WONG-ZAKAI APPROXIMATIONS WITH CONVERGENCE RATE FOR STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

TOSHIYUKI NAKAYAMA AND STEFAN TAPPE

ABSTRACT. The goal of this paper is to prove a convergence rate for Wong-Zakai approximations of semilinear stochastic partial differential equations driven by a finite dimensional Brownian motion. Several examples, including the HJMM equation from mathematical finance, illustrate our result.

1. INTRODUCTION

Consider a semilinear stochastic partial differential equation (SPDE) of the form

$$(1.1) \qquad \begin{cases} dX(t) & = & \big(AX(t) + b(X(t))\big)dt + \sum_{j=1}^{r} \sigma_j(X(t))dB^j(t) \\ X(0) & = & x_0 \end{cases}$$

on a separable Hilbert space $(H, \| \cdot \|)$ driven by a finite dimensional Brownian motion $B = (B^1, \ldots, B^r)$ for some positive integer $r \in \mathbb{N}$.

A natural method in order to approximate the SPDE (1.1) by a sequence of partial differential equations (PDEs) is to use the so-called Wong-Zakai approximations. More precisely, on a fixed time interval $[0, T]$ we replace the Brownian motions B^j by their polygonal approximations $(B_m^j)_{m \in \mathbb{N}}$ with step size $\frac{T}{m}$. For each $m \in \mathbb{N}$ the Wong-Zakai approximation $\xi_m(\cdot) = \xi_m(\cdot, \omega) : [0, T] \to H$ is the mild solution to the deterministic PDE

$$(1.2) \qquad \begin{cases} \dot{\xi}_m(t) & = & A\xi_m(t) + b(\xi_m(t)) - \frac{1}{2}\sum_{j=1}^{r} D\sigma_j(\xi_m(t))\sigma_j(\xi_m(t)) \\ & & + \sum_{j=1}^{r} \sigma_j(\xi_m(t))\dot{B}_m^j(t) \\ \xi_m(0) & = & x_0 \end{cases}$$

for each $\omega \in \Omega$. Under appropriate regularity conditions, the Wong-Zakai approximations $(\xi_m)_{m \in \mathbb{N}}$ converge to the solution X of the SPDE (1.1). More precisely, for every $p > 1$ we have the convergence

$$(1.3) \qquad \lim_{m \to \infty} \mathbb{E}\left[\sup_{t \in [0,T]} \|\xi_m(t) - X(t)\|^{2p} \right] = 0,$$

see [23, Thm. 2.1].

Such a convergence result has first been proven, in the case of finite dimensional SDEs, by Wong and Zakai, see [37, 38]. Their approximation result has been generalized into several directions; namely, to the infinite dimensional case, e.g. in [2, 3, 7, 8, 13, 17, 18, 19, 20, 23, 29] and [31]–[35], with a view to support theorems, e.g. in [3, 4, 15, 16, 22, 23] (we also mention the related viability result from [24]), with a view to the theory of rough paths, e.g. in [12], for driving processes with jumps, e.g. in [20, 27], and with a driving fractional Brownian motion, e.g. in [30].

However, there are only very few reference dealing with convergence rates for the Wong-Zakai approximations. In [17] and [18] the authors consider the particular situation where the SPDE (1.1) is a second-order SPDE of parabolic type, and in [20] it is assumed that the operator A appearing in (1.1) is the infinitesimal generator of a compact and analytic semigroup.

Our goal in the present paper is to establish a convergence rate for (1.3) without imposing restrictions on the generator A appearing in (1.1), that is, A is allowed to be the infinitesimal generator of an arbitrary strongly continuous semigroup.

In order to present our main result, let us briefly outline the assumptions on the drift $b : H \to H$ and the volatilities $\sigma_1, \ldots, \sigma_r : H \to H$; the precise mathematical framework is stated in Section 2. First, we assume that these coefficients satisfy standard regularity conditions:

1.1. **Assumption.** *We suppose that the following conditions are fulfilled:*
   (1) *The drift b is Lipschitz continuous and bounded.*
   (2) *We have $\sigma_j \in C_b^2(H)$ for each $j = 1, \ldots, r$.*

Here $C_b^2(H)$ denotes the space of all $\sigma \in C^2(H)$ such that σ, $D\sigma$ and $D^2\sigma$ are bounded. Then the volatilities $\sigma_1, \ldots, \sigma_r$ are Lipschitz continuous and bounded, and the mapping

$$(1.4) \qquad \rho : H \to H, \quad \rho(x) := \sum_{j=1}^r D\sigma_j(x)\sigma_j(x)$$

appearing in the PDE (1.2) is Lipschitz continuous and bounded, too, which ensures existence and uniqueness of mild solutions to the SPDE (1.1) and the PDE (1.2).

Furthermore, we assume that the conditions stated above are also fulfilled when we consider the coefficients as mappings on the domain $\mathcal{D}(A)$ of the generator with respect to the graph norm

$$(1.5) \qquad \|x\|_{\mathcal{D}(A)} = \sqrt{\|x\|^2 + \|Ax\|^2}, \quad x \in \mathcal{D}(A).$$

More precisely:

1.2. **Assumption.** *We suppose that the following conditions are fulfilled:*
   (1) *We have $b(\mathcal{D}(A)) \subset \mathcal{D}(A)$ and $\sigma_j(\mathcal{D}(A)) \subset \mathcal{D}(A)$ for each $j = 1, \ldots, r$.*
   (2) *The drift $b|_{\mathcal{D}(A)}$ is Lipschitz continuous and bounded with respect to the graph norm $\| \cdot \|_{\mathcal{D}(A)}$.*
   (3) *We have $\sigma_j|_{\mathcal{D}(A)} \in C_b^2(\mathcal{D}(A))$ for each $j = 1, \ldots, r$ with respect to the graph norm $\| \cdot \|_{\mathcal{D}(A)}$.*

Then our main result reads as follows:

1.3. **Theorem.** *Suppose that Assumptions 1.1 and 1.2 are fulfilled, and let $T > 0$, $p > 1$ and $x_0 \in \mathcal{D}(A)$ be arbitrary. Then there is a constant $C > 0$ such that for each $m \in \mathbb{N}$ we have*

$$\mathbb{E}\left[\sup_{t \in [0,T]} \|\xi_m(t) - X(t)\|^{2p} \right] \leq \frac{C}{m^{p-1}},$$

where X denotes the mild solution to the SPDE (1.1) with $X(0) = x_0$, and the $(\xi_m)_{m \in \mathbb{N}}$ denote the mild solutions to the PDEs (1.2) with $\xi_m(0) = x_0$.

The proof of Theorem 1.3 will be a consequence of the following two results:
   (1) First, we will prove the stated convergence rate for the Euler-Maruyama approximations; see Theorem 3.1.
   (2) Then, we will prove the stated convergence rate for the difference between the Euler-Maruyama approximations and the Wong-Zakai approximations; see Theorem 4.1.

For both steps, we will use and extend some results from [23].

The remainder of this text is organized as follows. In Section 2 we introduce the mathematical framework and present some preliminary results. In Section 3 we provide the stated convergence rate for the Euler-Maruyama approximations, and in Section 4 we provide the stated convergence rate for the difference between the Euler-Maruyama approximations and the Wong-Zakai approximations. In the remaining sections we present examples of SPDEs where our main result (Theorem 1.3) applies. These are the HJMM equation from mathematical finance in Section 5, and two further examples arising from natural sciences in Section 6.

2. General framework and notation

In this section, we introduce the mathematical framework and present some preliminary results. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions. Let $B^1, \ldots, B^r$ be independent standard Brownian motions for some positive integer $r \in \mathbb{N}$. Let H be a separable Hilbert space and let $(S_t)_{t \geq 0}$ be a C_0-semigroup on H with infinitesimal generator $A : \mathcal{D}(A) \subset H \to H$. Furthermore, let $b : H \to H$ and $\sigma_1, \ldots, \sigma_r : H \to H$ be measurable mappings.

2.1. **Lemma.** *Suppose that $\sigma_1, \ldots, \sigma_r \in C_b^2(H)$. Then the mapping $\rho : H \to H$ defined in (1.4) is Lipschitz continuous and bounded.*

Proof. By assumption, there exists a constant $C > 0$ such that

$$\max\{\|\sigma_j(x)\|, \|D\sigma_j(x)\|, \|D^2\sigma_j(x)\|\} \leq C \quad \text{for all } x \in H \text{ and } j = 1, \ldots, r.$$

Therefore, for each $x \in H$ we obtain

$$\|\rho(x)\| \leq \sum_{j=1}^r \|D\sigma_j(x)\sigma_j(x)\| \leq \sum_{j=1}^r \|D\sigma_j(x)\| \, \|\sigma_j(x)\| \leq rC^2,$$

proving that ρ is bounded. Now, let $x_1, x_2 \in H$ be arbitrary. Then we have

$$\|\rho(x_1) - \rho(x_2)\| \leq \sum_{j=1}^r \|D\sigma_j(x_1)\sigma_j(x_1) - D\sigma_j(x_2)\sigma_j(x_2)\|$$

$$\leq \sum_{j=1}^r \|D\sigma_j(x_1)\| \, \|\sigma_j(x_1) - \sigma_j(x_2)\| + \sum_{j=1}^r \|\sigma_j(x_2)\| \, \|D\sigma_j(x_1) - D\sigma_j(x_2)\|$$

$$\leq 2rC^2 \|x_1 - x_2\|,$$

showing that ρ is Lipschitz continuous. $\qquad \square$

We fix a finite time horizon $T > 0$, and define the quantities

$$\delta_m := \frac{T}{m}, \quad m \in \mathbb{N},$$
$$[t]_m^- := k\delta_m, \quad k\delta_m \leq t < (k+1)\delta_m, \quad k = 0, \ldots, m-1,$$
$$[t]_m^+ := (k+1)\delta_m, \quad k\delta_m \leq t < (k+1)\delta_m, \quad k = 0, \ldots, m-1,$$

and the real-valued processes $(B_m^j(t))_{t \in [0,T]}$ for $m \in \mathbb{N}$ and $j = 1 \ldots, r$ as

$$B_m^j(t) := B^j([t]_m^-) + \frac{t - [t]_m^-}{\delta_m}(B^j([t]_m^+) - B^j([t]_m^-)), \quad t \in [0, T].$$

Note that for all $m \in \mathbb{N}$ we have

$$[t]_m^- \leq t < [t]_m^+, \quad t \in [0, T],$$

and that for all $m \in \mathbb{N}$ and all $j = 1 \ldots, r$ we have

$$(2.1) \qquad \dot{B}_m^j(t) = \frac{B^j([t]_m^+) - B^j([t]_m^-)}{\delta_m}, \quad t \in [0, T].$$

For what follows, we suppose that Assumptions 1.1 and 1.2 are fulfilled, and consider the SPDE

$$(2.2) \qquad \begin{cases} dX(t) & = & \big(AX(t) + \hat{b}(X(t))\big)dt + \sum_{j=1}^r \sigma_j(X(t))dB^j(t) \\ X(0) & = & x_0, \end{cases}$$

where the mapping $\hat{b} : H \to H$ is given by $\hat{b} := b + \frac{\rho}{2}$ with $\rho : H \to H$ being defined in (1.4), and for each $m \in \mathbb{N}$ we consider the Wong-Zakai approximation $\xi_m(\cdot) = \xi_m(\cdot, \omega) : [0, T] \to H$ given by the PDE

$$(2.3) \qquad \begin{cases} \dot{\xi}_m(t) & = & A\xi_m(t) + b(\xi_m(t)) + \sum_{j=1}^r \sigma_j(\xi_m(t))\dot{B}_m^j(t) \\ \xi_m(0) & = & x_0 \end{cases}$$

for each $\omega \in \Omega$.

