ALMOST SURE CONVERGENCE OF A SEMIDISCRETE MILSTEIN SCHEME FOR SPDE'S OF ZAKAI TYPE

ANNIKA LANG, PAO-LIU CHOW, AND JÜRGEN POTTHOFF

ABSTRACT. A semidiscrete Milstein scheme for stochastic partial differential equations of Zakai type on a bounded domain of $\mathbb{R}^d$ is derived. It is shown that the order of convergence of this scheme is 1 for convergence in mean square sense. For almost sure convergence the order of convergence is proved to be $1 - \epsilon$ for any $\epsilon > 0$.

1. INTRODUCTION

Let D be a bounded domain in $\mathbb{R}^d$, $d \in \mathbb{N}$, with piecewise smooth boundary ∂D. Consider a Zakai equation (cf. [16]) in D of the form

$$du_t(x) = L^* u_t(x)\, dt + u_t(x)\, \big(g(x), dw_t\big), \qquad t \geq 0,\ x \in D, \tag{1.1}$$

with zero Dirichlet boundary conditions on ∂D, and initial condition $u_{t=0} = v$. L^* is a second order elliptic differential operator of the form

$$L^* u = \frac{1}{2} \sum_{i,j=1}^{d} D_i\big(D_j(a_{ij} u)\big) - \sum_{i=1}^{d} D_i(f_i u), \qquad u \in C_c^2(D), \tag{1.2}$$

where D_i denotes the partial derivative in coordinate direction $i \in \{1, \dots, d\}$. f is a differentiable vector field on D, and the $d \times d$–matrix valued function $a = (a_{ij},\ i, j = 1, \dots, d)$ on D is assumed to be twice differentiable, symmetric, and positive definite. (For a more complete listing of our assumptions we refer the reader to Section 2.) Moreover, $w = (w_t,\ t \geq 0)$ is a standard β–dimensional Brownian motion, and g is a mapping from D into $\mathbb{R}^\beta$. Finally, $(\cdot, \cdot)$ stands for the Euclidean inner product in $\mathbb{R}^\beta$.

It is well-known (e.g., [12], [13], [10], [2], [5] and the references given there) that under appropriate conditions on the data, equation (1.1) has a unique strong solution.

In the present paper we address the problem of deriving a discretization scheme with respect to the time parameter $t \geq 0$, which in analogy with the well-known Milstein scheme for ordinary SDE's (e.g., [11], [9] and references quoted there) converges faster than the usual Euler–Maruyama scheme as the time step size decreases to zero. The almost sure convergence of the latter for parabolic equations with additive noise has been treated in [3], and for the Zakai equation (1.1) it has been given

---

2000 *Mathematics Subject Classification.* 60G60, 60H15, 60H35, 65C30, 41A25.

Key words and phrases. Milstein scheme, stochastic partial differential equations, Zakai equation, numerical scheme.

in [4]. In these papers it is shown that the order of almost sure convergence for the Euler-Maruyama scheme is $1/2 - \epsilon$ for any $\epsilon > 0$. (Moreover, in [4] a Galerkin approximation in the space variable in terms of a certain eigenfunction expansion is given.) The Milstein scheme for the SPDE (1.1) derived in Section 2 will exhibit for almost sure convergence an order $1 - \epsilon$ for any $\epsilon > 0$. We find it interesting to observe that in order to achieve this order of convergence it is *not* enough to add the analogue of the expression which is well-known from the case of ordinary SDE's, but an additional term has to be built in. (For details cf. section 2.)

Concerning the discretization of the space variable, appropriate Galerkin approximations, their combination with the Milstein scheme presented here, and the generalization to the case of Lévy noise are currently under investigation.

The paper is organized as follows. In Section 2, we setup our framework, derive the Milstein scheme, and state the main results. The proofs of these results are done in Sections 3.

Acknowledgement. It is a pleasure to thank Andrea Barth, Vadim Kostrykin, and Martin Schmidt for helpful discussions. Two of us (A. L. and J. P.) acknowledge gratefully the very warm hospitality extended to them during their stay at the Department of Mathematics, Wayne State University, Detroit, Michigan, in 2008.

