda \in \Lambda_0} \left(\alpha_t^{\lambda 0} + \alpha_t^{0\lambda} \right) B_n \partial_\lambda^{n+1} \\ & + (n+1)^{-1} \sum_{\lambda \in \Lambda_0} \left(p_t^\lambda + (-1)^{n+1} q_t^\lambda \right) \partial_\lambda^{n+1}, \end{aligned} \quad (5.3)$$

$$\mathcal{M}_t^{(n)r} = (n+1)^{-1} \sum_{\lambda \in \Lambda_0} b_t^{\lambda r} B_n \partial_\lambda^{n+1} \quad (5.4)$$

for $n \geq 1$. Here and later on the constants A_{nj} and B_n are defined for integers $n \geq 0$ and $j \in [0, n]$ as follows:

$$B_n = \begin{cases} 0 & \text{if } n \text{ is odd} \\ 1 & \text{if } n \text{ is even} \end{cases}, \quad A_{nj} = \begin{cases} 0 & \text{if } n \text{ or } j \text{ is odd} \\ \frac{n!}{(j+1)!(n-j+1)!} & \text{if } n \text{ and } j \text{ are even} \end{cases}. \quad (5.5)$$

Note that, as we can see by Lemma 5.2 below, for sufficiently smooth φ

$$\begin{aligned} L_t^h \varphi(x) & \rightarrow \mathcal{L}_t^{(0)} \varphi(x), \quad M_t^h \varphi(x) \rightarrow \mathcal{M}_t^{(0)} \varphi(x), \\ \frac{\partial^n}{\partial h^n} L_t^h \varphi(x) & \rightarrow \mathcal{L}_t^{(n)} \varphi(x), \quad \frac{\partial^n}{\partial h^n} M_t^h \varphi(x) \rightarrow \mathcal{M}_t^{(n)} \varphi(x), \end{aligned}$$

as $h \rightarrow 0$, and the above system of SPDEs is obtained by differentiating formally n -times Eq. (2.1) in the parameter h , and taking $h \rightarrow 0$.

The notion of solution is understood in the sense of Definition 2.1. To formulate a suitable existence and uniqueness theorem we use for integers $l \geq 1$ the notation $\mathbb{C}^l(T)$ for the space of H^l -valued weakly continuous adapted processes $v = (v_t)_{t \in [0, T]}$ which are strongly continuous as H^{l-1} -valued processes.

Theorem 5.1 *Let Assumptions 2.2 through 2.5 hold with $m \geq 2k + 1$ for an integer $k \geq 1$. Then there is a unique solution $v^{(1)}, \dots, v^{(k)}$ to (5.1) with initial condition $v_0^{(1)} = \dots = v_0^{(k)} = 0$. Moreover, $v^{(n)} \in \mathbb{C}^{m-2n}(T)$ and*

$$E \sup_{t \leq T} |v_t^{(n)}|^p \Big|_{m-2n} \leq N \left(E |\psi|_m^p + EK_m^p(T) \right) \quad \text{for } n = 1, \dots, k, \quad (5.6)$$

for $p > 0$, with a constant $N = N(m, k, T, K, p, |\Lambda_0|)$. If k is an odd number, and $p^\lambda = q^\lambda = 0$ for $\lambda \in \Lambda_0$, then Assumptions 2.2 through 2.5 need only be satisfied with $m \geq 2k - 1$ for the above conclusions to hold. In this case we have $v^{(n)} = 0$ for odd $n \leq k$.

Proof Though this theorem can be proved in the same way as Theorem 3.7 from [3], for the convenience of the reader we present its proof here. For each $n = 1, \dots, k$ the equation for $v_t^{(n)}$ involves only the unknown functions $v^{(1)}, \dots, v^{(n)}$. Therefore we prove the theorem recursively on $n \leq k$. Set

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

and consider first the case $n = 1$. By Theorem 2.1 we have $v^{(0)} \in \mathbb{C}^m(T)$, and estimate (2.9) holds. Since $\mathcal{M}^{(1)} = 0$, we have $R^{(1)} = 0$, and due to Assumption 2.3 we have

$$\left| \mathcal{L}_t^{(1)} v_t^{(0)} \right|_{m-2} \leq N \left| v_t^{(0)} \right|_m.$$