2.2. **Remark.** *From now on, we consider the SPDE (2.2) and the PDEs (2.3) instead of (1.1) and (1.2). For our purposes, this is more convenient, as then we are directly in the framework of [23, Sec. 2], and it does not mean a restriction by virtue of Lemma 2.1.*

For each $m \in \mathbb{N}$ we define the Euler-Maruyama approximation Y_m inductively as follows. We set $Y_m(0) := x_0$, and, provided that Y_m is defined on the interval $[0, k\delta_m]$ for some $k \in \{0, \ldots, m - 1\}$, we set

$$(2.4) \qquad \begin{aligned} Y_m(t) &:= S_{t-k\delta_m} Y_m(k\delta_m) + \int_{k\delta_m}^t S_{t-s} \hat{b}(Y_m(k\delta_m))ds \\ &+ \sum_{j=1}^r \int_{k\delta_m}^t S_{t-s} \sigma_j(Y_m(k\delta_m))dB^j(s), \quad t \in [k\delta_m, (k+1)\delta_m]. \end{aligned}$$

Note that for $A = 0$ (that is $S_t = \mathrm{Id}$ for each $t \geq 0$) the processes $(Y_m)_{m \in \mathbb{N}}$ coincide with the well-known Euler-Maruyama approximations with step sizes δ_m for SDEs.

2.3. **Remark.** *As pointed out in [5], the naive implementation*

$$(2.5) \qquad \begin{aligned} Y_m(t) &:= Y_m(k\delta_m) + \int_{k\delta_m}^t \big(AY_m(k\delta_m) + \hat{b}(Y_m(k\delta_m))\big)ds \\ &+ \sum_{j=1}^r \int_{k\delta_m}^t \sigma_j(Y_m(k\delta_m))dB^j(s), \quad t \in [k\delta_m, (k+1)\delta_m]. \end{aligned}$$

of the Euler-Maruyama method does not work, because it might immediately lead to some $Y_m(k\delta_m) \notin \mathcal{D}(A)$. Even in our situation, where we have a well-defined strong solution (see Proposition 2.8 below), there is no reason why the discrete approximation (2.5) should always stay in $\mathcal{D}(A)$.

2.4. **Lemma.** *For each $m \in \mathbb{N}$ we have*

$$(2.6) \qquad \begin{aligned} Y_m(t) &= S_t x_0 + \int_0^t S_{t-s} \hat{b}(Y_m([s]_m^-))ds \\ &+ \sum_{j=1}^r \int_0^t S_{t-s} \sigma_j(Y_m([s]_m^-))dB^j(s), \quad t \in [0, T]. \end{aligned}$$

Proof. We prove identity (2.6) inductively on each interval $[0, k\delta_m]$ for $k = 0, \ldots, m$. The identity (2.6) holds true for $k = 0$, because $Y_m(0) = x_0$. For the induction step $k \to k+1$ note that for $t = k\delta_m$ identity (2.6) yields

$$
\begin{aligned}
(2.7) \qquad Y_m(k\delta_m) = S_{k\delta_m}x_0 &+ \int_0^{k\delta_m} S_{k\delta_m - s}\hat{b}(Y_m([s]_m^-))ds \\
&+ \sum_{j=1}^r \int_0^{k\delta_m} S_{k\delta_m - s}\sigma_j(Y_m([s]_m^-))dB^j(s).
\end{aligned}
$$

Therefore, by (2.4) and (2.7), and noting that

$$
[s]_m^- = k\delta_m, \quad s \in [k\delta_m, (k+1)\delta_m),
$$

for each $t \in [k\delta_m, (k+1)\delta_m]$ we obtain

$$
\begin{aligned}
Y_m(t) = S_{t-k\delta_m}\bigg(&S_{k\delta_m}x_0 + \int_0^{k\delta_m} S_{k\delta_m - s}\hat{b}(Y_m([s]_m^-))ds \\
&+ \sum_{j=1}^r \int_0^{k\delta_m} S_{k\delta_m - s}\sigma_j(Y_m([s]_m^-))dB^j(s)\bigg) \\
&+ \int_{k\delta_m}^t S_{t-s}\hat{b}(Y_m([s]_m^-))ds + \sum_{j=1}^r \int_{k\delta_m}^t S_{t-s}\sigma_j(Y_m([s]_m^-))dB^j(s),
\end{aligned}
$$

proving (2.6). $\qquad\square$

2.5. **Remark.** *With the terminology from* [9], *the Euler-Maruyama approximations (2.6) are so-called* accelerated exponential Euler approximations, *whereas the so-called* exponential Euler approximations *would be given by*

$$
\begin{aligned}
Y_m(t) = S_t x_0 &+ \int_0^t S_{t-[s]_m^-}\hat{b}(Y_m([s]_m^-))ds \\
&+ \sum_{j=1}^r \int_0^t S_{t-[s]_m^-}\sigma_j(Y_m([s]_m^-))dB^j(s), \quad t \in [0, T].
\end{aligned}
$$

The domain $\mathcal{D}(A)$ equipped with the graph norm (1.5) is a separable Hilbert space, too, and the restriction $(S_t|_{\mathcal{D}(A)})_{t\geq 0}$ is a C_0-semigroup on $(\mathcal{D}(A), \|\cdot\|_{\mathcal{D}(A)})$ with infinitesimal generator A on the domain $\mathcal{D}(A^2)$. In the upcoming results, the notation $(\mathcal{D}(A)-)\int_0^t$ indicates that we consider the respective integral on the state space $(\mathcal{D}(A), \|\cdot\|_{\mathcal{D}(A)})$; see, for example, the right-hand sides of (2.10) and (2.13). Otherwise, the integral is considered on the state space $(H, \|\cdot\|)$, as usual; see, for example, the left-hand sides of (2.10) and (2.13).

2.6. **Lemma.** *Let* $\Phi : [0, T] \to \mathcal{D}(A)$ *be a function such that*

$$
(2.8) \qquad \int_0^T \|\Phi(s)\|_{\mathcal{D}(A)}ds < \infty.
$$

Then the following statements are true:

(1) *We have*

$$
(2.9) \qquad \int_0^T \|\Phi(s)\|ds < \infty.
$$

(2) *For each* $t \in [0, T]$ *we have*

$$
(2.10) \qquad \int_0^t \Phi(s)ds = (\mathcal{D}(A)-)\int_0^t \Phi(s)ds.
$$

Proof. Relation (2.9) is an immediate consequence of (2.8). There is a sequence $(\Phi^n)_{n \in \mathbb{N}}$ of simple functions $\Phi^n : [0,T] \to \mathcal{D}(A)$ such that $\|\Phi - \Phi^n\|_{\mathcal{D}(A)} \leq \|\Phi\|_{\mathcal{D}(A)}$ for each $n \in \mathbb{N}$, and we have $\|\Phi - \Phi^n\|_{\mathcal{D}(A)} \to 0$ for $n \to \infty$. Let $t \in [0,T]$ be arbitrary. By Lebesgue's dominated convergence theorem we obtain

$$\left\| \big(\mathcal{D}(A)-\big) \int_0^t \Phi(s)ds - \big(\mathcal{D}(A)-\big) \int_0^t \Phi^n(s)ds \right\|_{\mathcal{D}(A)} \to 0 \quad \text{as } n \to \infty.$$

We have $\|\Phi - \Phi^n\| \leq \|\Phi\|_{\mathcal{D}(A)}$ for each $n \in \mathbb{N}$, and $\|\Phi - \Phi^n\| \to 0$ for $n \to \infty$. Therefore, by Lebesgue's dominated convergence theorem we also have

$$\left\| \int_0^t \Phi(s)ds - \int_0^t \Phi^n(s)ds \right\| \to 0 \quad \text{as } n \to \infty.$$

Noting that

$$\int_0^t \Phi^n(s)ds = \big(\mathcal{D}(A)-\big) \int_0^t \Phi^n(s)ds \quad \text{for each } n \in \mathbb{N},$$

we arrive at (2.10). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

2.7. Lemma. *Let Ψ be a $\mathcal{D}(A)$-valued predictable process such that $\mathbb{P}$-almost surely*

$$(2.11) \qquad\qquad \int_0^T \|\Psi(s)\|_{\mathcal{D}(A)}^2 ds < \infty.$$

Then the following statements are true:

    (1) *We have*

$$(2.12) \qquad\qquad \int_0^T \|\Psi(s)\|^2 ds < \infty.$$

    (2) *For each $t \in [0,T]$ and each $j = 1, \ldots, r$ we have*

$$(2.13) \qquad\qquad \int_0^t \Psi(s)dB^j(s) = \big(\mathcal{D}(A)-\big) \int_0^t \Psi(s)dB^j(s).$$

Proof. The proof is similar to that of Lemma 2.6, and therefore omitted. $\qquad$ □

The following three results show that for each starting point $x_0 \in \mathcal{D}(A)$ the mild solution X to the SPDE (2.2) with $X(0) = x_0$, the Wong-Zakai approximations $(\xi_m)_{m \in \mathbb{N}}$ given by the PDEs (2.3) with $\xi_m(0) = x_0$ and the Euler-Maruyama approximations $(Y_m)_{m \in \mathbb{N}}$ given by $Y_m(0) = x_0$ and (2.4) take their values in $\mathcal{D}(A)$.

2.8. Proposition. *For each $x_0 \in \mathcal{D}(A)$ there exists a unique mild solution X to the SPDE (2.2) on the state space $(\mathcal{D}(A), \|\cdot\|_{\mathcal{D}(A)})$, and it is a strong solution to the SPDE (2.2) on the state space H.*

Proof. By a standard result (see, for example, [10, Thm. 7.2]), there is a a unique mild solution X to the SPDE (2.2) on the state space $(\mathcal{D}(A), \|\cdot\|_{\mathcal{D}(A)})$; that is, a continuous adapted process such that $\mathbb{P}$-almost surely

$$X(t) = S_t x_0 + \big(\mathcal{D}(A)-\big) \int_0^t S_{t-s}\hat{b}(X(s))ds$$

$$+ \sum_{j=1}^r \big(\mathcal{D}(A)-\big) \int_0^t S_{t-s}\sigma_j(X(s))dB^j(s), \quad t \in \mathbb{R}_+.$$

This implies that X is also continuous in H, and by Lemmas 2.6 and 2.7 we obtain $\mathbb{P}$-almost surely

$$X(t) = S_t x_0 + \int_0^t S_{t-s} \hat{b}(X(s)) ds$$
$$+ \sum_{j=1}^r \int_0^t S_{t-s} \sigma_j(X(s)) dB^j(s), \quad t \in \mathbb{R}_+,$$

showing that X is also a mild solution to the SPDE (2.2) on the state space H. Furthermore, since X is continuous in $(\mathcal{D}(A), \|\cdot\|_{\mathcal{D}(A)})$ we have $\mathbb{P}$-almost surely

$$\int_0^t \|AX(s)\| ds \le \int_0^t \|X(s)\|_{\mathcal{D}(A)} ds < \infty \quad \text{for each } t \in \mathbb{R}_+.$$

Therefore X is also a strong solution to the SPDE (2.2) on the state space H. $\square$

2.9. **Proposition.** *For each $x_0 \in \mathcal{D}(A)$ and each $m \in \mathbb{N}$ there exists a unique mild solution ξ_m to the PDE (2.3) on the state space $(\mathcal{D}(A), \|\cdot\|_{\mathcal{D}(A)})$, and it is a strong solution to the PDE (2.3) on the state space H. Moreover, for each $k \in \{0, \ldots, m-1\}$ we have*

$$\xi_m(t) = S_{t-k\delta_m} \xi_m(k\delta_m) + \int_{k\delta_m}^t S_{t-s} b(\xi_m(s)) ds$$
$$+ \sum_{j=1}^r \int_{k\delta_m}^t S_{t-s} \sigma_j(\xi_m(s)) \dot{B}_m^j(s) ds, \quad t \in [k\delta_m, (k+1)\delta_m].$$