2. FRAMEWORK, MILSTEIN SCHEME, AND MAIN RESULTS

First we reformulate the SPDE (1.1) as an Itô equation with values in a Hilbert space (e.g., [2, 5]). To this end, let H denote the Hilbert space $L^2(D)$, and denote by H_0^m the Sobolev space of order $m \in \mathbb{N}$ on D with zero boundary values. (I.e., H_0^m is the completion of $C_c^\infty(D)$ with respect to the Sobolev norm of order m.) We denote $V = H_0^1$, and V' is the dual of V. As usual, we use Riesz' theorem to identify H with its dual, and thereby obtain the dense, continuous embeddings $V \subset H \subset V'$. Moreover, in the sequel we shall follow the common abuse of language to call elements in these spaces "functions". Moreover, we denote $K = \mathbb{R}^\beta$.

Following [4], we write the operator L^* (cf. Equation (1.2)) in divergence form

$$L^* = A + B, \tag{2.1a}$$

where for $u \in C_c^2(D)$

$$Au = \frac{1}{2} \sum_{i,j=1}^{d} D_i(a_{ij} D_j u), \tag{2.1b}$$

$$Bu = \sum_{i=1}^{d} D_i(b_i u), \tag{2.1c}$$

$$b_i = \frac{1}{2} \sum_{j}^{d} D_j a_{ij} - f_i. \tag{2.1d}$$

Similarly as in [4] we make the following

Assumptions 2.1. The coefficients of L^* and the initial condition v satisfy the following conditions:

(a) $a_{ij}, f_i, g_k, i, j = 1, \ldots, d, k = 1, \ldots, \beta$ of L^* belong to $C_b^3(D)$,

(b) the diffusion matrix a is strictly positive definite, i.e., there exists $\delta > 0$ so that for all $x \in D, \xi \in \mathbb{R}^d$,

$$\sum_{i,j=1}^{d} a_{ij}(x)\,\xi_i\xi_j \geq \delta\|\xi\|^2$$

(c) v is positive and belongs to $C_b^2(D)$.

We shall make use of the following lemma — whose statement is also known as Kato's conjecture — which has been proved in [1].

Lemma 2.2. *Let $A_{1/2} := (-A)^{1/2}$ and $A_{-1/2} := A_{1/2}^{-1}$, then $\mathcal{D}(A_{1/2}) = H_0^1$ and the norm $\|A_{1/2}\cdot\|_H$ is equivalent to $\|\cdot\|_{H^1}$, i.e.*

$$\|A_{1/2}\,\phi\|_H \leq C\|\phi\|_{H^1} \quad \textit{and} \quad \|\phi\|_{H^1} \leq C\|A_{1/2}\,\phi\|_H$$

for all $\phi \in H^1$.

For $u \in H$, we define a linear operator $G(u)$ from K into H by

$$\big(G(u)y\big)(x) = \big(g(x), y\big)\,u(x), \qquad y \in K,\ x \in D. \tag{2.2}$$

Then we can interpret the Zakai equation (1.1) as the following H–valued Itô SDE:

$$du_t = (Au_t + Bu_t)\,dt + G(u_t)\,dw_t, \tag{2.3}$$

subject to the initial condition $u_0 = v$. Assumption 2.1(b) implies that the operator A is dissipative, see [8]. Then by the Lumer–Phillips-Theorem, e.g. [6], A generates a strongly continuous contraction semigroup on H which we denote by $S = (S_t,\ t \geq 0)$. We may rewrite this initial value problem in the form of the following integral equation:

$$u_t = S_t v + \int_0^t S_{t-s}Bu_s\,ds + \int_0^t S_{t-s}G(u_s)\,dw_s. \tag{2.4}$$

The integral equation (2.4) is the starting point for our derivation of the Milstein scheme for (1.1).

We shall always consider a finite time horizon: $t \in [0, T]$ with $T < +\infty$. Let $\mathcal{T} = (\mathcal{T}_m,\ m \in \mathbb{N})$ be a sequence of partitions $\mathcal{T}_m$, $m \in \mathbb{N}$, of the interval $[0, T]$ whose mesh Δ_m tends to zero as m tends to $+\infty$. We write $\mathcal{T}_m$ as $\{t_0^m, t_1^m, \ldots, t_{n_m}^m\}$ with $n_m \in \mathbb{N}$, $0 = t_0^m < t_1^m < \cdots < t_{n_m}^m = T$, and

$$\Delta_m = \max_i (t_{i+1}^m - t_i^m),$$

the maximum being taken over $i \in \{0, \ldots, n_m - 1\}$. In the sequel we assume that Δ_m converges to zero at least polynomially in $1/m$, i.e., we suppose that there exists $\delta > 0$ so that $\Delta_m = O(m^{-\delta})$. For $m \in \mathbb{N}$, we define the map $\pi_m : [0, T] \to \{t_i^m,\ i = 0, \ldots, n_m\}$ by $\pi_m(s) = t_i$ if $t_i \leq s < t_{i+1}$.