Hence

$$\int_0^T \left| S_t^{(1)} \right|_{m-2}^2 + \left| R_t^{(1)} \right|_{m-1}^2 dt \leq N \sup_{t \leq T} \left| v_t^{(0)} \right|_m^2.$$

Thus by Theorem 2.1 there exists a unique $v^{(1)} \in \mathbb{C}^{m-2}(T)$ satisfying (5.1) with zero initial condition, and (5.6) holds with $n = 1$. For $n \in \{2, \dots, k\}$ we assume that there exist $v^{(1)}, \dots, v^{(n-1)}$ with the asserted properties. Then $\mathcal{M}^{(1)} v^{(n-1)} = 0$ and

$$\left| \mathcal{L}_t^{(1)} v_t^{(n-1)} \right|_{m-2n} \leq N \left| v_t^{(n-1)} \right|_{m-2n+2} = N \left| v_t^{(n-1)} \right|_{m-2(n-1)},$$

and for $i \geq 2$

$$\begin{aligned} \left| \mathcal{L}_t^{(i)} v_t^{(n-i)} \right|_{m-2n} &\leq N \left| v_t^{(n-i)} \right|_{m-2n+(i+2)} \leq N \left| v_t^{(n-i)} \right|_{m-2(n-i)}, \\ &\sum_{r=1}^{\infty} \left| \mathcal{M}^{(i)r} v^{(n-i)} \right|_{m-2n+1}^2 = N \left| v^{(n-i)} \right|_{m-2n+1+(i+1)}^2 \\ &\leq N \left| v_t^{(n-i)} \right|_{m-2(n-i)}. \end{aligned}$$

It follows from the induction hypothesis that almost surely

$$\int_0^T \left| S_t^{(n)} \right|_{\mathfrak{m}-2n}^2 + \left| R_t^{(n)} \right|_{\mathfrak{m}-2n+1}^2 dt \leq N \sum_{i=1}^n \sup_{t \leq T} \left| v_t^{(n-i)} \right|_{\mathfrak{m}-2(n-i)}^2 < \infty, \quad (5.7)$$

which by Theorem 2.1 yields the existence of a unique $v^{(n)} \in \mathbb{C}^{\mathfrak{m}-2n}(T)$ satisfying (5.1) with zero initial condition, and estimate (2.9) applied to $v^{(n)}$ via (5.7) gives (5.6). The proof of the existence of $v^{(1)}, \dots, v^{(k)}$ with the stated properties is complete. Since for any given $v^{(i)} \in \mathbb{C}^{\mathfrak{m}-2i}(T)$, $i = 1, \dots, n-1$, the solution $v^{(n)}$ to Eq. (5.1) with zero initial condition is uniquely determined, the uniqueness of the solution $(v^{(1)}, \dots, v^{(k)})$ also holds.

Notice that $\mathcal{M}^{(n)} = 0$ for odd $n \leq k$ by (5.5). 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 (5.5). Hence $S^{(1)} = 0$ and $R^{(1)} = 0$, which implies $v^{(1)} = 0$. Assume that $k \geq 2$ and that for an odd $n \leq k$ we have $v^{(l)} = 0$ for all odd $l < n \leq k$. Then $\mathcal{L}^{(i)} v^{(n-i)} = 0$ and $\mathcal{M}^{(i)} v^{(n-i)} = 0$ for all $i = 1, \dots, n$, since either i or $n-i$ is odd. Thus $S^{(n)} = 0$ and $R^{(n)} = 0$, and hence $v^{(n)} = 0$ for all odd $n \leq k$. In particular, when k is odd we need only solve (5.1) for $n = 1, \dots, k-1$, since for $n = k$ the right-hand side of Eq. (5.1) is zero. $\square$

We evoke now Lemma 3.6 from [3].