Proof. The proof is similar to that of Proposition 2.8 and Lemma 2.4, and therefore omitted. $\square$

2.10. **Proposition.** *For each $x_0 \in \mathcal{D}(A)$ and each $m \in \mathbb{N}$ we have $Y_m \in \mathcal{D}(A)$.*

Proof. Noting that $Y_m(0) = x_0$ and (2.4), this is an immediate consequence of Lemmas 2.6 and 2.7. $\square$

2.11. **Lemma.** *Let Ψ be an H-valued predictable process, and let $p > 1$ be such that*

$$\mathbb{E}\left[\int_0^T \|\Psi(s)\|^{2p} ds \right] < \infty.$$

Then there is a constant $C > 0$ such that for each $j = 1, \ldots, r$ we have

$$\mathbb{E}\left[\sup_{t \in [0,T]} \left\| \int_0^T S_{t-s} \Psi(s) dB^j(s) \right\|^{2p} \right] \le C \, \mathbb{E}\left[\int_0^T \|\Psi(s)\|^{2p} ds \right].$$

Proof. This follows, for example, from [14, Lemma 3.3]. $\square$

2.12. **Lemma.** *There is a constant $C > 0$ such that for all $t_1, t_2 \in [0,T]$ with $t_1 \le t_2$ and all $x \in \mathcal{D}(A)$ we have*

(2.14) $$\|S_{t_2} x - S_{t_1} x\| \le C \|x\|_{\mathcal{D}(A)} |t_2 - t_1|.$$

Proof. According to [25, Thm. 2.2] there are constants $M \ge 1$ and $\omega \in \mathbb{R}$ such that

$$\|S_t\| \le M e^{\omega t} \quad \text{for all } t \ge 0.$$

Therefore, by [25, Thm. 2.4] we obtain

$$\|S_{t_2} x - S_{t_1} x\| = \left\| \int_{t_1}^{t_2} S_s A x \, ds \right\| \le \int_{t_1}^{t_2} \|S_s A x\| ds$$
$$\le \int_{t_1}^{t_2} \|S_s\| \, \|Ax\| ds \le M e^{\omega T} \|x\|_{\mathcal{D}(A)} |t_2 - t_1|,$$

proving (2.14) with $C = Me^{\omega T}$. □

3. Convergence rate for the Euler-Maruyama approximations

In this section, we prove the stated convergence rate for the Euler-Maruyama approximations. The general mathematical framework is that of Section 2.

3.1. **Theorem.** *Suppose that Assumptions 1.1 and 1.2 are fulfilled, and let $T > 0$, $p > 1$ and $x_0 \in \mathcal{D}(A)$ be arbitrary. Then there is a constant $C > 0$ such that for each $m \in \mathbb{N}$ we have*

$$\mathbb{E}\left[\sup_{t \in [0,T]} \|Y_m(t) - X(t)\|^{2p} \right] \leq \frac{C}{m^{p-1}},$$

where X denotes the mild solution to the SPDE (2.2) with $X(0) = x_0$, and the $(Y_m)_{m \in \mathbb{N}}$ denote Euler-Maruyama approximations given by $Y_m(0) = x_0$ and (2.4).

We will provide the proof of Theorem 3.1 at the end of this section. For each $m \in \mathbb{N}$ we introduce the processes $\bar{Y}_m$ and $\bar{X}_m$ as

$$\bar{Y}_m(t) := S_{t-[t]_m^-} Y_m([t]_m^-), \quad t \in \mathbb{R}_+,$$
$$\bar{X}_m(t) := S_{t-[t]_m^-} X([t]_m^-), \quad t \in \mathbb{R}_+.$$

Then, for each $m \in \mathbb{N}$ we have

(3.1) $Y_m - X = (Y_m - \bar{Y}_m) + (\bar{Y}_m - \bar{X}_m) + (\bar{X}_m - X)$.

In the upcoming proofs, we will denote by C a suitable positive constant, possibly different from line to line, but only depending on T, p, x_0 and the parameters (A, b, σ) of the SPDE (2.2).

3.2. **Proposition.** *There is a constant $C > 0$ such that for each $m \in \mathbb{N}$ we have*

$$\mathbb{E}\left[\sup_{t \in [0,T]} \|Y_m(t) - \bar{Y}_m(t)\|^{2p} \right] \leq C\delta_m^{p-1}.$$

Proof. We have

$$Y_m(t) - \bar{Y}_m(t) = \int_{[t]_m^-}^t S_{t-s}\hat{b}(Y_m([s]_m^-))ds + \sum_{j=1}^r \int_{[t]_m^-}^t S_{t-s}\sigma_j(Y_m([s]_m^-))dB^j(s).$$

Since

$$\left\| \int_{[t]_m^-}^t S_{t-s}\hat{b}(Y_m([s]_m^-))ds \right\| \leq C\delta_m,$$

applying [23, Lemma 2.4] completes the proof. □

3.3. **Proposition.** *There is a constant $C > 0$ such that for each $m \in \mathbb{N}$ we have*

$$\mathbb{E}\left[\sup_{t \in [0,T]} \|\bar{X}_m(t) - X(t)\|^{2p} \right] \leq C\delta_m^{p-1}.$$

Proof. See [23, Lemma 2.5]. □

3.4. **Lemma.** *There is a constant $C > 0$ such that for each $m \in \mathbb{N}$ we have*

$$\mathbb{E}\left[\sup_{t \in [0,T]} \|\bar{Y}_m(t) - Y_m([t]_m^-)\|^{2p} \right] \leq C\delta_m^{2p}.$$

Proof. Note that

$$\bar{Y}_m(t) - Y_m([t]_m^-) = (S_{t-[t]_m^-} - \mathrm{Id})Y_m([t]_m^-).$$

Therefore, by Lemma 2.12 we obtain

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|\bar{Y}_m(t) - Y_m([t]_m^-)\|^{2p}\right] \leq C\delta_m^{2p}\mathbb{E}\left[\sup_{t\in[0,T]} \|Y_m([t]_m^-)\|_{\mathcal{D}(A)}^{2p}\right],$$

which, by virtue of Lemma 2.11 – applied with the separable Hilbert space $(\mathcal{D}(A), \|\cdot\|_{\mathcal{D}(A)})$ – and Assumption 1.2 completes the proof. $\qquad\square$

3.5. **Proposition.** *There is a constant $C > 0$ such that for each $m \in \mathbb{N}$ and each $v \in [0,T]$ we have*

$$\mathbb{E}\left[\sup_{t\in[0,v]} \|\bar{Y}_m(t) - \bar{X}_m(t)\|^{2p}\right] \leq C\left(\int_0^v \mathbb{E}\left[\sup_{t\in[0,u]} \|Y_m(t) - X(t)\|^{2p}\right]du + \delta_m^{p-1}\right).$$

Proof. Note that

$$\bar{Y}_m(t) - \bar{X}_m(t) = \int_0^{[t]_m^-} S_{t-s}\big(\hat{b}(Y([s]_m^-)) - \hat{b}(X(s))\big)ds$$

$$+ \sum_{j=1}^r \int_0^{[t]_m^-} S_{t-s}\big(\sigma_j(Y([s]_m^-)) - \sigma_j(X(s))\big)dB^j(s), \quad t \in [0,T].$$

We have

$$\mathbb{E}\left[\sup_{t\in[0,v]} \left\|\int_0^{[t]_m^-} S_{t-s}\big(\hat{b}(Y([s]_m^-)) - \hat{b}(X(s))\big)ds\right\|^{2p}\right]$$

$$\leq C\int_0^v \mathbb{E}[\|Y([s]_m^-) - X(s)\|^{2p}]ds.$$

Furthermore, by Lemma 2.11 we have

$$\mathbb{E}\left[\sup_{t\in[0,v]} \left\|\int_0^{[t]_m^-} S_{t-s}\big(\sigma_j(Y([s]_m^-)) - \sigma_j(X(s))\big)dB^j(s)\right\|^{2p}\right]$$

$$\leq C\int_0^v \mathbb{E}[\|Y([s]_m^-) - X(s)\|^{2p}]ds.$$

Therefore, we obtain

$$\mathbb{E}\left[\sup_{t\in[0,v]} \|\bar{Y}_m(t) - \bar{X}_m(t)\|^{2p}\right] \leq C\big(I_1(v) + I_2(v) + I_3(v)\big),$$

where we have set

$$I_1(v) := \int_0^v \mathbb{E}\left[\sup_{t\in[0,u]} \|Y([t]_m^-) - \bar{Y}_m(t)\|^{2p}\right]du,$$

$$I_2(v) := \int_0^v \mathbb{E}\left[\sup_{t\in[0,u]} \|\bar{Y}_m(t) - Y_m(t)\|^{2p}\right]du,$$

$$I_3(v) := \int_0^v \mathbb{E}\left[\sup_{t\in[0,u]} \|Y_m(t) - X(t)\|^{2p}\right]du.$$

Therefore, applying Proposition 3.2 and Lemma 3.4 completes the proof. $\qquad\square$

Now, the proof of Theorem 3.1 is an immediate consequence of the decomposition (3.1), Propositions 3.2, 3.3, 3.5 and Gronwall's inequality.

4. Distance between the Euler-Maruyama approximations and the Wong-Zakai approximations

In this section, we prove the stated convergence rate for the difference between the Euler-Maruyama approximations and the Wong-Zakai approximations. The general mathematical framework is that of Section 2.

4.1. **Theorem.** *Suppose that Assumptions 1.1 and 1.2 are fulfilled, and let $T > 0$, $p > 1$ and $x_0 \in \mathcal{D}(A)$ be arbitrary. Then there is a constant $C > 0$ such that for each $m \in \mathbb{N}$ we have*

$$\mathbb{E}\left[\sup_{t \in [0,T]} \|\xi_m(t) - Y_m(t)\|^{2p}\right] \leq \frac{C}{m^{p-1}},$$

where the $(\xi_m)_{m \in \mathbb{N}}$ denote the mild solutions to the PDEs (2.3) with $\xi_m(0) = x_0$, and the $(Y_m)_{m \in \mathbb{N}}$ denote the Euler-Maruyama approximations given by $Y_m(0) = x_0$ and (2.4).