As in [4], the integral equation (2.4) gives us directly the *Euler–Maruyama* scheme for approximate solutions u^m

$$u_t^m = S_t v + \int_0^t S_{t-\pi_m(s)} B u^m_{\pi_m(s)}\, ds + \int_0^t S_{t-\pi_m(s)} G\big(u^m_{\pi_m(s)}\big)\, dw_s. \tag{EM}$$

If one restricts to times $t \in \mathcal{T}_m$, this equation is readily converted into a recursive scheme (cf. [4]) which can be simulated on a computer, provided that one has made an appropriate discretization of the semigroup S, viz. of the differential operators in the space variables, too. It is proved in [4] that the approximate solutions u^m converge almost surely to u in such a way that

$$\sup_{0\le t\le T} \|u_t - u_t^m\|_H = O(\Delta_m^{1/2-\epsilon}), \qquad P\text{–a.s.}, \tag{2.5}$$

for any $\epsilon > 0$.

With Equation (2.4) and the semigroup property of S we get for $r,\ t \in [0,T]$, $r < t$,

$$u_t = S_{t-r} u_r + \int_r^t S_{t-s} B u_s\, ds + \int_r^t S_{t-s} G(u_s)\, dw_s. \tag{2.6}$$

Inserting this equation back into (2.4), and putting $r = \pi_m(s)$ we find

$$\begin{aligned} u_t = S_t v &+ \int_0^t S_{t-s} B\Big(u_{\pi_m(s)} + \int_{\pi_m(s)}^s S_{s-r} B u_r\, dr \\ &\qquad\qquad + \int_{\pi_m(s)}^s S_{s-r} G(u_r)\, dw_r\Big)\, ds \\ &+ \int_0^t S_{t-s} G\Big(u_{\pi_m(s)} + \int_{\pi_m(s)}^s S_{s-r} B u_r\, dr \\ &\qquad\qquad + \int_{\pi_m(s)}^s S_{s-r} G(u_r)\, dw_r\Big)\, dw_s. \end{aligned} \tag{2.7}$$

Formula (2.7) suggests the following *Milstein* scheme for an approximate solution u^m:

$$\begin{aligned} u_t^m = S_t v + \int_0^t S_{t-\pi_m(s)} B u^m_{\pi_m(s)}\, ds &+ \int_0^t S_{t-\pi_m(s)} G\big(u^m_{\pi_m(s)}\big)\, dw_s \\ &+ \int_0^t S_{t-\pi_m(s)} B\Big(\int_{\pi_m(s)}^s G\big(u^m_{\pi_m(s)}\big)\, dw_r\Big)\, ds \\ &+ \int_0^t S_{t-\pi_m(s)} G\Big(\int_{\pi_m(s)}^s G\big(u^m_{\pi_m(s)}\big)\, dw_r\Big)\, dw_s. \end{aligned} \tag{M}$$

Observe that in comparison to the Euler–Maruyama scheme we have the last two terms in addition. Superficially, the first of these may appear as being of order $\Delta_m^{3/2}$, but — roughly speaking — the combination of the semigroup with the first order differential operator B introduces an integrable singularity of the type $(t-\pi_m(s))^{-1/2}$, yielding a contribution of the order Δ_m of this term. (Of course, this rough argument gets its precise meaning in the statement and the proof of Theorems 2.3, 2.4, and 2.6 below.)

Note that all random variables involved with Equation (2.8) can easily be simulated on a computer. The double Itô integral can be simulated for each component as known for standard Milstein schemes (e.g., [9]): If B is a one-dimensional Brownian motion, we have

$$\int_{t_i^m}^{t_{i+1}^m} \int_{t_i^m}^{s} dB_r\, dB_s = \frac{1}{2}\left((B_{t_{i+1}^m} - B_{t_i}^m)^2 - (t_{i+1}^m - t_i^m)\right).$$