Lemma 5.2 *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*

$$\delta_{0,\lambda}\phi = \partial_\lambda\phi = \lambda^i D_i\phi, \quad \partial_{\lambda\mu} = \partial_\lambda\partial_\mu.$$

Then we have

$$\frac{\partial^n}{(\partial h)^n} \delta_{h,\lambda}\phi(x) = \int_0^1 \theta^n \partial_\lambda^{n+1} \phi(x + h\theta\lambda) d\theta, \quad (5.8)$$

$$\frac{\partial^n}{(\partial h)^n} \delta_\lambda^h \phi(x) = \frac{1}{2} \int_{-1}^1 \theta^n \partial_\lambda^{n+1} \phi(x + h\theta\lambda) d\theta, \quad (5.9)$$

$$\frac{\partial^n}{(\partial h)^n} \delta_\lambda^h \delta_\mu^h \psi(x) = \frac{1}{4} \int_{-1}^1 \int_{-1}^1 \left(\theta_1 \partial_\lambda + \theta_2 \partial_\mu \right)^n \partial_{\lambda\mu} \psi(x + h(\theta_1\lambda + \theta_2\mu)) d\theta_1 d\theta_2, \quad (5.10)$$

for each $h \neq 0$, for almost all $x \in \mathbb{R}^d$. Furthermore,

$$\frac{\partial^n}{(\partial h)^n} \delta_{h,\lambda}\phi \Big|_{h=0} = \frac{1}{n+1} \partial_\lambda^{n+1} \phi, \quad \frac{\partial^n}{(\partial h)^n} \delta_\lambda^h \phi \Big|_{h=0} = \frac{B_n}{n+1} \partial_\lambda^{n+1} \phi, \quad (5.11)$$

$$\frac{\partial^n}{(\partial h)^n} \delta_\lambda^h \delta_\mu^h \psi \Big|_{h=0} = \sum_{r=0}^n A_{n,r} \partial_\lambda^{r+1} \partial_\mu^{n-r+1} \psi, \quad (5.12)$$

and for integers $l \geq 0$, functions $\phi \in H^{n+2+l}$ and $\psi \in H^{n+3+l}$ we have

$$\left| \delta_{h,\lambda} \phi - \sum_{i=0}^n \frac{h^i}{(i+1)!} \partial_\lambda^{i+1} \phi \right|_l \leq \frac{|h|^{n+1}}{(n+2)!} |\partial_\lambda^{n+2} \phi|_l \quad (5.13)$$

$$\left| \delta_\lambda^h \phi - \sum_{i=0}^n \frac{h^i}{(i+1)!} B_i \partial_\lambda^{i+1} \phi \right|_l \leq \frac{|h|^{n+1}}{(n+2)!} |\partial_\lambda^{n+2} \phi|_l \quad (5.14)$$

$$\left| \delta_\lambda^h \delta_\mu^h \psi - \sum_{i=0}^n h^i \sum_{j=0}^i A_{i,j} \partial_\lambda^{j+1} \partial_\mu^{i-j+1} \psi \right|_l \leq N |h|^{n+1} |\psi|_{l+n+3}, \quad (5.15)$$

for every $h \neq 0$, with a constant $N = N(|\lambda|, |\mu|, d, n)$, where $A_{i,j}$ and B_i are defined by (5.5).

Proof It suffices to prove the lemma for $\phi, \psi \in C_0^\infty(\mathbb{R}^d)$. Formulas (5.8) and (5.9) for $n = 0$ can be obtained by applying the Newton-Leibniz formula to the function $\phi(x + \theta h \lambda)$, for $\theta \in [0, 1]$ and $\theta \in [-1, 1]$, respectively. We get (5.10) for $n = 0$ by applying (5.9) with $n = 0$ twice. After that we obtain (5.8)–(5.10) for $n \geq 1$ by differentiating these equations written with $n = 1$. Formulas in (5.11) and (5.12) follow from those in (5.8)–(5.9) and (5.10), respectively. We obtain estimates (5.13)–(5.15) by applying Taylor's formula to the functions

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

and

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

with remainder terms given in integral form, whose H^l -norm (in x) we estimate by Minkowski's inequality. For more details we refer to [3]. $\square$

Introduce the operators

$$\mathcal{O}_t^{h(n)} = L_t^h - \sum_{i=0}^n \frac{h^i}{i!} \mathcal{L}_t^{(i)}, \quad \mathcal{R}_t^{h(n)r} = M_t^{h,r} - \sum_{i=0}^n \frac{h^i}{i!} \mathcal{M}_t^{(i)r}, \quad r \geq 1$$

for integers $n \geq 0$, where $\mathcal{L}^{(i)}$ and $\mathcal{M}^{(i),r}$ are defined by (5.2) through (5.4).