We will provide the proof of Theorem 4.1 at the end of this section. For each $m \in \mathbb{N}$ we introduce the processes $\bar{\xi}_m$ and $\bar{Y}_m$ as

$$\bar{\xi}_m(t) := S_{t-[t]_m^-} \xi_m([t]_m^-), \quad t \in \mathbb{R}_+,$$
$$\bar{Y}_m(t) := S_{t-[t]_m^-} Y_m([t]_m^-), \quad t \in \mathbb{R}_+.$$

Then, for each $m \in \mathbb{N}$ we have

(4.1) $$\xi_m - Y_m = (\xi_m - \bar{\xi}_m) + (\bar{\xi}_m - \bar{Y}_m) + (\bar{Y}_m - Y_m).$$

4.2. **Proposition.** *There is a constant $C > 0$ such that for each $m \in \mathbb{N}$ we have*

$$\mathbb{E}\left[\sup_{t \in [0,T]} \|\xi_m - \bar{\xi}_m\|^{2p}\right] \leq C\delta_m^{p-1}.$$

Proof. See [23, Lemma 2.2].   □

We have the identity

$$\bar{\xi}_m(t) - \bar{Y}_m(t) = S_{t-[t]_m^-}\left(\xi_m([t]_m^-) - Y_m([t]_m^-)\right),$$

and hence

(4.2)

$$\bar{\xi}_m(t) - \bar{Y}_m(t) = S_{t-[t]_m^-}\bigg(\int_0^{[t]_m^-} S_{t-s}b(\xi_m(s))ds$$
$$+ \sum_{j=1}^r \int_0^{[t]_m^-} S_{t-s}\sigma_j(\xi_m(s))\dot{B}_m^j(s)ds - \int_0^{[t]_m^-} S_{t-s}b(Y_m([s]_m^-))ds$$
$$- \frac{1}{2}\sum_{j=1}^r \int_0^{[t]_m^-} S_{t-s}D\sigma_j(Y_m([s]_m^-))\sigma_j(Y_m([s]_m^-))ds$$
$$- \sum_{j=1}^r \int_0^{[t]_m^-} S_{t-s}\sigma_j(Y_m([s]_m^-))dB^j(s)\bigg).$$

Fix an arbitrary $j \in \{1, \ldots, r\}$. By equation (2.6) in [23], for each $m \in \mathbb{N}$ we have

(4.3) $$\sigma_j(\xi_m(t)) = \gamma_{m,1}^j(t) + \sum_{l=1}^r \gamma_{m,2}^{j,l}(t) + \gamma_{m,3}^j(t), \quad t \in [0,T],$$

where the quantity $\gamma_{m,1}^j(t)$ is given by

$$\gamma_{m,1}^j(t) := \sigma_j(\bar{\xi}_m(t)),$$

the quantities $\gamma_{m,2}^{j,l}(t)$ for $l = 1, \ldots, r$ are given by

$$\gamma_{m,2}^{j,l}(t) := D\sigma_j(\bar{\xi}_m(t)) \int_{[t]_m^-}^t S_{t-u}\sigma_l(\bar{\xi}_m(u))\dot{B}_m^l(u)du,$$

and the quantity $\gamma_{m,3}^j(t)$ is given by

$$\begin{aligned}
\gamma_{m,3}^j(t) := &\ D\sigma_j(\bar{\xi}_m(t)) \int_{[t]_m^-}^t S_{t-u}b(\xi_m(u))du \\
&+ \sum_{l=1}^r D\sigma_j(\bar{\xi}_m(t)) \int_{[t]_m^-}^t S_{t-u}\bigg(\int_0^1 D\sigma_l(\bar{\xi}_m(u)) \\
&+ v(\xi_m(u) - \bar{\xi}_m(u))(\xi_m(u) - \bar{\xi}_m(u))dv \bigg)\dot{B}_m^l(u)du \\
&+ \int_0^1 \bigg(\int_0^{v_1} D^2\sigma_j(\bar{\xi}_m(t) + v_2(\xi_m(t) - \bar{\xi}_m(t))) \\
&\quad (\xi_m(t) - \bar{\xi}_m(t), \xi_m(t) - \bar{\xi}_m(t)))dv_2 \bigg)dv_1.
\end{aligned}$$

We introduce the quantity $\zeta_{m,1}^j(t)$ as

$$\zeta_{m,1}^j(t) := \int_0^{[t]_m^-} S_{t-s}\gamma_{m,3}^j(s)\dot{B}_m^j(s)ds,$$

the quantity $\zeta_{m,2}^j(t)$ as

$$\zeta_{m,2}^j(t) := \int_0^{[t]_m^-} S_{t-s}\gamma_{m,1}^j(s)\dot{B}_m^j(s)ds - \int_0^{[t]_m^-} S_{t-s}\sigma_j(Y_m([s]_m^-))dB^j(s),$$

the quantities $\zeta_{m,3}^{j,l}(t)$ for $l = 1, \ldots, r$ as

$$\begin{aligned}
\zeta_{m,3}^{j,j}(t) := &\ \int_0^{[t]_m^-} S_{t-s}\gamma_{m,2}^{j,j}(s)\dot{B}_m^j(s)ds \\
&- \frac{1}{\delta_m} \int_0^{[t]_m^-} S_{t-s}D\sigma_j(\bar{\xi}_m(s)) \bigg(\int_{[s]_m^-}^s S_{s-u}\sigma_j(\bar{\xi}_m(u))du \bigg)ds
\end{aligned}$$

and

$$\zeta_{m,3}^{j,l}(t) := \int_0^{[t]_m^-} S_{t-s}\gamma_{m,2}^{j,l}(s)\dot{B}_m^j(s)ds \quad \text{for } l \neq j,$$

the quantity $\zeta_{m,4}^j(t)$ as

$$\begin{aligned}
\zeta_{m,4}^j(t) := &\ \frac{1}{\delta_m} \int_0^{[t]_m^-} S_{t-s}D\sigma_j(\bar{\xi}_m(s)) \bigg(\int_{[s]_m^-}^s S_{s-u}\sigma_j(\bar{\xi}_m(u))du \bigg)ds \\
&- \frac{1}{2} \int_0^{[t]_m^-} S_{t-s}D\sigma_j(Y_m([s]_m^-))\sigma_j(Y_m([s]_m^-))ds,
\end{aligned}$$

and the quantity $\zeta_{m,5}(t)$ as

$$\zeta_{m,5}(t) := \int_0^{[t]_m^-} S_{t-s}\big(b(\xi_m(s)) - b(Y_m([s]_m^-))\big)ds.$$

Then we obtain

$$\bar{\xi}_m(t) - \bar{Y}_m(t) = S_{t-[t]_m^-}$$

$$(4.4) \qquad \left(\sum_{j=1}^{r} \zeta_{m,1}^j(t) + \sum_{j=1}^{r} \zeta_{m,2}^j(t) + \sum_{j=1}^{r} \sum_{l=1}^{r} \zeta_{m,3}^{j,l}(t) + \sum_{j=1}^{r} \zeta_{m,4}^j(t) + \zeta_{m,5}(t) \right).$$

4.3. Proposition. *There is a constant $C > 0$ such that for each $m \in \mathbb{N}$ we have*

$$\mathbb{E}\left[\sup_{t \in [0,T]} \|\zeta_{m,1}^j(t)\|^{2p} \right] \leq C\delta_m^p.$$

Proof. This follows from [23, Lemma 2.6]. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

4.4. Proposition. *There is a constant $C > 0$ such that for each $m \in \mathbb{N}$ and each $v \in [0,T]$ we have*

$$\mathbb{E}\left[\sup_{t \in [0,v]} \|\zeta_{m,2}^j(t)\|^{2p} \right] \leq C\left(\int_0^v \mathbb{E}\left[\sup_{t \in [0,u]} \|\xi_m(t) - Y_m(t)\|^{2p} \right] du + \delta_m^{2p} \right).$$

Proof. An analogous calculation as in the proof of [23, Lemma 2.9] shows that

$$\mathbb{E}\left[\sup_{t \in [0,v]} \|\zeta_{m,2}^j(t)\|^{2p} \right] \leq C\mathbb{E}\left[\int_0^{[v]_m^-} \|\tilde{Y}_m^j(u)\|^{2p} du \right], \quad v \in [0,T],$$

where

$$\tilde{Y}_m^j(u) = \frac{1}{\delta_m} \int_{[u]_m^-}^{[u]_m^+} S_{[u]_m^+ - \tilde{u}} \sigma_j(\bar{\xi}_m(\tilde{u})) d\tilde{u} - S_{[u]_m^+ - u} \sigma_j(Y_m([u]_m^-)).$$

Now, we have

$$\tilde{Y}_m^j(u) = \frac{1}{\delta_m} \int_{[u]_m^-}^{[u]_m^+} S_{[u]_m^+ - \tilde{u}} \sigma_j(S_{\tilde{u} - [u]_m^-} \xi_m([u]_m^-)) d\tilde{u}$$

$$- \frac{1}{\delta_m} \int_{[u]_m^-}^{[u]_m^+} S_{[u]_m^+ - u} \sigma_j(Y_m([u]_m^-)) d\tilde{u},$$

and hence, we obtain

$$\tilde{Y}_m^j(u) = \frac{1}{\delta_m} \int_{[u]_m^-}^{[u]_m^+} S_{[u]_m^+ - \tilde{u}} \big(\sigma_j(S_{\tilde{u} - [u]_m^-} \xi_m([u]_m^-)) - \sigma_j(S_{\tilde{u} - [u]_m^-} Y_m([u]_m^-)) \big) d\tilde{u}$$

$$+ \frac{1}{\delta_m} \int_{[u]_m^-}^{[u]_m^+} S_{[u]_m^+ - \tilde{u}} \big(\sigma_j(S_{\tilde{u} - [u]_m^-} Y_m([u]_m^-)) - \sigma_j(Y_m([u]_m^-)) \big) d\tilde{u}$$

$$+ \frac{1}{\delta_m} \int_{[u]_m^-}^{[u]_m^+} \big(S_{[u]_m^+ - \tilde{u}} - S_{[u]_m^+ - u} \big) \sigma_j(Y_m([u]_m^-)) d\tilde{u}.$$

Therefore, and by Lemma 2.12 and Assumption 1.2 we obtain

$$\|\tilde{Y}_m^j(u)\| \leq C\|\xi_m([u]_m^-) - Y_m([u]_m^-)\|$$

$$+ C \sup_{\tilde{u} \in [0,\delta_m]} \|(S_{\tilde{u}} - \mathrm{Id}) Y_m([u]_m^-)\|$$

$$+ C \sup_{\substack{u_1, u_2 \in [0,T] \\ |u_1 - u_2| \leq \delta_m}} \|(S_{u_1} - S_{u_2}) \sigma_j(Y_m([u]_m^-))\|$$

$$\leq C\|\xi_m([u]_m^-) - Y_m([u]_m^-)\| + C\delta_m \big(\|Y_m([u]_m^-)\|_{\mathcal{D}(A)} + 1 \big).$$

By Lemma 2.11 – applied with the separable Hilbert space $(\mathcal{D}(A), \|\cdot\|_{\mathcal{D}(A)})$ – and Assumption 1.2 we obtain

$$
\mathbb{E}\left[\sup_{t \in [0,v]} \|\zeta_{m,2}^j(t)\|^{2p}\right] \leq C \int_0^{[v]_m^-} \mathbb{E}\left[\|\tilde{Y}_m^j(u)\|^{2p}\right] du
$$

$$
\leq C \int_0^v \left(\mathbb{E}[\|\xi_m([u]_m^-) - Y_m([u]_m^-)\|^{2p}] + \delta_m^{2p}\left(\mathbb{E}\left[\|Y_m([u]_m^-)\|_{\mathcal{D}(A)}^{2p}\right] + 1\right)\right) du
$$

$$
\leq C\left(\int_0^v \mathbb{E}\left[\sup_{t \in [0,u]} \|\xi_m(t) - Y_m(t)\|^{2p}\right] du + \delta_m^{2p}\right),
$$

finishing the proof. $\qquad\square$

4.5. Lemma. *Let* $X \sim \mathrm{N}(0, \sigma^2)$ *be a normally distributed random variable with variance* $\sigma^2 > 0$. *Then, for each positive real number* $q > 0$ *we have*

$$
\mathbb{E}\left[|X|^{2q}\right] = 2^q \frac{\Gamma(q + \frac{1}{2})}{\Gamma(\frac{1}{2})} \sigma^{2q}.
$$

Proof. We have $X = \sigma Y$ with a random variable $Y \sim \mathrm{N}(0,1)$. Therefore, we have $Y^2 \sim \chi^2$, and by [36, Sec. 7.8.1] we obtain

$$
\mathbb{E}\left[|X|^{2q}\right] = \mathbb{E}\left[|\sigma Y|^{2q}\right] = \sigma^{2q}\mathbb{E}\left[(Y^2)^q\right] = 2^q \frac{\Gamma(q + \frac{1}{2})}{\Gamma(\frac{1}{2})} \sigma^{2q},
$$

completing the proof. $\qquad\square$

4.6. Corollary. *Let* $q > 0$ *be a positive real number. Then there is a constant* $C > 0$ *such that for each* $m \in \mathbb{N}$ *we have*

$$
\mathbb{E}\left[|B^j((k+1)\delta_m) - B^j(k\delta_m)|^{2q}\right] \leq C\delta_m^q.
$$

Proof. Since $B^j((k+1)\delta_m) - B^j(k\delta_m) \sim \mathrm{N}(0, \delta_m)$ for each $k = 0, \ldots, m-1$, this is an immediate consequence of Lemma 4.5. $\qquad\square$

The proof of the following auxiliary result is similar to that of Lemma 2.2 in [23].