A simple simulation method for the mixed ("$dw_r\, ds$") integral can be found in [9]: The components of this integral are normally distributed with mean zero, variance $1/3\,(t_{i+1}^m - t_i^m)^3$, and covariance with the given Brownian motion $1/2\,(t_{i+1}^m - t_i^m)^2$. Let U_1, U_2 be independent centered normally distributed random variables with variance $t_{i+1}^m - t_i^m$. Then samples of

$$\int_{t_i}^{t_{i+1}} \int_{t_i}^{s} dB_r\, ds$$

can be generated via

$$\frac{1}{2}\,(t_{i+1} - t_i)\left(U_1 + \frac{1}{\sqrt{3}}\,U_2\right).$$

If we restrict the time variable t to the grid $\mathcal{T}_m$ and set $\Delta t_j = t_{j+1} - t_j$ and $\Delta w_{t_j}^i = w_{t_{j+1}}^i - w_{t_j}^i$, we get the following iterative scheme

$$\begin{aligned} u_{t_{j+1}}^m &= S_{t_{j+1}-t_j} u_{t_j}^m + S_{t_{j+1}-t_j} B u_{t_j}^m\, \Delta t_j + S_{t_{j+1}-t_j} G\big(u_{t_j}^m\big)\, \Delta w_{t_j} \\ &\quad + S_{t_{j+1}-t_j} BG\big(u_{t_j}^m\big)\, \eta(t_j, t_{j+1}) \\ &\quad + S_{t_{j+1}-t_j}\Big(u_{t_j}^m \sum_{i=1}^{\beta} g_i^2\, \frac{1}{2}\left((\Delta w_{t_j}^i)^2 - \Delta t_j\right)\Big), \end{aligned} \tag{2.8}$$

where we have set $\eta(a, b) = \int_a^b \int_a^s dw_r\, ds$.

For a mapping ϕ from $[0, T]$ into H we set

$$\|\phi\|_{H,T} = \sup_{0\le t\le T} \|\phi_t\|_H,$$

and for an H–valued stochastic process Φ, and $p \ge 1$, we define the norm

$$\|\Phi\|_{p,H,T} = \Big(\mathbb{E}\big(\sup_{0\le t\le T} \|\Phi_t\|_H^p\big)\Big)^{1/p}.$$

Now we can state our main results for the approximation $(u^m,\ m \in \mathbb{N})$ defined by the Milstein scheme (M):

Theorem 2.3 (Mean Square Convergence)**.** *There is a constant C so that for all $m \in M$ large enough,*

$$\|u - u^m\|_{2,H,T} \le C\, \Delta_m. \tag{2.9}$$

Theorem 2.4 (Almost Sure Convergence)**.** *For every $\epsilon > 0$ there is a constant C_ϵ so that for all $m \in \mathbb{N}$ large enough,*

$$\|u - u^m\|_{H,T} \leq C_\epsilon \, \Delta_m^{1-\epsilon}, \qquad P\text{–a.s.} \tag{2.10}$$

Finally we consider the situation where our choices of $K = \mathbb{R}^\beta$ and w are generalized to K being a separable Hilbert space, and w being a Q–Wiener process in K in the sense of [2,5]. Here Q is a trace class operator from K into K, and we assume furthermore that the following condition is satisfied:

Assumption 2.5. The coefficients of L^* and the initial condition v satisfy Assumption 2.1. Furthermore the operators G and Q satisfy the following conditions:

(a) G is a linear mapping from H into the Hilbert–Schmidt operators from K into H that satisfies for $\phi \in H$: $\|G(\phi)Q^{1/2}\|_{L_{HS}(K,H)} \leq C\|\phi\|_H$,
(b) $\|AG(\phi)Q^{1/2}\|_{L_{HS}(K,H)} \leq C\|\phi\|_{H^2}$ for $\phi \in \mathcal{D}(A)$,
(c) $\|A_{1/2}G(\phi)Q^{1/2}\|_{L_{HS}(K,H)} \leq C\|\phi\|_{H^1}$ for $\phi \in \mathcal{D}(A_{1/2})$.

Theorem 2.6 (Infinite Dimensional Noise)**.** *Under Assumption 2.5 the statements of Theorem 2.3 and 2.4 remain true for a Q–Wiener process w.*

The proofs of these theorems are carried out in the next section.

3. Proofs

The essential machinery for the proofs of all three theorems is contained in the proof of Theorem 2.4. Therefore we begin with its proof.