Corollary 5.3 *Let l be a non-negative integer. Let $\alpha^{\lambda,\mu}$, $\mathfrak{b}^\lambda$, $\mathfrak{p}^\kappa$, $\mathfrak{q}^\kappa$ and their derivatives in x up to order l be functions and be bounded by a constant C for all $\lambda, \mu \in \Lambda_1$ and $\kappa \in \Lambda_0$. Then for every $n \geq 0$*

$$\left| \mathcal{O}_t^{h(n)} \right|_l \leq N |h|^{n+1} \left| \phi \right|_{l+n+3} \quad (5.16)$$

$$\left| \mathcal{R}_t^{h(n)} \phi \right|_l^2 = \sum_{\rho=1}^{\infty} \left| \mathcal{R}_t^{h(n)\rho} \phi \right|_l^2 \leq N |h|^{2n+2} \left| \phi \right|_{l+n+2}^2, \quad (5.17)$$

where N stands for constants depending only on n, d, l, C and $|\Lambda_0|$.

Proof This corollary follows immediately from (5.13) through (5.15). $\square$

6 Proof of Theorem 2.2

Assume that the conditions of Theorem 5.1 hold and set

$$\bar{r}_t^h = v_t^h - \sum_{j=1}^k \frac{h^j}{j!} v_t^{(j)}, \quad (6.1)$$

where v^h is the unique L_2 -valued solution of (4.1)–(4.2), $v^{(0)}$ is the solution of (2.1)–(2.2), and $(v^{(n)})_{n=1}^k$ is the solution of the system of stochastic partial differential equations (5.1).

Lemma 6.1 *Let Assumptions 2.2 through 2.5 hold with*

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

for some integers $k \geq 0$ and $l \geq 0$. Then $r_0^h = 0$, $r^h \in \mathbb{C}^l(T)$, and

$$d\bar{r}_t^h = \left(L_t^h \bar{r}_t^h + F_t^h \right) dt + \left(M_t^{h,r} \bar{r}_t^h + G_t^{h,r} \right) dw_t^r, \quad (6.2)$$

where

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

Moreover, we have that almost surely

$$\int_0^T \left| F_t^h \right|_l^2 dt < \infty \quad \int_0^T \left| G_t^h \right|_{l+1}^2 dt < \infty. \quad (6.3)$$

If $k \geq 1$ is odd and $\mathfrak{p}^\lambda = \mathfrak{q}^\lambda = 0$ for $\lambda \in \Lambda_0$ then the above assertions remain true if Assumptions 2.2 through 2.5 hold only with $m = l + 2k$ in place of $m = l + 2k + 2$.

Proof We have $v^{(h)} \in \mathbb{C}^l(T)$ and $v^{(j)} \in \mathbb{C}^l(T)$, for $j \leq k$ by Theorems 2.1 and 5.1. Hence $r^h \in \mathbb{C}^l(T)$. Using the equations for v^h and $v^{(n)}$ for $n = 0, \dots, k$, we can easily see that (6.2) holds with $\hat{F}^h$ and $\hat{G}^h$ in place of F^h and G^h , respectively, where

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

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

with

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

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

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

$$\begin{aligned} I^h &= \sum_{i=1}^k \sum_{j=i}^k \frac{1}{i!(j-i)!} \mathcal{L}^{(i)} v^{(j-i)} h^j \\ &= \sum_{i=1}^k \sum_{l=0}^{k-i} \frac{1}{i!l!} \mathcal{L}^{(i)} v^{(l)} h^{l+i} = \sum_{l=0}^{k-1} \frac{h^l}{l!} \sum_{i=1}^{k-l} \frac{h^i}{i!} \mathcal{L}^{(i)} v^{(l)} \\ &= \sum_{j=0}^k \frac{h^j}{j!} \sum_{i=1}^{k-j} \frac{h^i}{i!} \mathcal{L}^{(i)} v^{(j)}, \end{aligned}$$

and similarly,

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

Hence we get $\hat{F} = F$ and $\hat{G} = G$ by simple calculations, and since by Lemma 5.2 for $j = 0, 1, \dots, k$, $(\omega, t) = \Omega \times [0, T]$ we have

$$\begin{aligned} \left| \mathcal{O}_t^{h(k-j)} v_t^{(j)} \right|_l &\leq N \left| v_t^{(j)} \right|_{l+k-j+2} \leq N \left| v_t^{(j)} \right|_{m-2j}, \\ \left| \mathcal{R}_t^{h(k-j)} v_t^{(j)} \right|_l &\leq N \left| v_t^{(j)} \right|_{l+k-j+1} \leq N \left| v_t^{(j)} \right|_{m-2j-1}, \end{aligned}$$