4.7. Lemma. *Let* $q > 0$ *be a positive real number. Then there is a constant* $C > 0$ *such that for each* $m \in \mathbb{N}$ *we have*

$$
\mathbb{E}\left[\sup_{t \in [0,T]} |\dot{B}_m^j(t)|^{2q}\right] \leq C\delta_m^{-(q+1)}.
$$

Proof. Taking into account (2.1), by Corollary 4.6 we obtain

$$
\mathbb{E}\left[\sup_{t \in [0,T]} |\dot{B}_m^j(t)|^{2q}\right] = \frac{1}{\delta_m^{2q}}\mathbb{E}\left[\sup_{t \in [0,T]} |B^j([t]_m^+) - B^j([t]_m^-)|^{2q}\right]
$$

$$
\leq \frac{1}{\delta_m^{2q}}\sum_{k=0}^{m-1}\mathbb{E}\left[|B^j((k+1)\delta_m) - B^j(k\delta_m)|^{2q}\right] \leq \frac{C}{\delta_m^{2q}}m\delta_m^q = C\delta_m^{-(q+1)},
$$

completing the proof. $\qquad\square$

4.8. Corollary. *Let* $q > 0$ *be a positive real number. Then there is a constant* $C > 0$ *such that for each* $m \in \mathbb{N}$ *we have*

$$
\mathbb{E}\left[\sup_{t \in [0,T]} \|\xi_m(t)\|_{\mathcal{D}(A)}^{2q}\right] \leq C\delta_m^{-(q+1)}.
$$

Proof. By Proposition 2.9, for each $m \in \mathbb{N}$ the process ξ_m is a solution to the $\mathcal{D}(A)$-valued integral equation

$$\xi_m(t) = S_t x_0 + \big(\mathcal{D}(A)-\big) \int_0^t S_{t-s}\bigg(b(\xi_m(s)) + \sum_{j=1}^r \sigma_j(\xi_m(s))\dot{B}_m^j(s) \bigg)ds, \quad t \in [0,T].$$

Therefore, taking into account Assumption 1.2, the stated estimate is an immediate consequence of Lemma 4.7. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

The following result contributes to [23, Lemma 2.11], where it was shown that merely under Assumption 1.1 (that is, without imposing Assumption 1.2) for each each $\alpha \in (\frac{1}{2p}, \frac{1}{2})$ there is a constant $C > 0$ such that for each $l \in \{1, \dots, r\}$ and each $m \in \mathbb{N}$ we have

$$\mathbb{E}\bigg[\sup_{t \in [0,T]} \|\zeta_{m,3}^{j,l}(t)\|^{2p} \bigg] \leq C\delta_m^{p(1-2\alpha)}.$$

4.9. **Proposition.** *For each $\kappa \in (0, p)$ there is a constant $C > 0$ such that for each $l \in \{1, \dots, r\}$ and each $m \in \mathbb{N}$ we have*

$$\mathbb{E}\bigg[\sup_{t \in [0,T]} \|\zeta_{m,3}^{j,l}(t)\|^{2p} \bigg] \leq C\delta_m^{\kappa}.$$

Proof. Let

$$I_m(t) := \delta_m^{-2} \int_0^{[t]_m^-} S_{t-s} D\sigma_j(\bar{\xi}_m(s)) \int_{[s]_m^-}^s S_{s-u}\sigma_l(\bar{\xi}_m(u))du K_m(s)ds,$$

where

$$K_m(s) := \begin{cases} (B^j([s]_m^+) - B^j([s]_m^-))^2 - \delta_m, & \text{if } j = l, \\ (B^j([s]_m^+) - B^j([s]_m^-))(B^l([s]_m^+) - B^l([s]_m^-)), & \text{if } j \neq l. \end{cases}$$

Then we have $\zeta_{m,3}^{j,l}(t) = I_m(t)$, and hence, it has to be shown that

$$\mathbb{E}\bigg[\sup_{t \in [0,T]} \|I_m(t)\|^{2p} \bigg] \leq C\delta_m^{\kappa}.$$

Note that

$$I_m(t) = \delta_m^{-2}(I_m^1(t) + I_m^2(t) + I_m^3(t)),$$

where the quantity $I_m^1(t)$ is defined as

$$I_m^1(t) := \int_0^{[t]_m^-} S_{t-s} D\sigma_j(\bar{\xi}_m(s))$$
$$\int_{[s]_m^-}^s S_{s-u}(\sigma_l(\bar{\xi}_m(u)) - \sigma_l(\xi_m([u]_m^-)))du K_m(s)ds,$$

the quantity $I_m^2(t)$ is defined as

$$I_m^2(t) := \int_0^{[t]_m^-} S_{t-s}\big(D\sigma_j(\bar{\xi}_m(s)) - D\sigma_j(\xi_m([s]_m^-))\big)$$
$$\int_{[s]_m^-}^s S_{s-u}\sigma_l(\xi_m([u]_m^-))du K_m(s)ds,$$

and the quantity $I_m^3(t)$ is defined as

$$I_m^3(t) := \int_0^{[t]_m^-} S_{t-s} D\sigma_j(\xi_m([s]_m^-))$$

$$\int_{[s]_m^-}^s S_{s-u}\sigma_l(\xi_m([u]_m^-)) du K_m(s) ds.$$

Next, we define

$$\tilde{K}_m(k) := \begin{cases} (B^j((k+1)\delta_m) - B^j(k\delta_m))^2 - \delta_m, & \text{if } j = l, \\ (B^j((k+1)\delta_m) - B^j(k\delta_m))(B^l((k+1)\delta_m) - B^l(k\delta_m)), & \text{if } j \neq l, \end{cases}$$

and we choose constants $\pi, \theta \in (1, \infty)$ such that

$$p - \frac{1}{\pi} > \kappa \quad \text{and} \quad \frac{1}{\pi} + \frac{1}{\theta} = 1.$$

Noting that B^j and B^l are independent for $j \neq l$, by Corollary 4.6 (applied with $q = 2p\theta$ in case $j = l$, and applied twice with $q = p\theta$ in case $j \neq l$) we have

$$\mathbb{E}\left[|\tilde{K}_m(k)|^{2p\theta}\right]^{1/\theta} \leq C\delta_m^{2p} \quad \text{for all } k = 0, \ldots, m-1.$$

Now, we set

$$U_m(t) := \sup_{0 \leq u \leq \delta_m} \|(S_u - \text{Id})\xi_m(t)\|.$$

By Lemma 2.12, for each $t \in [0, T]$ we have

$$U_m(t)^{2p\pi} = \sup_{0 \leq u \leq \delta_m} \|(S_u - \text{Id})\xi_m(t)\|^{2p\pi} \leq C \sup_{0 \leq u \leq \delta_m} u^{2p\pi}\|\xi_m(t)\|_{\mathcal{D}(A)}^{2p\pi}$$

$$\leq C\delta_m^{2p\pi}\|\xi_m(t)\|_{\mathcal{D}(A)}^{2p\pi}.$$

Therefore, by Corollary 4.8 (applied with $q = p\pi$) we obtain

$$\mathbb{E}\left[\max_{0 \leq k \leq m} U_m(k\delta_m)^{2p\pi}\right] \leq \mathbb{E}\left[\sup_{t \in [0,T]} U_m(t)^{2p\pi}\right] \leq C\delta_m^{2p\pi}\mathbb{E}\left[\sup_{t \in [0,T]} \|\xi_m(t)\|_{\mathcal{D}(A)}^{2p\pi}\right]$$

$$\leq C\delta_m^{2p\pi}\delta_m^{-(p\pi+1)} = C\delta_m^{p\pi-1}.$$

Now we have

$$\|I_m^1(t)\| \leq C\int_0^T \int_{[s]_m^-}^s \|(S_{u-[s]_m^-} - \text{Id})\xi_m([s]_m^-)\| du |K_m(s)| ds$$

$$\leq C\int_0^T (s - [s]_m^-)U_m([s]_m^-)|K_m(s)| ds$$

$$= C\frac{\delta_m^2}{2}\sum_{k=0}^{m-1} U_m(k\delta_m)|\tilde{K}_m(k)|.$$

Hence we obtain

$$\|I_m^1(t)\|^{2p} \leq C\left(\frac{\delta_m^2}{2}\right)^{2p} m^{2p-1} \times \sum_{k=0}^{m-1} U_m(k\delta_m)^{2p}|\tilde{K}_m(k)|^{2p}$$

$$\leq C\left(\frac{\delta_m^2}{2}\right)^{2p} m^{2p} \max_{0 \leq k \leq m} U_m(k\delta_m)^{2p}|\tilde{K}_m(k)|^{2p}.$$

By Hölder's inequality we conclude that

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|I_m^1(t)\|^{2p}\right]$$

$$\leq C\left(\frac{\delta_m^2}{2}\right)^{2p}m^{2p}\mathbb{E}\left[\max_{0\leq k\leq m}U_m(k\delta_m)^{2p\pi}\right]^{1/\pi}\mathbb{E}\left[|\tilde{K}_m(k)|^{2p\theta}\right]^{1/\theta}$$

$$\leq C\delta_m^{4p}\delta_m^{-2p}\delta_m^{p-\frac{1}{\pi}}\delta_m^{2p}\leq C\delta_m^{4p}\delta_m^{\kappa}.$$

Now we have

$$\|I_m^2(t)\|\leq C\int_0^T\|(S_{s-[s]_m^-}-\mathrm{Id})\xi_m([s]_m^-)\|(s-[s]_m^-)|K_m(s)|ds$$

$$\leq C\int_0^T U_m([s]_m^-)(s-[s]_m^-)|K_m(s)|ds.$$

So by the same argument as above, we also have

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|I_m^2(t)\|^{2p}\right]\leq C\delta_m^{4p}\delta_m^{\kappa}.$$

Next we have

$$I_m^3(t)=\sum_{k=0}^{\delta_m^{-1}[t]_m^- - 1}\int_{k\delta_m}^{(k+1)\delta_m}S_{t-s}D\sigma_j(\xi_m(k\delta_m))\int_{k\delta_m}^s S_{s-u}\sigma_l(\xi_m(k\delta_m))duK_m(s)ds.$$

Let us consider the process $M^m=(M_n^m)_{n=0,1,\ldots,m}$ defined by

$$M_n^m:=\sum_{k=0}^{n-1}\int_{k\delta_m}^{(k+1)\delta_m}S_{t-s}D\sigma_j(\xi_m(k\delta_m))\int_{k\delta_m}^s S_{s-u}\sigma_l(\xi_m(k\delta_m))duK_m(s)ds$$

for $n=1,2,\ldots,m$ with $M_0^m=0$. Then M^m is a $(\mathcal{F}_{n\delta_m})$-martingale, because $\mathbb{E}[K_m(s)]=0$ for all $s\in[0,T]$, and $\xi_m(k\delta_m)$ is $\mathcal{F}_{k\delta_m}$-measurable and $\tilde{K}_m(k)$ is independent of $\mathcal{F}_{k\delta_m}$ for all $k=0,\ldots,m-1$. Furthermore, by the Lemma 2.10 in the paper [23], we have