3.1. **Proof of Theorem 2.4.** In this section we show almost sure convergence of the Milstein scheme by proving L^p convergence and applying a Borel–Cantelli argument. We will divide the proof into several lemmas. First one shows by direct calculations the following properties of the operators of Equation (2.3):

Lemma 3.1. *For A, B, and G defined as in Section 2 it holds that*

(a) $\|A\phi\|_H \leq C\|\phi\|_{H^2}$ *for all* $\phi \in H^2$,
(b) $\|B\phi\|_H \leq C\|\phi\|_{H^1}$ *for all* $\phi \in H^1$,
(c) $\|B\phi\|_{H^1} \leq C\|\phi\|_{H^2}$ *for all* $\phi \in H^2$,
(d) $\|G(\phi)\|_{L_{HS}(K,H)} \leq C\|\phi\|_H$ *for all* $\phi \in H$,
(e) $\|G(\phi)\|_{L_{HS}(K,H^1)} \leq C\|\phi\|_{H^1}$ *for all* $\phi \in H^1$,
(f) $\|A\,G(\phi)\|_{L_{HS}(K,H)} \leq C\|\phi\|_{H^2}$ *for all* $\phi \in H^2$.

While S generates a contraction semigroup on H, in general it does not generate a semigroup of that type on H^1 and H^2. The following lemma states some properties of S that are direct consequences of Lemma 3.1 and Theorem 2.6.13 in [15].

Lemma 3.2. *The following hold true:*

(a) $\|S_t\phi\|_{H^1} \leq C\|\phi\|_{H^1}$ *for* $\phi \in H^1$,
(b) $\|S_t\phi\|_{H^2} \leq C\|\phi\|_{H^2}$ *for* $\phi \in H^2$,
(c) $\|(S_t - \mathbb{1})\,G(\phi)\|_{L_{HS}(K,H)} \leq C\,t^{\alpha/2}\|\phi\|_{H^\alpha}$ *for* $\phi \in H^\alpha$ *and* $\alpha = 1, 2$.

In particular, S generates a strongly continuous semigroup on the two mentioned Sobolev spaces.

Finally, before we can start to prove the error estimates, we have to look at the solution of Equation (2.3). By Example 6.31 in [5] there exists a unique solution in H^2 and by Lemma 3 in [4] $\|u\|_{p,H^1,T}^p$ is bounded. Similarly one can show that $\|u\|_{p,H^2,T}^p$ is also bounded because the solution of Equation (2.3) is smoother than the one in [4], i.e., it is in H^2. The regularity of the solution is given in the following lemma, where we set for a mapping Φ from $[0,T]$ into H

$$\|\Phi\|_{p,H,r,R} = \Big(\mathbb{E}\big(\sup_{r\le t\le R}\|\Phi_t\|_H^p\big)\Big)^{1/p}.$$

Lemma 3.3. *If u is solution of Equation* (2.3)*, then*

$$\|u-u_r\|_{p,H,r,R}^p \le C\,\|u\|_{p,H,T}^p\,(R-r)^{p/2}, \qquad 0\le r\le R\le T,\ p\ge 2.$$

Proof. Let

$$\phi_t = \int_0^t S_{t-s}\,Bu_s\,ds$$

and

$$\psi_t = \int_0^t S_{t-s}\,G(u_s)\,dw_s,$$

then we can estimate with Lemma 3.1, Theorem 2.6.13 in [14], the theorem in [15], and Lemma 2.2

$$\begin{aligned}
\|u-u_r\|_{p,H,r,R}^p &\le 4^{p-1}\Big(\|(S-S_r)v\|_{p,H,r,R}^p + \|(S_{\cdot-r}-\mathbb{1})(\phi_r+\psi_r)\|_{p,H,r,R}^p \\
&\qquad\qquad + \|\phi-\phi_r\|_{p,H,r,R}^p + \|\psi-\psi_r\|_{p,H,r,R}^p\Big) \\
&\le C\,\big(\mathbb{E}(\|A_{1/2}v\|_H^p) + \|u\|_{p,H,T}^p\big)(R-r)^{p/2} \\
&\le C\,\|u\|_{p,H,T}^p\,(R-r)^{p/2}. \qquad \square
\end{aligned}$$