(6.3) follows by virtue of Theorems 2.1 and 5.1. The last statement follows from the fact that $v^{(k)} = 0$ for odd k when $\mathbf{p}^\lambda = \mathbf{q}^\lambda = 0$ for $\lambda \in \Lambda_0$. $\square$

Now we present a theorem which is more general than Theorem 2.2.

Theorem 6.2 *Let Assumptions 2.2 through 2.5 hold with*

$$m = l + 2k + 3 \tag{6.4}$$

for some integer $k \geq 0$. Then for $h > 0$ and any $p > 0$ we have

$$E \sup_{t \leq T} |\bar{r}_t^h|_l^p \leq N|h|^{p(k+1)} (E|\psi|_m^p + EK_m^p(T)) \quad (6.5)$$

with a constant $N = N(T, K, l, d, k, p, |\Lambda_0|)$. Moreover, if $p^\lambda = q^\lambda = 0$ for $\lambda \in \Lambda_0$, then $v^{(j)} = 0$ in (6.1) for odd $j \leq k$, and if k is odd then it is sufficient to assume $m = l + 2k + 2$ instead of (6.4) to have estimate (6.5).

Proof Let $p > 0$ such that $E|\psi|_m^p + EK_m^p(T) < \infty$. Lemma 6.1 and Theorem 4.4 yield

$$E \sup_{t \leq T} |\bar{r}_t^h|_l^p \leq N \left(|F^h|_{\mathbb{H}_p^l}^p + |G^h|_{\mathbb{H}_p^{l+1}}^p \right) \quad (6.6)$$

with a constant $N = N(T, k, p, l, d, K, |\Lambda_0|)$. Let (6.4) hold. Then for $j = 0, \dots, k$

$$l + k - j + 3 \leq m - 2j,$$

and by Corollary 5.3 we have

$$\begin{aligned} \left| \mathcal{O}_t^{h(k-j)} v_t^{(j)} \right|_l + \left| \mathcal{R}_t^{h(k-j)} v_t^{(j)} \right|_{l+1} &\leq N|h|^{k-j+1} \left| v_t^{(j)} \right|_{l+k-j+3} \\ &\leq N|h|^{k-j+1} |v^{(j)}|_{m-2j}. \end{aligned} \quad (6.7)$$

Hence, using Theorem 5.1 we see that

$$\left| F^h \right|_{\mathbb{H}_p^l}^p + \left| G^h \right|_{\mathbb{H}_p^{l+1}}^p \leq N|h|^{p(k+1)} (E|\psi|_m^p + EK_m^p(T)),$$

which by (6.6) implies estimate (6.5). If $p^\lambda = q^\lambda = 0$ for $\lambda \in \Lambda_0$ then by Theorem 5.1 we know that $v^{(j)} = 0$ for odd $j \leq k$. In particular, $v^{(k)} = 0$ for odd k and (6.7) obviously holds for $j = k$ and to have it also for $j \leq k - 1$ we need only $m = l + 2k + 2$. $\square$

Set $r_t^h = J\bar{r}_t^h$, where $J : H^l \rightarrow C_b$, is the Sobolev embedding operator from H^l to C_b for $l > d/2$. Set $\Lambda^0 = \{0\}$ and recall that $\delta_{h,0}$ is the identity operator and $\delta_{h,\lambda} = \delta_{h,\lambda_1}, \dots, \delta_{h,\lambda_n}$ for $(\lambda_1, \dots, \lambda_n) \in \Lambda^n$, $n \geq 1$. Then we have the following corollary of Theorem 6.2

