An exponential integrator scheme for time discretization of nonlinear stochastic wave equation

Xiaojie Wang

Received: date / Accepted: date

Abstract This work is devoted to convergence analysis of an exponential integrator scheme for semi-discretization in time of nonlinear stochastic wave equation. A unified framework is first set forth, which covers important cases of additive and multiplicative noises. Within this framework, the proposed scheme is shown to converge uniformly in the strong $L^p$-sense with precise convergence rates given. The abstract results are then applied to several concrete examples. Further, weak convergence rates of the scheme are examined for the case of additive noise. To analyze the weak error for the nonlinear case, techniques based on the Malliavin calculus were usually exploited in the literature. Under certain appropriate assumptions on the nonlinearity, this paper provides a weak error analysis, which does not rely on the Malliavin calculus. The rates of weak convergence can, as expected, be improved in comparison with the strong rates. Both strong and weak convergence results obtained here show that the proposed method achieves higher convergence rates than the implicit Euler and Crank-Nicolson time discretizations. Numerical results are finally reported to confirm our theoretical findings.

Keywords nonlinear stochastic wave equation · multiplicative noise · exponential Euler scheme · strong convergence · weak convergence

Mathematics Subject Classification (2000) 60H35 · 60H15 · 65C30

1 Introduction

In the present paper, we consider stochastic evolution equation of Itô type in a separable Hilbert space $(U, \langle \cdot, \cdot \rangle_U, \| \cdot \|_U)$, given by

$$
\begin{cases}
    d\dot{u}(t) = -\Lambda u(t) \, dt + F(u(t)) \, dt + G(u(t)) \, dW(t), & t \in [0, T], \\
    u(0) = u_0, \dot{u}(0) = v_0,
\end{cases}
$$

(1.1)

where $T \in (0, \infty)$ and $\Lambda: \mathcal{D}(\Lambda) \subseteq U \to U$ is a densely defined, linear unbounded, positive self-adjoint operator with compact inverse (e.g., $\Lambda = -\Delta$ with homogeneous Dirichlet boundary condition). Under this assumption, there exists an increasing sequence of real numbers $\{\lambda_i\}_{i=1}^{\infty}$ and an orthonormal basis $\{e_i\}_{i=1}^{\infty}$ such that $\Lambda e_i = \lambda_i e_i$ and

$$0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n (\to \infty).$$

(1.2)

In Equation (1.1), $\{u(t)\}_{t \in [0, T]}$ is regarded as a $U$-valued stochastic process and $\{\dot{u}(t)\}_{t \in [0, T]}$ stands for the time derivative of $u$. Moreover, $F: U \to U$, $G(u): U \to U$ for $u \in U$ are deterministic mappings and

Xiaojie Wang (Corresponding author)
School of Mathematics and Statistics and School of Geosciences and Info-Physics, Central South University, Changsha 410083, Hunan, PR China
E-mail: x.j.wang7@csu.edu.cn, x.j.wang7@gmail.com
\[ \{W(t)\}_{t \in [0,T]} \] is a cylindrical \( Q \)-Wiener process on a given probability space \( (\Omega, \mathcal{F}, \mathbb{P}) \) with normal filtration \( \{\mathcal{F}_t\}_{t \in [0,T]} \).

The abstract equation (1.1) includes many stochastic wave equations (SWEs) in applications [5,11,35]. Since their true solutions are rarely known explicitly, numerical simulations are often used to understand the behavior of the solutions. To do this, one has to discretize both the time interval \([0,T]\) and the infinite dimensional space \( U \). As pointed out in [32], the main difficulty in studying numerical schemes for stochastic partial differential equations (SPDEs) of evolutionary type lies in the treatment of the time discretization. On the one hand, development of effective high order time-stepping schemes are important since better time integration technology gives significant performance improvements in the numerical solution of SPDEs. This issue is, however, essentially difficult, even in the context of finite dimensional stochastic ordinary differential equations (SODEs) [21]. Time discretizations of SPDEs encounter all the difficulties that arise in the time approximations of both deterministic PDEs and finite dimensional SODEs as well as many more due to the infinite dimensional nature of the driving noise processes. On the other hand, as one can see in our forthcoming work, the arguments used in the following analysis of pure time discretization can be extended to the analysis of fully discretized scheme when combined with arguments used in the deterministic theory. For instance, optimal error estimates for the deterministic wave equation obtained in [24] allow us to analyze a fully discretized scheme with finite element spatial discretization. Therefore, and also for the sake of simplicity, we first concentrate on the time discretization of (1.1) in this paper.