The error estimates of $\|u-u^m\|_{p,H,T}$ are done in several steps. First we observe that the difference of the solution and the approximate solution can be split into the deterministic and the stochastic part of the mild solution

$$u-u^m = \xi^m + \eta^m,$$

where the deterministic part is split again into three parts

$$\xi^m = \xi_1^m + \xi_2^m + \xi_3^m$$

with

$$\xi_1^m(t) = \int_0^t \Big(S_{t-s}\, B\, S_{s-\pi_m(s)} u_{\pi_m(s)} - S_{t-\pi_m(s)}\, B u^m_{\pi_m(s)} \Big)\, ds,$$

$$\xi_2^m(t) = \int_0^t \Big(S_{t-s}\, B \int_{\pi_m(s)}^s S_{s-r}\, B u_r\, dr \Big)\, ds,$$

$$\xi_3^m(t) = \int_0^t \Big(S_{t-s}\, B \int_{\pi_m(s)}^s S_{s-r}\, G(u_r)\, dw_r - S_{t-\pi_m(s)}\, B \int_{\pi_m(s)}^s G\big(u^m_{\pi_m(s)}\big)\, dw_r \Big)\, ds,$$

and similarly, the stochastic integral is decomposed as

$$\eta^m = \eta_1^m + \eta_2^m + \eta_3^m$$

with

$$\eta_1^m(t) = \int_0^t \Big(S_{t-s}\, G\big(S_{s-\pi_m(s)} u_{\pi_m(s)}\big) - S_{t-\pi_m(s)}\, G\big(u^m_{\pi_m(s)}\big) \Big)\, dw_s,$$

$$\eta_2^m(t) = \int_0^t \Big(S_{t-s}\, G\big(\int_{\pi_m(s)}^s S_{s-r}\, B u_r\, dr \big) \Big)\, dw_s,$$

$$\eta_3^m(t) = \int_0^t \Big(S_{t-s}\, G\big(\int_{\pi_m(s)}^s S_{s-r}\, G(u_r)\, dw_r \big) - S_{t-\pi_m(s)}\, G\big(\int_{\pi_m(s)}^s G\big(u^m_{\pi_m(s)}\big)\, dw_r \big) \Big)\, dw_s.$$

We will only give explicit estimates for ξ_3^m which exhibit all the techniques needed for the estimates of the other five terms.

Lemma 3.4. *The term ξ_3^m satisfies*

$$\|\xi_3^m\|_{p,H,T}^p \leq C_1\, \Delta_m^p + C_2 \int_0^T \|u - u^m\|_{p,H,s}^p\, ds.$$

Proof. For an H-valued random field Φ indexed by $T \times T$ we set

$$|\Phi|_{p,H,T} = \mathbb{E}\Big(\sup_{0 \leq t \leq T} \big(\int_0^t \|\Phi(t,s)\|_H\, ds \big)^p \Big)^{1/p}.$$

We split the integral into four parts and get with Hölder's inequality

$$
\begin{aligned}
\|\xi_3^m(t)\|_{p,H,T}^p \le 4^{p-1}\Big(&\big|S_{t-s}\,(\mathbb{1}-S_{s-\pi_m(s)})B\int_{\pi_m(s)}^s S_{s-r}G(u_r)\,dw_r\big|_{p,H,T}^p \\
&+\big|S_{t-\pi_m(s)}\,B\int_{\pi_m(s)}^s (S_{s-r}-\mathbb{1})\,G(u_r)\,dw_r\big|_{p,H,T}^p \\
&+\big|S_{t-\pi_m(s)}\,B\int_{\pi_m(s)}^s G(u_r-u_{\pi_m(s)})\,dw_r\big|_{p,H,T}^p \\
&+\big|S_{t-\pi_m(s)}\,B\int_{\pi_m(s)}^s G\big(u_{\pi_m(s)}-u^m_{\pi_m(s)}\big)\,dw_r\big|_{p,H,T}^p\Big).
\end{aligned}
$$

Next we use the commutativity of the semigroup and apply Theorem 2.6.13 in [14] and Lemma 3.1 to the first term. We remark that for all terms we can use Lemma 2.2 and Theorem 2.6.13 in [14] to show that for $\phi \in H$

$$
\|S_s\,B\phi\|_H = \|S_s\,A_{1/2}\,A_{-1/2}\,B\phi\|_H \le C\,s^{-1/2}\,\|\phi\|_H.
$$