Corollary 6.3 *Let the assumptions of Theorem 6.2 hold and assume that $l > n + d/2$ for an integer $n \geq 0$ in (6.4). Then for $\lambda \in \Lambda^n$ we have*

$$E \sup_{t \in [0, T]} \sup_{x \in \mathbb{R}^d} \left| \delta_{h,\lambda} r_t^h(x) \right|^p \leq Nh^{p(k+1)} E(|\psi|_m^p + \mathcal{K}_m^p(T)), \quad (6.8)$$

and when Assumption 2.1 also holds,

$$E \sup_{t \in [0, T]} \left\{ \sum_{x \in \mathbb{G}_h} \left| \delta_{h,\lambda} r_t^h(x) \right|^2 h^d \right\}^{p/2} \leq Nh^{p(k+1)} E(|\psi|_m^p + \mathcal{K}_m^p(T)) \quad (6.9)$$

for $h > 0$ with a constant $N = N(d, n, k, K, T, p, \Lambda_0)$.

Proof Set $j = l - n$. Then $j > d/2$ and by Sobolev's theorem on embedding H^j into C_b and by Lemma 3.2, from Theorem 6.2 we get

$$\begin{aligned} E \sup_{t \in [0, T]} \sup_{x \in \mathbb{R}^d} \left| \delta_{h, \lambda} r_t^h(x) \right|^p &\leq C_1 E \sup_{t \in [0, T]} |\delta_{h, \lambda} r_t^h|_j^p \\ &\leq C_2 E \sup_{t \in [0, T]} |r_t^h|_l^p \leq N h^{p(k+1)} E(|\psi|_m^p + \mathcal{K}_m^p(T)). \end{aligned}$$

Similarly, by Lemma 4.1 and Lemma 3.2

$$\begin{aligned} E \sup_{t \in [0, T]} \left\{ \sum_{x \in \mathbb{G}_h} \left| \delta_{h, \lambda} r_t^h(x) \right|^2 h^d \right\}^{p/2} &\leq C_1 E \sup_{t \in [0, T]} |\delta_{h, \lambda} r_t^h|_l^p \\ &\leq C_2 E \sup_{t \in [0, T]} |R_t^h|_l^p \leq N h^{p(k+1)} E(|\psi|_m^p + \mathcal{K}_m^p(T)). \end{aligned}$$

□

Now we can easily see that Theorem 2.2 follows from the above corollary.

Let the conditions of Theorem 6.2 hold with an integer $l > n + d/2$ for some integer $n \geq 0$ and define

$$r^h = J\bar{r}^h, \quad u^h = Jv^h, \quad u^{(j)} = Jv^{(j)}, \quad j = 0, \dots, k,$$

v^h is a continuous H^l -valued process, and by Theorem 5.1 $v^{(j)}$, $j = 1, 2, \dots, k$, are H^{m-2k} -valued continuous processes. By Proposition 4.2 we know that u^h restricted to $\mathbb{G}_h$ is the unique $l_{2,h}$ -valued solution of the finite difference scheme and from the previous corollary we have that for each $h > 0$ almost surely

$$\delta_{h, \lambda} u_t^h(x) = \sum_{j=0}^k \frac{h^j}{j!} \delta_{h, \lambda} u_t^{(j)}(x) + h^{k+1} \delta_{h, \lambda} r_t^h(x), \quad (6.10)$$

for all $t \in [0, T]$ and $x \in \mathbb{G}_h$, and (6.8)–(6.9) hold for any $\lambda \in \Lambda^n$, for integers $n \geq 0$ such that $l > n + d/2$. Moreover, by Theorem 6.2 we have that $u^{(j)} = 0$ for odd $j \leq k$ provided $\mathbf{p}^\lambda = \mathbf{q}^\lambda = 0$ for $\lambda \in \Lambda_0$, and in this case for odd k it is sufficient to assume that Assumptions 2.2 through 2.5 hold with $m \geq l + k + 2$ and $l > n + d/2$, to have (6.10) and estimates (6.8) and (6.9) for all $h \neq 0$. Hence Theorem 2.2 follows immediately.

Acknowledgments The author is grateful to Chris Farmer and Terry Lyons in Oxford for useful discussions on the topic of the paper, and to Máté Gerencsér in Edinburgh for pointing out a few misprints.