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|I_m^3(t)\|^{2p}\right]=\mathbb{E}\left[\max_{1\leq n\leq m}\|M_n^m(t)\|^{2p}\right]\leq C_p\mathbb{E}\left[\left(\sum_{k=0}^{m-1}\|M_{k+1}^m-M_k^m\|^2\right)^p\right]$$

$$=C_p\mathbb{E}\left[\left(\sum_{k=0}^{m-1}\left\|\int_{k\delta_m}^{(k+1)\delta_m}S_{t-s}D\sigma_j(\xi_m(k\delta_m))\right.\right.\right.$$

$$\left.\left.\left.\int_{k\delta_m}^s S_{s-u}\sigma_l(\xi_m(k\delta_m))duK_m(s)ds\right\|^2\right)^p\right],$$

where C_p is depending only on p, and hence

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|I_m^3(t)\|^{2p}\right]\leq C\mathbb{E}\left[\left(\sum_{k=0}^{m-1}\left(\int_{k\delta_m}^{(k+1)\delta_m}(s-k\delta_m)|K_m(s)|ds\right)^2\right)^p\right]$$

$$=C\left(\frac{\delta_m^2}{2}\right)^{2p}\mathbb{E}\left[\left(\sum_{k=0}^{m-1}|\tilde{K}_m(k)|^2\right)^p\right]\leq C\left(\frac{\delta_m^2}{2}\right)^{2p}m^{p-1}\sum_{k=0}^{m-1}\mathbb{E}\left[|\tilde{K}_m(k)|^{2p}\right]$$

$$\leq C\left(\frac{\delta_m^2}{2}\right)^{2p}m^p\delta_m^{2p}\leq C\delta_m^{5p}.$$

Combining above results, we have

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|I_m(t)\|^{2p}\right] \leq C\delta_m^{-4p}(\delta_m^{4p}\delta_m^{\kappa} + \delta_m^{5p}) = C(\delta_m^{\kappa} + \delta_m^{p}) \leq C\delta_m^{\kappa},$$

completing the proof. □

4.10. **Proposition.** *There is a constant $C > 0$ such that for each $m \in \mathbb{N}$ and each $v \in [0, T]$ we have*

$$\mathbb{E}\left[\sup_{t\in[0,v]}\|\zeta_{m,4}^{j}(t)\|^{2p}\right] \leq C\left(\int_0^v \mathbb{E}\left[\sup_{t\in[0,u]}\|\xi_m(t) - Y_m(t)\|^{2p}\right]du + \delta_m^{2p}\right).$$

Proof. The proof is similar to that of [23, Lemma 2.12]. Note that

$$\zeta_{m,4}^{j}(t) = I_{m,1}^{j}(t) + I_{m,2}^{j}(t) + I_{m,3}^{j}(t) + I_{m,4}^{j}(t),$$

where the quantity $I_{m,1}^{j}(t)$ is given by

$$I_{m,1}^{j}(t) = \frac{1}{\delta_m}\int_0^{[t]_m^-} S_{t-s}D\sigma_j(\bar{\xi}_m(s))\left(\int_{[s]_m^-}^s S_{s-u}\sigma_j(\bar{\xi}_m(u))du\right)ds$$
$$- \frac{1}{\delta_m}\int_0^{[t]_m^-} S_{t-s}D\sigma_j(\bar{Y}_m(s))\left(\int_{[s]_m^-}^s S_{s-u}\sigma_j(\bar{Y}_m(u))du\right)ds,$$

the quantity $I_{m,2}^{j}(t)$ is given by

$$I_{m,2}^{j}(t) = \frac{1}{\delta_m}\int_0^{[t]_m^-} S_{t-s}D\sigma_j(\bar{Y}_m(s))\left(\int_{[s]_m^-}^s S_{s-u}\sigma_j(\bar{Y}_m(u))du\right)ds$$
$$- \frac{1}{\delta_m}\int_0^{[t]_m^-} S_{t-s}D\sigma_j(Y_m([s]_m^-))\left(\int_{[s]_m^-}^s S_{s-u}\sigma_j(Y_m([u]_m^-))du\right)ds,$$

the quantity $I_{m,3}^{j}(t)$ is given by

$$I_{m,3}^{j}(t) = \frac{1}{\delta_m}\int_0^{[t]_m^-} S_{t-s}D\sigma_j(Y_m([s]_m^-))\left(\int_{[s]_m^-}^s S_{s-u}\sigma_j(Y_m([u]_m^-))du\right)ds$$
$$- \frac{1}{\delta_m}\int_0^{[t]_m^-} (s - [s]_m^-)S_{t-[s]_m^-}D\sigma_j(Y_m([s]_m^-))\sigma_j(Y_m([s]_m^-))ds,$$

and the quantity $I_{m,4}^{j}(t)$ is given by

$$I_{m,4}^{j}(t) = \frac{1}{\delta_m}\int_0^{[t]_m^-} (s - [s]_m^-)S_{t-[s]_m^-}D\sigma_j(Y_m([s]_m^-))\sigma_j(Y_m([s]_m^-))ds$$
$$- \frac{1}{2}\int_0^{[t]_m^-} S_{t-s}D\sigma_j(Y_m([s]_m^-))\sigma_j(Y_m([s]_m^-))ds.$$

We have

$$\|I_{m,1}^{j}(t)\| \leq \frac{1}{\delta_m}\left\|\int_0^{[t]_m^-} S_{t-s}D\sigma_j(\bar{\xi}_m(s))\right.$$
$$\left(\int_{[s]_m^-}^s S_{s-u}\big(\sigma_j(\bar{\xi}_m(u)) - \sigma_j(\bar{Y}_m(u))\big)du\right)ds\Big\|$$
$$+ \frac{1}{\delta_m}\left\|\int_0^{[t]_m^-} S_{t-s}\big(D\sigma_j(\bar{\xi}_m(s)) - D\sigma_j(\bar{Y}_m(s))\big)\right.$$
$$\left(\int_{[s]_m^-}^s S_{s-u}\sigma_j(\bar{Y}_m(u))du\right)ds\Big\|,$$

and hence

$$\|I_{m,1}^j(t)\| \le C\bigg[\frac{1}{\delta_m} \int_0^{[t]_m^-} \bigg(\int_{[s]_m^-}^s \|\bar{\xi}_m(u) - \bar{Y}_m(u)\| du \bigg) ds$$
$$+ \int_0^{[t]_m^-} \|\bar{\xi}_m(s) - \bar{Y}_m(s)\| ds \bigg]$$
$$\le C \int_0^{[v]_m^-} \|\xi_m([s]_m^-) - Y_m([s]_m^-)\| ds \le C \int_0^{[v]_m^-} \sup_{t \in [0,u]} \|\xi_m(t) - Y_m(t)\| du.$$

for all $0 \le t \le v \le T$. Therefore, we get

$$\mathbb{E}\bigg[\sup_{t \in [0,v]} \|I_{m,1}^j(t)\|^{2p} \bigg] \le C \int_0^v \mathbb{E}\bigg[\sup_{t \in [0,u]} \|\xi_m(t) - Y_m(t)\|^{2p} \bigg] du.$$

Furthermore, we have

$$\|I_{m,2}^j(t)\| = \frac{1}{\delta_m} \bigg\| \int_0^{[t]_m^-} S_{t-s} D\sigma_j(\bar{Y}_m(s))$$
$$\bigg(\int_{[s]_m^-}^s S_{s-u} \big(\sigma_j(\bar{Y}_m(u)) - \sigma_j(Y_m([u]_m^-)) \big) du \bigg) ds \bigg\|$$
$$+ \frac{1}{\delta_m} \bigg\| \int_0^{[t]_m^-} S_{t-s} \big(D\sigma_j(\bar{Y}_m(s)) - D\sigma_j(Y_m([s]_m^-)) \big)$$
$$\bigg(\int_{[s]_m^-}^s S_{s-u} \sigma_j(Y_m([u]_m^-)) du \bigg) ds \bigg\|,$$

and hence

$$\|I_{m,2}^j(t)\| \le C\bigg[\frac{1}{\delta_m} \int_0^{[t]_m^-} \bigg(\int_{[s]_m^-}^s \|(S_{u-[u]_m^-} - \mathrm{Id})Y_m([u]_m^-)\| du \bigg) ds$$
$$+ \int_0^{[t]_m^-} \|(S_{s-[s]_m^-} - \mathrm{Id})Y_m([s]_m^-)\| ds \bigg]$$
$$\le C \int_0^{[t]_m^-} \sup_{[s]_m^- \le u \le s} \|(S_{u-[u]_m^-} - \mathrm{Id})Y_m([u]_m^-)\| ds.$$

Therefore, by Lemma 2.12 we obtain

$$\mathbb{E}\bigg[\sup_{t \in [0,T]} \|I_{m,2}^j(t)\|^{2p} \bigg] \le C\mathbb{E}\bigg[\sup_{t \in [0,T]} \|(S_{t-[t]_m^-} - \mathrm{Id})Y_m([t]_m^-)\|^{2p} \bigg]$$
$$\le C\delta_m^{2p} \mathbb{E}\bigg[\sup_{t \in [0,T]} \|Y_m([t]_m^-)\|_{\mathcal{D}(A)}^{2p} \bigg].$$

Therefore, by virtue of Lemma 2.11 – applied with the separable Hilbert space $(\mathcal{D}(A), \|\cdot\|_{\mathcal{D}(A)})$ – and Assumption 1.2 we obtain

$$\mathbb{E}\bigg[\sup_{t \in [0,T]} \|I_{m,2}^j(t)\|^{2p} \bigg] \le C\delta_m^{2p}.$$

Furthermore, we have

$$\|I_{m,3}^j(t)\| \leq \frac{1}{\delta_m} \left\| \int_0^{[t]_m^-} S_{t-s} D\sigma_j(Y_m([s]_m^-)) \right.$$
$$\left. \left(\int_{[s]_m^-}^s (S_{s-u} - \mathrm{Id})\sigma_j(Y_m([s]_m^-))du \right)ds \right\|$$
$$+ \frac{1}{\delta_m} \left\| \int_0^{[t]_m^-} (s - [s]_m^-)\left(S_{t-s} - S_{t-[s]_m^-}\right) \right.$$
$$\left. D\sigma_j(Y_m([s]_m^-))\sigma_j(Y_m([s]_m^-))ds \right\|,$$

and hence

$$\|I_{m,3}^j(t)\| \leq C\left[\frac{1}{\delta_m} \int_0^{[t]_m^-} \left(\int_{[s]_m^-}^s \|(S_{s-u} - \mathrm{Id})\sigma_j(Y_m([s]_m^-))\|du \right)ds \right.$$
$$\left. + \int_0^{[t]_m^-} \|(S_{t-s} - S_{t-[s]_m^-})D\sigma_j(Y_m([s]_m^-))\sigma_j(Y_m([s]_m^-))\|ds \right].$$

Therefore, by Lemma 2.12 and Assumption 1.2, we obtain

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|I_{m,3}^j(t)\|^{2p} \right] \leq C\delta_m^{2p}.$$