Hence we obtain the bound

$$
\begin{aligned}
\|\xi_3^m(t)\|_{p,H,T}^p \le C\,\Big(&\Delta_m^{p/2}\,\big|(t-s)^{-1/2}\,B\int_{\pi_m(s)}^s S_{s-r}\,G(u_r)\,dw_r\big|_{p,H,T}^p \\
&+\big|(t-s)^{-1/2}\int_{\pi_m(s)}^s (S_{s-r}-\mathbb{1})\,G(u_r)\,dw_r\big|_{p,H,T}^p \\
&+\big|(t-s)^{-1/2}\int_{\pi_m(s)}^s G\big(u_r-u_{\pi_m(s)}\big)\,dw_r\big|_{p,H,T}^p \\
&+|(t-s)^{-1/2}\int_{\pi_m(s)}^s G\big(u_{\pi_m(s)}-u^m_{\pi_m(s)}\big)\,dw_r\big|_{p,H,T}^p\Big).
\end{aligned}
$$

Hölder's inequality for $p > 2$, Fubini's theorem, and the Burkholder–Davis–Gundy inequality for Q-Wiener processes [7] lead to

$$
\begin{aligned}
\|\xi_3^m(t)\|_{p,H,T}^p \le C\int_0^T\Big(&\Delta_m^{p/2}\,\mathbb{E}\Big(\int_{\pi_m(s)}^s \|G(u_r)\|^2_{L_{HS}(K,H^1)}\,dr\Big)^{p/2} \\
&+\mathbb{E}\Big(\int_{\pi_m(s)}^s \|(S_{s-r}-\mathbb{1})\,G(u_r)\|^2_{L_{HS}(K,H)}\,dr\Big)^{p/2} \\
&+\mathbb{E}\Big(\int_{\pi_m(s)}^s \|G(u_r-u_{\pi_m(s)})\|^2_{L_{HS}(K,H)}\,dr\Big)^{p/2} \\
&+\mathbb{E}\Big(\int_{\pi_m(s)}^s \|G\big(u_{\pi_m(s)}-u^m_{\pi_m(s)}\big)\|^2_{L_{HS}(K,H)}\,dr\Big)^{p/2}\Big)\,ds,
\end{aligned}
$$

since by the results in [1] the norm $\|A_{1/2}\cdot\|_H$ is equivalent to the H^1–norm. Finally we use Lemma 3.1, apply Lemma 3.2 to the second term, Lemma 3.3 to the third one, and to the last term Hölder's inequality one more time. So we get

$$\begin{aligned}\|\xi_3^m(t)\|_{p,H,T}^p &\le C\Big(\|u\|_{p,H^1,T}^p\,\Delta_m^p + \|u\|_{p,H^1,T}^p\,\Delta_m^p + \|u\|_{p,H,T}^p\,\Delta_m^p \\ &\qquad + \Delta_m^{p/2}\int_0^T \|u-u^m\|_{p,H,s}^p\,ds\Big) \\ &= C_1\,\Delta_m^p + C_2\int_0^T \|u-u^m\|_{p,H,s}^p\,ds\end{aligned}$$

and the lemma is proved. □

Now we can proceed to the

Proof of Theorem 2.4. Similar estimates as in Lemma 3.4 for the other five terms give for $p > 2$

$$\begin{aligned}\|u-u^m\|_{p,H,T}^p &\le 6^{p-1}\Big(\|\xi_1^m\|_{p,H,T}^p + \|\xi_2^m\|_{p,H,T}^p + \|\xi_3^m\|_{p,H,T}^p \\ &\qquad + \|\eta_1^m\|_{p,H,T}^p + \|\eta_2^m\|_{p,H,T}^p + \|\eta_3^m\|_{p,H,T}^p\Big) \\ &\le C_1\,\|u\|_{p,H^2,T}^p\,\Delta_m^p + C_2\int_0^T \|u-u^m\|_{p,H,s}^p\,ds.\end{aligned}$$

An application of Gronwall's inequality yields

$$\|u-u^m\|_{p,H,T}^p \le C\,\Delta_m^p. \tag{3.1}$$

Finally let $\epsilon > 0$, then Chebyshev's inequality implies

$$P(\|u-u^m\|_{H,T} \ge \Delta_m^{1-\epsilon}) \le \Delta_m^{-(1-\epsilon)p}\,\|u-u^m\|_{p,H,T}^p \le C\,\Delta_m^{p\epsilon}.$$

Since $\Delta_m = O(m^{-\delta})$, the corresponding series is convergent for any $p > (\epsilon\delta)^{-1}$ and therefore by the Borel–Cantelli lemma we get that

$$\|u-u^m\|_{H,T} \le C_\epsilon\,\Delta_m^{1-\epsilon}, \qquad P\text{–a.s.},$$

and the proof is concluded. □

3.2. **Proof of Theorem 2.3.** Mean square convergence of the Milstein scheme follows immediately as a special case from the proof of Theorem 2.4 in the previous subsection. Therefore we will just show how to derive the result from the proof of almost sure convergence which actually is a proof of L^p convergence for all $p > 2$. In that proof, Equation 3.1 and Hölder's inequality yield for $p > 2$

$$\|u-u^m\|_{2,H,T} \le \|u-u^m\|_{p,H,T} \le C\Delta_m,$$

which proves Theorem 2.3.