References

1. Krylov, N.V.: The rate of convergence of finite-difference approximations for Bellman equations with Lipschitz coefficients. *Appl. Math. Optim.* **52**(3), 365–399 (2005)

2. Gyöngy, I., Krylov, N.V.: Accelerated finite difference schemes for stochastic parabolic partial differential equations in the whole space. *SIAM J. Math. Anal.* **42**(5), 2275–2296 (2010)
3. Gyöngy, I.: On finite difference schemes for degenerate stochastic parabolic partial differential equations. *J. Math. Sci.* **179**, 100–126 (2011)
4. Olejnik, O.A., Radkevich, E.V.: *Second Order Equations with Nonnegative Characteristic Form*. AMS, Providence (1973)
5. Gyöngy, I., Krylov, N.V.: Accelerated finite difference schemes for second order degenerate elliptic and parabolic problems in the whole space. *Math. Comput.* **80**, 1431–1458 (2011)
6. Dong, H., Krylov, N.V.: On the rate of convergence of finite-difference approximations for degenerate linear parabolic equations with C^1 and C^2 coefficients. *Electron. J. Diff. Equ.* **2005**(102), 1–25 (2005). <http://ejde.math.txstate.edu>
7. Richardson, L.F.: The approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a Masonry dam. *Phil. Trans. R. Soc. London, Ser. A* **210**, 307–357 (1911)
8. Richardson, L.F., Gaunt, J.A.: The deferred approach to the limit. *Phil. Trans. Roy. Soc. London Ser. A* **226**, 299–361 (1927)
9. Romberg, W.: Vereinfachte numerische Integration. *Det Kongelige Norske Videnskabers Selskab Forhandling* (Trondheim) **28**, 30–36 (1955)
10. Marschuk, G.I., Shaidurov, V.V.: *Difference Methods and Their Extrapolations*. Springer, New York (1983)
11. Sidi, A.: *Practical Extrapolation Methods: Theory and applications*. Cambridge Monographs on Applied and Computational Mathematics, vol. 10. Cambridge University Press, Cambridge (2003)
12. Brezinski, C.: Convergence acceleration during the 20th century. *J. Comput. Appl. Math.* **122**, 1–21 (2000)
13. Joyce, D.C.: Survey of extrapolation processes in numerical analysis. *SIAM Rev.* **13**(4), 435–490 (1971)
14. Talay, D., Tubaro, L.: Extension of the global error for numerical schemes solving stochastic differential equations. *Stoch. Anal. Appl.* **8**, 483–509 (1990)
15. Malliavin, P., Thalmaier, A.: Numerical error for SDE: asymptotic expansion and hyperdistributions. *C. R. Math. Acad. Sci. Paris* **336**(10), 851–856 (2003)
16. Hall, E.J.: Accelerated spatial approximations for time discretized stochastic partial differential equations. *SIAM J. Math. Anal.* **44**, 3162–3185 (2012)
17. Hall, E.J.: Higher order spatial approximations for degenerate parabolic stochastic partial differential equations. *SIAM J. Math. Anal.* **45**(4), 2071–2098 (2013)
18. Krylov, N.V., Rozovskii, B.L.: Characteristics of degenerating second-order parabolic Ito equations. *J. Sov. Maths.* **32**, 336–348. (1986). (Translated from *Trudy Seminara imeni I.G. Petrovskogo*, No. 8. pp. 153–168, 1982.).
19. Gyöngy, I., Krylov, N.V.: On the splitting-up method and stochastic partial differential equations. *Ann. Probab.* **31**, 564–591 (2003)
20. Yoo, H.: *An analytic approach to stochastic partial differential equations and its applications*, Ph.D. thesis, University of Minnesota, Minneapolis 1998.
21. Gyöngy, I.: *An introduction to the theory of stochastic partial differential equations*, in preparation.
22. Gyöngy, I., Shmatkov, A.: Rate of convergence of Wong-Zakai approximations for stochastic partial differential equations. *Appl. Math. Optim.* **54**(3), 315–341 (2006)
23. Krylov, N.V., Rozovskii, B.L.: Stochastic evolution equations. *J. Sov. Math.* **16**, 1233–1277 (1981)
24. Rozovskii, B.L.: *Stochastic Evolution Systems. Linear theory and applications to nonlinear filtering*. Kluwer, Dordrecht (1990)
25. Gyöngy, I., Krylov, N.V.: On the rate of convergence of splitting-up approximations for SPDEs. *Progress in Probability*, vol. 56, pp. 301–321, 2003. Birkhäuser, Basel.