On the interval \([0,T]\), we construct a uniform mesh \( T_M = \{0, t_1, \cdots, t_M\} \) satisfying
\[
T_M : 0 = t_0 < t_1 < t_2 < \cdots < t_M = T
\] (1.3)
with \( \tau = T/M, M \in \mathbb{N} \), being the time stepsize. This article is concerned with the following scheme for the time discretization of (1.1):
\[
\begin{aligned}
&u_{m+1} = C(\tau)u_m + \Lambda^{-\frac{1}{2}}S(\tau)v_m + \tau \Lambda^{-\frac{1}{2}}S(\tau)F(u_m) + \Lambda^{-\frac{1}{2}}S(\tau)G(u_m)\Delta W_m, \\
v_{m+1} = -\Lambda^{\frac{1}{2}}S(\tau)u_m + C(\tau)v_m + \tau C(\tau)F(u_m) + C(\tau)G(u_m)\Delta W_m,
\end{aligned}
\] (1.4)
where \( C(t) = \cos(t\Lambda^{\frac{1}{2}}) \) and \( S(t) = \sin(t\Lambda^{\frac{1}{2}}) \) for \( t \in [0,T] \) are the cosine and sine operators and \( \Delta W_m = W(t_{m+1}) - W(t_m) \) is the Wiener increment. Here \( u_m \) and \( v_m \) are, respectively, temporal approximations of \( u(t) \) and \( \dot{u}(t) \) at the grid points \( t_m \in T_M \). Rewriting (1.4) in another abstract form (see (2.13)), one can observe that the scheme (1.4) can be identified with a version of the stochastic exponential Euler scheme in the literature [16,19,20,27,28], where the schemes were designed for parabolic SPDEs. By choosing particular filter functions, the stochastic exponential Euler scheme can be viewed as a stochastic trigonometric method for nonlinear problems [7, Example 2]. Recently, Cohen, Larsson and Sigg [8] applied the stochastic trigonometric method to the linear SWE with additive noise, which coincides with the scheme (1.4) when applied to the linear case. In the first part of this article, we establish the uniform \( L^p \)-convergence of the scheme (1.4) in a general framework for nonlinear problems, which of course extends the \( L^2 \)-convergence results in [3] for the linear case.

To get started, we make the following assumptions.

**Assumption 1.1** Assume \( F : U \to U \) and \( G(u) : U \to U \) for \( u \in U \) are deterministic mappings satisfying
\[
\begin{aligned}
&\|F(u)\|_U + \left\| \Lambda^{-\frac{1}{2}}G(u) \right\|_{L_{\infty}^2} \leq L(\|u\|_U + 1),
\end{aligned}
\] (1.5)
\[
\begin{aligned}
&\|F(u_1) - F(u_2)\|_U + \left\| \Lambda^{\frac{1}{2}}(G(u_1) - G(u_2)) \right\|_{L_{\infty}^2} \leq L\|u_1 - u_2\|_U
\end{aligned}
\] (1.6)
for all \( u, u_1, u_2 \in U \) and some positive constants \( \delta, L \in (0, \infty) \).

For notations, we refer to the next section. Throughout this article, we choose \( L \geq 1 \) to be a universal constant in all assumptions. Under Assumption 1.1 (1.1) possesses a unique mild solution (Theorem 2.1).

Moreover, the spatial regularity properties of the mild solution and its numerical approximations are investigated under the above assumption (Proposition 3.1). Further, we establish the main strong convergence
result, Theorem 5.1 which shows that the scheme (1.4) is convergent uniformly in the strong $L^p$-sense, $p \in [2, \infty)$. In particular, we show
\[
\sup_{t_m \in \mathcal{M}} \left\| u(t_m) - u_m \right\|_{L^p(\Omega, \mathbb{R})} \leq C \tau^{\min(\delta, 1)}.
\] (1.7)