Moreover, we have

$$\frac{1}{\delta_m} \int_0^{[t]_m^-} (s - [s]_m^-)S_{t-[s]_m^-}D\sigma_j(Y_m([s]_m^-))\sigma_j(Y_m([s]_m^-))ds$$
$$= \frac{1}{\delta_m} \sum_{k=0}^{\delta_m^{-1}[t]_m^- - 1} \int_{k\delta_m}^{(k+1)\delta_m} (s - k\delta_m)S_{t-k\delta_m}D\sigma_j(Y_m(k\delta_m))\sigma_j(Y_m(k\delta_m))ds$$
$$= \frac{\delta_m}{2} \sum_{k=0}^{\delta_m^{-1}[t]_m^- - 1} S_{t-k\delta_m}D\sigma_j(Y_m(k\delta_m))\sigma_j(Y_m(k\delta_m))$$
$$= \frac{1}{2} \sum_{k=0}^{\delta_m^{-1}[t]_m^- - 1} \int_{k\delta_m}^{(k+1)\delta_m} S_{t-k\delta_m}D\sigma_j(Y_m(k\delta_m))\sigma_j(Y_m(k\delta_m))ds$$
$$= \frac{1}{2} \int_0^{[t]_m^-} S_{t-[s]_m^-}D\sigma_j(Y_m([s]_m^-))\sigma_j(Y_m([s]_m^-))ds.$$

Therefore, we have

$$I_{m,4}^j(t) = \frac{1}{2} \int_0^{[t]_m^-} \left(S_{t-[s]_m^-} - S_{t-s}\right)D\sigma_j(Y_m([s]_m^-))\sigma_j(Y_m([s]_m^-))ds.$$

Consequently, by Lemma 2.12 and Assumption 1.2, we obtain

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|I_{m,4}^j(t)\|^{2p} \right] \leq C\delta_m^{2p},$$

completing the proof. $\qquad \square$

4.11. **Proposition.** *There is a constant $C > 0$ such that for each $m \in \mathbb{N}$ and each $v \in [0,T]$ we have*

$$\mathbb{E}\left[\sup_{t\in[0,v]} \|\zeta_{m,5}(t)\|^{2p} \right] \leq C\left(\int_0^v \mathbb{E}\left[\sup_{t\in[0,u]} \|\xi_m(t) - Y_m(t)\|^{2p} \right]du + \delta_m^{p-1} \right).$$

Proof. Since b is Lipschitz continuous, we obtain

$$\mathbb{E}\left[\sup_{t\in[0,v]}\|\zeta_{m,5}(t)\|^{2p}\right] = \mathbb{E}\left[\sup_{t\in[0,v]}\left\|\int_0^{[t]_m^-}S_{t-s}\big(b(\xi_m(s)) - b(Y_m([s]_m^-))\big)ds\right\|^{2p}\right]$$

$$\leq C\int_0^v \mathbb{E}[\|\xi_m(s) - Y_m([s]_m^-)\|^{2p}]ds \leq C\big(I_1(v) + I_2(v) + I_3(v)\big),$$

where

$$I_1(v) = \int_0^v \mathbb{E}\left[\sup_{t\in[0,u]}\|\xi_m(t) - Y_m(t)\|^{2p}\right]du,$$

$$I_2(v) = \int_0^v \mathbb{E}\left[\sup_{t\in[0,u]}\|Y_m(t) - \bar{Y}_m(t)\|^{2p}\right]du,$$

$$I_3(v) = \int_0^v \mathbb{E}\left[\sup_{t\in[0,u]}\|\bar{Y}_m(t) - Y_m([t]_m^-)\|^{2p}\right]du.$$

Therefore, applying Proposition 3.2 and Lemma 3.4 completes the proof. $\square$

Now, the proof of Theorem 4.1 is an immediate consequence of the decompositions (4.1), (4.4), Propositions 3.2, 4.2, Propositions 4.3–4.11 and Gronwall's inequality.

5. AN EXAMPLE: THE HJMM EQUATION

As an example of our main result, let us consider the HJMM (Heath-Jarrow-Morton-Musiela) equation from mathematical finance. This is a SPDE which models the term structure of interest rates in a market of zero coupon bonds.

Let us briefly introduce the model. A zero coupon bond with maturity T is a financial asset that pays the holder one monetary unit at T. Its price at $t \leq T$ can be written as the continuous discounting of one unit of the domestic currency

$$P(t,T) = \exp\left(-\int_t^T f(t,s)ds\right),$$

where $f(t,T)$ is the rate prevailing at time t for instantaneous borrowing at time T, also called the forward rate for date T.

After transforming the original HJM (Heath-Jarrow-Morton) dynamics of the forward rates (see [21]) by means of the Musiela parametrization $r_t(x) = f(t,t+x)$ (see [6]), the instantaneous forward rate $r_t(x)$ with maturity time x from observing time t can be considered as a mild solution to the HJMM (Heath-Jarrow-Morton-Musiela) equation

$$(5.1) \quad \begin{cases} dr(t) &= \big(\frac{\partial}{\partial x}r(t) + \alpha(r(t),v(t))\big)dt + \sum_{j=1}^r \gamma_j(r(t),v(t))dB^j(t) \\ v(t) &= \mu(v(t))dt + \sum_{j=1}^r \lambda_j(v(t))dB^j(t) \\ r(0) &= r_0 \\ v(0) &= v_0. \end{cases}$$

Note that we consider the HJMM equation (5.1) with stochastic volatility. More precisely, the functions $(\gamma_j)_{j=1,\ldots,r}$ are called volatility functions; they represent the degree of variation of the instantaneous forward rate. If $v(t) = t$, $t \in \mathbb{R}_+$, which corresponds to $\mu \equiv 1$ and $\lambda_j \equiv 0$ for all $j = 1,\ldots,r$, then the model is called a local volatility model. In this case, the volatility of the instantaneous forward rate with each maturity depends on the observation time and the interest rate curve itself. If $v(t)$ is not deterministic, then the model is called a stochastic volatility model, which fits more with the real interest rate market.

In order to ensure absence of arbitrage in the bond market, we consider the HJMM equation (5.1) under a martingale measure. Then the drift term is given by the so-called HJM drift condition

$$(5.2) \qquad \alpha(h, v) = \sum_{j=1}^{r} \gamma_j(h, v) \int_0^\bullet \gamma_j(h, v)(\eta) d\eta.$$

We refer, e.g., to [11] for further details concerning the derivation of the HJMM equation (5.1) and the HJM drift condition (5.2).

The precise mathematical formulation of our model is as follows. We fix an arbitrary constant $\beta > 0$. Let $\tilde{H}_\beta$ be the space of all absolutely continuous functions $h : \mathbb{R}_+ \to \mathbb{R}$ such that

$$(5.3) \qquad \|h\|_\beta := \left(|h(0)|^2 + \int_{\mathbb{R}_+} |h'(x)|^2 e^{\beta x} dx \right)^{\frac{1}{2}} < \infty.$$

This kind of space was introduced in [11, Sec. 5.1], where the following properties have been proven:

- The space $(\tilde{H}_\beta, \|\cdot\|_\beta)$ is a separable Hilbert space.
- For each $x \in \mathbb{R}_+$ the point evaluation $h \mapsto h(x) : \tilde{H}_\beta \to \mathbb{R}$ is a continuous linear functional.
- The semigroup $(\tilde{S}_t)_{t \geq 0}$ of right shifts in $\tilde{H}_\beta$ defined by $\tilde{S}_t h(x) = h(x+t)$ is a C_0-semigroup with infinitesimal generator $\tilde{A}$ given by $\tilde{A}h(x) = h'(x)$ on the domain $\mathcal{D}_\beta(\tilde{A}) = \{h \in \tilde{H}_\beta : h' \in \tilde{H}_\beta\}$.
- For each $h \in \tilde{H}_\beta$ the limit $h(\infty) := \lim_{x \to \infty} h(x)$ exists, and

$$\tilde{H}_\beta^0 := \{h \in \tilde{H}_\beta : h(\infty) = 0\}$$

  is a closed subspace of $\tilde{H}_\beta$.

Let us fix an additional index $\beta' > \beta$. We have the following additional result.

5.1. Lemma. *The following statements are true:*

(1) *The multiplication operator $m : \tilde{H}_\beta \times \tilde{H}_\beta \to \tilde{H}_\beta$ given by $m(h, g) = hg$ is a continuous bilinear operator.*

(2) *We have $m(\mathcal{D}_\beta(\tilde{A}) \times \mathcal{D}_\beta(\tilde{A})) \subset \mathcal{D}_\beta(\tilde{A})$, and the restriction of m is a continuous bilinear operator with respect to the graph norm.*

(3) *We have $\tilde{H}_{\beta'} \subset \tilde{H}_\beta$ with continuous embedding.*

(4) *The integral operator $\mathcal{I} : \tilde{H}_{\beta'}^0 \to \tilde{H}_\beta$ given by $\mathcal{I}h := \int_0^\bullet h(\eta) d\eta$ is a continuous linear operator.*

(5) *We have $\mathcal{I}(\mathcal{D}_{\beta'}(\tilde{A}) \cap \tilde{H}_{\beta'}^0) \subset \mathcal{D}_\beta(\tilde{A})$, and the restriction is a continuous linear operator with respect to the corresponding graph norms.*

Proof. This is a consequence of [28, Thm. 4.1 and Lemmas 4.2, 4.3]. $\qquad \square$

Now, let us assume the following.

5.2. Assumption. *We suppose that the following conditions are satisfied:*

- *We have $\gamma_j \in C_b^2(\tilde{H}_\beta \times \mathbb{R}; \tilde{H}_{\beta'}^0)$ for each $j = 1, \ldots, r$.*
- *We have $\gamma_j(\mathcal{D}_\beta(\tilde{A}) \times \mathbb{R}) \subset \mathcal{D}_{\beta'}(\tilde{A})$ for each $j = 1, \ldots, r$.*
- *We have $\gamma_j|_{\mathcal{D}_\beta(\tilde{A}) \times \mathbb{R}} \in C_b^2(\mathcal{D}_\beta(\tilde{A}) \times \mathbb{R}; \mathcal{D}_{\beta'}(\tilde{A}))$ for each $j = 1, \ldots, r$ with respect to the corresponding graph norms.*
- *$\mu : \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous and bounded.*
- *We have $\lambda_j \in C_b^2(\mathbb{R})$ for each $j = 1, \ldots, r$.*

Now we can consider the HJM model in our SPDE framework, and rewrite the HJMM equation (5.1) as

$$(5.4) \qquad \begin{cases} dX(t) & = \; \big(AX(t) + b(X(t))\big)dt + \sum_{j=1}^{r} \sigma_j(X(t))dB^j(t) \\ X(0) & = \; x_0 \end{cases}$$

on the separable Hilbert space $H = \tilde{H}_\beta \times \mathbb{R}$ with the notation

$$x_0 := \begin{pmatrix} r_0 \\ v_0 \end{pmatrix} \in \mathcal{D}(A) := \mathcal{D}_\beta(\tilde{A}) \times \mathbb{R},$$

$$X(t) := \begin{pmatrix} r(t) \\ v(t) \end{pmatrix} \in H,$$

$$S_t x := \begin{pmatrix} \tilde{S}_t h \\ v \end{pmatrix} \text{ for } x = \begin{pmatrix} h \\ v \end{pmatrix} \in H,$$

$$Ax := \begin{pmatrix} \tilde{A}h \\ 0 \end{pmatrix} \text{ for } x = \begin{pmatrix} h \\ v \end{pmatrix} \in \mathcal{D}(A) := \mathcal{D}_\beta(\tilde{A}) \times \mathbb{R},$$

$$b := \begin{pmatrix} b^1 \\ b^2 \end{pmatrix} : H \to H, \quad b^1(h,v) = \alpha(h,v), \quad b^2(h,v) = \mu(v),$$

$$\sigma_j := \begin{pmatrix} \sigma_j^1 \\ \sigma_j^2 \end{pmatrix} : H \to H, \quad \sigma_j^1(h,v) = \gamma_j(h,v), \quad \sigma_j^2(h,v) = \lambda_j(v).$$