3.3. **Proof of Theorem 2.6.** The proof of Theorem 2.6 is similar to the proof of Theorem 2.4, and will not be done explicitly. The only differences that occur are terms of the form

$$\|G(\Phi)Q^{1/2}\|_{L_{HS}(K,H)}$$

with $\Phi \in H$, where the operator A may appear as well. But due to Assumption 2.5, these terms can be handled in the same way as in Theorem 2.4.

REFERENCES

1. P. Auscher and P. Tchamitchian, *Square roots of elliptic second order divergence operators on strongly Lipschitz domains: L^2 theory*, J. Anal. Math. **90** (2003), 1–12.
2. P.-L. Chow, *Stochastic Partial Differential Equations*, Chapman and Hall / CRC, Boca Raton, London, New York, 2007.
3. P. L. Chow and J. L. Jiang, *Almost Sure Convergence of Some Approximate Solutions for Random Parabolic Equations*, Random Partial Differential Equations (Basel), Birkhäuser Verlag, 1991, pp. 45–54.
4. P. L. Chow, J. L. Jiang, and J.L. Menaldi, *Pathwise convergence of approximate solutions to Zakai's equation in a bounded domain*, Stochastic partial differential equations and applications. (G. Da Prato and L. Tubaro, eds.), vol. 268, Longman Scientific & Technical. Pitman Res. Notes Math. Ser., 1992, pp. 111–123.
5. G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Encyclopedia of Mathematics and Its Applications. 44. Cambridge: Cambridge University Press. xviii, 454 p., 1992 (English).
6. Klaus-Jochen Engel and Rainer Nagel, *One-parameter semigroups for linear evolution equations.*, Graduate Texts in Mathematics. 194. Berlin: Springer. xxi, 586 p., 2000 (English).
7. E. Hausenblas and J. Seidler, *A note on maximal inequality for stochastic convolutions*, Czech. Math. J. **51** (2001), 785–790.
8. Tosio Kato, *Perturbation theory for linear operators. Corr. printing of the 2nd ed.*, Grundlehren der mathematischen Wissenschaften, 132. Berlin – Heidelberg – New York: Springer. xxi, 619 p., 1980 (English).
9. P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, Spinger-Verlag, Berlin, Heidelberg, New York, 1992.
10. N. V. Krylov and B. Rozovskii, *Stochastic partial differential equations and diffusion processes*, Russian Math. Surveys **37** (1982), 81–105.
11. G. Milstein, *Numerical integration of stochastic differential equations*, Kluwer Academic Publishers, Dordrecht, 1994.
12. E. Pardoux, *Stochastic partial differential equation for the density of conditional law of a diffusion process with boundary*, Stochastic Analysis (A. Friedman and M. Pinsky, eds.), Academic Press, New York, 1978.
13. ______, *Stochastic partial differential equations and filtering of diffusion processes*, Stochastics **3** (1979), 127–167.
14. A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Springer-Verlag, Berlin, Heidelberg, New York, 1983.
15. L. Tubaro, *An estimate of Burkholder type for stochastic processes defined by the stochastic integral*, Stochastic Anal. Appl. **2** (1984), 187–192.
16. M. Zakai, *On the optimal filtering of diffusion processes*, Z. Wahrsch. verw. Geb. **11** (1969), 230–243.

ANNIKA LANG
FAKULTÄT FÜR MATHEMATIK UND INFORMATIK, UNIVERSITÄT MANNHEIM
D–68131 MANNHEIM, GERMANY
E-mail address: lang@math.uni-mannheim.de

PAO-LIU CHOW
DEPARTMENT OF MATHEMATICS, WAYNE STATE UNIVERSITY
DETROIT, MI 48202, USA
E-mail address: plchow@math.wayne.edu

JÜRGEN POTTHOFF
FAKULTÄT FÜR MATHEMATIK UND INFORMATIK, UNIVERSITÄT MANNHEIM
D–68131 MANNHEIM, GERMANY
E-mail address: potthoff@math.uni-mannheim.de