Here, and throughout this work, $C$ is a constant that may vary from one place to another and depends on $T, \delta, L, p, \lambda$, and the initial data $u_0, \psi_0$, but is independent of $M$. To derive the uniform bound (1.7), a continuous-time extension of $u_m$ is introduced and the group property of the operator family $E(t) = e^{tA}$ is exploited. Then, several concrete examples including SWEs with general additive noise, multiplicative trace class noise and multiplicative space-time white noise are presented to fit in the above setting. For the three different cases, (1.7) is then applied with precise convergence rates given. It is shown that the spatial regularity of the noise term determines the order of strong convergence. Moreover, unlike the parabolic case, the exponential Euler scheme achieves higher strong order than the implicit Euler and Crank-Nicolson time discretizations. Instead of imposing very strong spatial regularity conditions on the nonlinearities [12, 23], whose solution can be written down explicitly, one can get rid of the irregular term involving the unbounded operator by a transformation of variables. This transformation, however, does not work for the nonlinear heat equation. Since the operator family $E(t) = e^{tA}$ is a group (see the next section), one can adapt the basic line of [12] to make the weak error analysis easier. Instead of imposing very strong spatial regularity conditions on the nonlinearities [12, Hypothesis 5.7, 5.8], we make certain relatively mild assumptions on the nonlinearity $F$ (c.f. Theorem 5.1 and Corollary 5.2). For instance, the assumptions do not require $F$ to be twice Fréchet differentiable in $U$, which is in general not fulfilled for Nemitskij operators but was commonly demanded in the weak error analysis [11, 12, 14, 37]. Moreover, a key condition on $F$ (see (5.3)) is used to avoid techniques involving the Malliavin calculus. Under such assumptions we obtain
\[
\left| \mathbb{E}[\phi(u(T))] - \mathbb{E}[\phi(u_M)] \right| \leq C \tau^{\min(2\beta, 1+\beta-\varepsilon, 1)}
\] (1.9)

with arbitrarily small $\varepsilon > 0$. As expected, the rate of weak convergence is, up to an arbitrarily small $\varepsilon > 0$ in some cases, twice that of strong convergence. Eventually, the assumptions we put are illustrated by a concrete setting.

In summary, the contribution of this article to the numerical analysis of SWEs is twofold. On the one hand, a unified abstract framework is formulated and a high order time-stepping scheme is proposed for nonlinear SWEs. Within the abstract framework, strong convergence of the exponential Euler scheme is studied thoroughly. The framework is general as it covers cases of additive noise, multiplicative trace-class noise as well as multiplicative space-time white noise. On the other hand, we provide a weak error analysis which does not rely on the Malliavin calculus, while techniques based on the Malliavin calculus were usually exploited in the course of weak error analysis for the nonlinear case [11, 12, 14, 37]. Both strong and weak convergence results show that the proposed scheme attains higher convergence rates than the implicit Euler and Crank-Nicolson time discretizations.

Finally, we mention that parabolic SPDEs have been extensively studied from the numerical point of view. The readers are referred to a review article [17], a relevant monograph [18] and more recent doctor
2 Preliminaries and framework

Let \((U, \langle \cdot, \cdot \rangle_U, \| \cdot \|_U)\) and \((H, \langle \cdot, \cdot \rangle_H, \| \cdot \|_H)\) be two separable Hilbert spaces. By \(\mathcal{L}(U, H)\) we denote the space of bounded linear operators from \(U\) to \(H\) with the usual operator norm \(\| \cdot \|_{\mathcal{L}(U, H)}\) and write \(\mathcal{L}(U) = \mathcal{L}(U, U)\) to lighten the notation. Alternatively, we need spaces of nuclear and Hilbert-Schmidt operators \([9, 31]\). The space of nuclear operators from \(U\) to \(H\) is denoted by \(\mathcal{L}_1(U, H)\) and we write \(\mathcal{L}_1(U) = \mathcal{L}_1(U, U)\). If \(\Gamma \in \mathcal{L}_1(U)\) is nonnegative and symmetric, then

\[
\| \Gamma \|_{\mathcal{L}_1(U)} = \text{Tr}(\Gamma) := \sum_{i=1}^{\infty} \langle \Gamma \psi_i, \psi_i \rangle_U,
\]

where \(\{\psi_i\}_{i \in \mathbb{N}}\) is an orthonormal basis of \(U\) and the trace of a nuclear operator, namely, \(\text{Tr}(\Gamma)\) for \(\Gamma \in \mathcal{L}_1(U)\), is independent of the particular choice of the basis \(\{\psi_i\}_{i \in \mathbb{N}}\). By \(\mathbb{N} := \{1, 2, 3, \ldots\}\) we denote the natural numbers and by \(\mathcal{L}_2(U, H)\) we denote the space of Hilbert-Schmidt operators from \(U\) to \(H\), equipped with the norm

\[
\| \Gamma \|_{\mathcal{L}_2(U, H)} = \left( \sum_{i=1}^{\infty} \| \Gamma \psi_i \|_H^2 \right)^{1/2},
\]