Note that the HJM drift term (5.2) has the representation

$$\alpha = \sum_{j=1}^{r} m(\gamma_j, \mathcal{I}\gamma_j).$$

Taking into account the Leibniz rule (see, for example [1, Thm. 2.4.4]), by virtue of Lemma 5.1 and Assumption 5.2 we obtain that Assumptions 1.1 and 1.2 are fulfilled. Consequently, Theorem 1.3 applies to the HJMM equation (5.4) and provides the convergence rate for the corresponding Wong-Zakai approximations. By Proposition 2.9 and identity (2.1) the Wong-Zakai approximations $\xi_m = (\varrho_m, \zeta_m)$ for $m \in \mathbb{N}$ are the solutions to the integral equations

$$\varrho_m(t,x) = \varrho_m(k\delta_m, x + t - k\delta_m) + \int_{k\delta_m}^{t} \Big(\alpha(\varrho_m(t), \zeta_m(s), x + t - s)$$
$$- \beta(\varrho_m(s), \zeta_m(s), x + t - s)$$
$$+ \sum_{j=1}^{r} \frac{B^j((k+1)\delta_m) - B^j(k\delta_m)}{\delta_m} \gamma_j(\varrho_m(s), \zeta_m(s), x + t - s) \Big) ds,$$

$$\zeta_m(t) = \zeta_m(k\delta_m) + \int_{k\delta_m}^{t} \Big(\mu(\zeta_m(s)) - \frac{1}{2} \sum_{j=1}^{r} \lambda_j(\zeta_m(s))\lambda_j'(\zeta_m(s))$$
$$+ \sum_{j=1}^{r} \frac{B^j((k+1)\delta_m) - B^j(k\delta_m)}{\delta_m} \lambda_j(\zeta_m(s)) \Big) ds$$

for $t \in [k\delta_m, (k+1)\delta_m]$ and $k = 0, \ldots, m-1$. Note that $x \in \mathbb{R}_+$ in the second argument of $\varrho_m(t,x)$ denotes the point evaluation of the interest rate curve. Moreover, we have used the notation

$$\beta(r,v) := \frac{1}{2} \sum_{j=1}^{r} \big(D_r \gamma_j(r,v)\gamma_j(r,v) + D_v \gamma_j(r,v)\lambda_j(v) \big).$$

6. FURTHER EXAMPLES

In this section, we treat two further examples arising from natural sciences. Before presenting these examples, let us recall an auxiliary result for the infinitesimal generators of strongly continuous semigroups. As in the previous sections, let $A : \mathcal{D}(A) \subset H \to H$ be the infinitesimal generator of a C_0-semigroup $(S_t)_{t \geq 0}$ on the separable Hilbert space H.

6.1. Lemma. [25, Thm. 3.1.1] *Let $R \in L(H)$ be a continuous linear operator, and let the linear operator $B : \mathcal{D}(B) \subset H \to H$ be given by $\mathcal{D}(B) := \mathcal{D}(A)$ and $B := A + R$. Then B is the generator of a C_0-semigroup on H.*

As our first example of this section, we consider the stochastic quantization of the free Euclidean quantum field (cf. [26, Ex. 1.0.1])

$$
(6.1) \qquad \begin{cases} dX(t) & = (\Delta - m^2)X(t)dt + \sum_{j=1}^r \sigma_j dB^j(t) \\ X(0) & = x_0, \end{cases}
$$

where $m \in \mathbb{R}_+$ denotes "mass", and the volatilities $\sigma_1, \ldots, \sigma_r \in \mathcal{D}(\Delta)$ are constant. Here we choose the state space $H = L^2(\mathbb{R}^d)$, and the Laplace operator Δ is defined on the domain $\mathcal{D}(\Delta) = W^2(\mathbb{R}^d)$. Taking into account Lemma 6.1, we see that Assumptions 1.1 and 1.2 are fulfilled, and hence Theorem 1.3 applies to the SPDE (6.1) and provides the convergence rate for the corresponding Wong-Zakai approximations.

Another example is the stochastic cable equation (cf. [10, Ex. 0.8])

$$
(6.2) \qquad \begin{cases} dV(t) & = \frac{1}{\tau}(\lambda^2 \Delta V(t) - V(t))dt + \sum_{j=1}^r \sigma_j dB^j(t) \\ V(0) & = x_0, \end{cases}
$$

where $\lambda > 0$ denotes the length constant, $\tau > 0$ denotes the time constant of the electric cable, and the volatilities $\sigma_1, \ldots, \sigma_r \in \mathcal{D}(\Delta)$ are constant. Here we choose the state space $H = L^2((0, \pi))$, and the Laplace operator Δ is defined on the domain $\mathcal{D}(\Delta) = W^2((0, \pi)) \cap H_0^1((0, \pi))$. Taking into account Lemma 6.1, we see that Assumptions 1.1 and 1.2 are fulfilled, and hence Theorem 1.3 applies to the SPDE (6.2) and provides the convergence rate for the corresponding Wong-Zakai approximations.

ACKNOWLEDGEMENT

REFERENCES

[1] R. Abraham, J. E. Marsden, and T. Ratiu. *Manifolds, tensor analysis and applications.* Springer, New York, 1988.

[2] P. Acquistapace and B. Terreni. An approach to Itô linear equations in Hilbert spaces by approximation of white noise with coloured noise. *Stoch. Anal. Appl.*, 2(2):131–186, 1984.

[3] S. Aida. Support theorem for diffusion processes on Hilbert spaces. *Publ. Res. Inst. Math. Sci.*, 26(6):947–965, 1990.

[4] V. Bally, A. Millet, and M. Sanz-Solé. Approximation and support theorem in Hölder norm for parabolic stochastic partial differential equations. *Ann. Probab.*, 23(1):178–222, 1995.

[5] C. Bayer and J. Teichmann. Cubature on Wiener space in infinite dimension. *Proc. R. Soc. of Lond. Ser. A Math. Phys. Eng. Sci.*, 464(2097):2493–2516, 2008.

[6] A. Brace and M. Musiela. A multifactor Gauss Markov implementation of Heath, Jarrow, and Morton. *Math. Finance*, 4(3):259–283, 1994.

[7] Z. Brzeźniak and A. Carroll. Approximations of the Wong-Zakai type for stochastic differential equations in M-type 2 Banach spaces with applications to loop spaces. In *Séminaire de Probabilités XXXVII*, number 1832 in Lecture Notes in Mathematics, pages 251–289. Springer, Berlin, 2003.

[8] Z. Brzeźniak and F. Flandoli. Almost sure approximation of Wong-Zakai type for stochastic partial differential equations. *Stoch. Process. Appl.*, 55(2):329–358, 1995.

[9] G. Da Prato, A. Jentzen, and M. Röckner. A mild Itô formula for SPDEs. *Trans. Am. Math. Soc.*, 2017. to appear.

[10] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*. Cambridge University Press, New York, second edition, 2014.

[11] D. Filipović. *Consistency problems for Heath-Jarrow-Morton interest rate models*. Springer, Berlin, 2001.

[12] P. Friz and H. Oberhauser. Rough path limits of the Wong-Zakai type with a modified drift term. *J. Funct. Anal.*, 256(10):3236–3256, 2009.

[13] A. Ganguly. Wong-Zakai type convergence in infinite dimensions. *Electron. J. Probab.*, 18(31):1–34, 2013.

[14] L. Gawarecki and V. Mandrekar. *Stochastic differential equations in infinite dimensions with applications to SPDEs*. Springer, Berlin, 2011.

[15] I. Gyöngy, D. Nualart, and M. Sanz-Solé. Approximation and support theorems in modulus spaces. *Probab. Theory Relat. Fields*, 101(4):495–509, 1995.

[16] I. Gyöngy and T. Pröhle. On the approximation of stochastic differential equation and on Stroock-Varadhan's support theorem. *Comput. Math. Appl.*, 19(1):65–70, 1990.

[17] I. Gyöngy and A. Shmatkov. Rate of convergence of Wong-Zakai approximations for stochastic partial differential equations. *Appl. Math. Optim.*, 54(3):315–341, 2006.

[18] I. Gyöngy and P. R. Stinga. Rate of convergence of Wong-Zakai approximations for stochastic partial differential equations. In *Seminar on Stochastic Analysis, Random Fields and Applications VII*, number 67, pages 95–130. Springer, Basel, 2013.

[19] M. Hairer and E. Pardoux. A Wong-Zakai theorem for stochastic PDEs. *J. Math. Soc. Jpn.*, 67(4):1551–1604, 2015.

[20] E. Hausenblas. Wong-Zakai type approximation of SPDEs of Lévy noise. *Acta Appl. Math.*, 98(2):99–134, 2007.

[21] D. Heath, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica*, 60(1):77–105, 1992.

[22] A. Millet and M. Sanz-Solé. The support of the solution to a hyperbolic SPDE. *Probab. Theory Relat. Fields*, 98(3):361–387, 1994.

[23] T. Nakayama. Support theorem for mild solutions of SDE's in Hilbert space. *J. Math. Sci. Univ. Tokyo*, 11(3):245–311, 2004.

[24] T. Nakayama. Viability Theorem for SPDE's including HJM framework. *J. Math. Sci. Univ. Tokyo*, 11(3):313–324, 2004.

[25] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*. Springer, New York, 1983.

[26] C. Prévôt and M. Röckner. *A concise course on stochastic partial differential equations*. Springer, Berlin, 2007.

[27] C. Proppe. The Wong-Zakai theorem for dynamical systems with parametric Poisson white noise excitation. *Int. J. Eng. Sci.*, 40(10):1165–1178, 2002.

[28] S. Tappe. An alternative approach on the existence of affine realizations for HJM term structure models. *Proc. R. Soc. of Lond. Ser. A Math. Phys. Eng. Sci.*, 466(2122):3033–3060, 2010.

[29] G. Tessitore and J. Zabczyk. Wong-Zakai approximations of stochastic evolution equations. *J. Evol. Equ.*, 6(4):621–655, 2006.

[30] C. Tudor. Wong-Zakai type approximations for stochastic differential equations driven by a fractional Brownian motion. *Z. Anal. ihre Anwend.*, 28(2):165–182, 2009.

[31] K. Twardowska. An extension of the Wong-Zakai theorem for stochastic evolution equations in Hilbert spaces. *Stoch. Anal. Appl.*, 10(4):471–500, 1992.

[32] K. Twardowska. *Approximation theorems of Wong-Zakai type for stochastic differential equations in infinite dimensions*. PhD thesis, Warsaw Technical University, Institute of Mathematics, 1993.

[33] K. Twardowska. An approximation theorem of Wong-Zakai type for nonlinear stochastic partial differential equations. *Stoch. Anal. Appl.*, 13(5):601–626, 1995.

[34] K. Twardowska. An approximation theorem of Wong-Zakai type for stochastic Navier-Stokes equations. *Rend. Sem. Mat. Univ. Padova*, 96:15–36, 1996.

[35] K. Twardowska. Wong-Zakai approximations for stochastic differential equations. *Acta Appl. Math.*, 43(3):317–359, 1996.

[36] S. Wilks. *Mathematical Statistics*. Wiley, New York, 1962.

[37] E. Wong and M. Zakai. On the convergence of ordinary integrals to stochastic integrals. *Ann. Math. Stat.*, 36(5):1560–1564, 1965.

[38] E. Wong and M. Zakai. On the relationship between ordinary and stochastic differential equations. *Int. J. Eng. Sci.*, 3(2):213–229, 1965.

MUFG Bank, Ltd., Otemachi Financial City Grand Cube 20F 9-2, Otemachi 1-chome, Chiyoda-ku, Tokyo 100-0004, Japan

E-mail address: tnkym376736@ch.em-net.ne.jp

Albert Ludwig University of Freiburg, Department of Mathematical Stochastics, Ernst-Zermelo-Strasse 1, D-79104 Freiburg, Germany

E-mail address: stefan.tappe@math.uni-freiburg.de