also not depending on the particular choice of the basis. Analogously, we write \(\mathcal{L}_2(U) = \mathcal{L}_2(U, U)\). If \(\Gamma_1 \in \mathcal{L}_2(U, H)\) and \(\Gamma_2 \in \mathcal{L}_2(U, H)\), \(j = 1, 2\), then \(\Gamma_1 \Gamma_2 \in \mathcal{L}_2(U, H)\) for \(j = 1, 2\), and

\[
\| \Gamma_1 \Gamma_2 \|_{\mathcal{L}_2(U, H)} \leq \| \Gamma_1 \|_{\mathcal{L}_2(U, H)} \cdot \| \Gamma_2 \|_{\mathcal{L}_2(U, H)}, \quad j = 1, 2.
\]

Moreover, if \(\Gamma_1 \in \mathcal{L}_2(H)\) and \(\Gamma_2 \in \mathcal{L}_2(H)\), then both \(\Gamma_1 \Gamma_2\) and \(\Gamma_2 \Gamma_1\) belong to \(\mathcal{L}_2(H)\) and

\[
\text{Tr}(\Gamma_1 \Gamma_2) = \text{Tr}(\Gamma_2 \Gamma_1).
\]

Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a probability space with a normal filtration \(\{\mathcal{F}_t\}_{0 \leq t \leq T}\) and by \(L^p(\Omega, U)\) we denote the space of \(U\)-valued integrable random variables with the norm defined by \(\|X\|_{L^p(\Omega, U)} = \left( \mathbb{E} [\|X \|_p^p] \right)^{1/p}\) for \(p \in [2, \infty)\). Furthermore, the driven stochastic process \(W(t)\) in (2.1) is assumed to be a cylindrical \(Q\)-Wiener process in the stochastic basis \((\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{0 \leq t \leq T})\), with a covariance operator \(Q : U \to U\), which can be represented as follows \([9, 31]\):

\[
W(t) = \sum_{i=1}^{\infty} \sqrt{q_i} \beta_i(t) \phi_i, \quad t \in [0, T],
\]

where \(\{\beta_i(t)\}_{i \in \mathbb{N}}\) is a family of mutually independent real Brownian motions and \(\{\phi_i\}_{i \in \mathbb{N}}\) form an orthonormal basis of \(U\) consisting of eigenfunctions of \(Q\) with \(Q \phi_i = q_i \phi_i, q_i \geq 0, i \in \mathbb{N}\). The covariance operator \(Q \in \mathcal{L}(U)\) is nonnegative and symmetric, but not necessarily of finite trace. As a result, the series in (2.5) may not converge in \(U\), but in some space \(U_1\) into which \(U\) can be embedded, see, e.g., \([9, 31]\). Let \(Q^{\frac{1}{2}}\) denote the unique positive square root of \(Q\). Now we are able to introduce the separable Hilbert space \(U_0 := Q^{\frac{1}{2}}(U)\) endowed with the inner product \((u, \hat{u})_0 = (Q^{-\frac{1}{2}} u, Q^{-\frac{1}{2}} \hat{u})_U\) for \(u, \hat{u} \in U\), where \(Q^{-\frac{1}{2}}\) is the pseudo inverse of \(Q^{\frac{1}{2}}\) in the case when \(Q\) is not one-to-one. For lighter notation, we use \(\mathcal{L}_0^2\) to denote the separable Hilbert spaces \(\mathcal{L}_2(U_0, U)\) and also \(\mathcal{L}_2(U_0, H)\) when it causes no confusion.
At this point, we are ready to discuss the existence and uniqueness of the so-called mild solution of (1.1). To this end, we rewrite (1.1) as a stochastic evolution equation of first order in a product space. Specifically, introducing a new variable $v = u$ transforms (1.1) into the following Cauchy problem
\[
\begin{align*}
\text{d}X(t) &= AX(t)\text{d}t + F(X(t))\text{d}t + G(X(t))\text{d}W(t), & t \in (0,T], \\
X(0) &= X_0,
\end{align*}
\] (2.6)
where we denote
\[
X_0 = \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}, \quad X = \begin{bmatrix} u \\ v \end{bmatrix}, \quad A = \begin{bmatrix} 0 & I \\ -\Lambda & 0 \end{bmatrix}, \quad F(X) = \begin{bmatrix} 0 \\ F(u) \end{bmatrix}, \quad G(X) = \begin{bmatrix} 0 \\ G(u) \end{bmatrix}.
\] (2.7)

Here and below by $I$ we mean the identity operator in $U$. In this way, we transfer the existence and uniqueness of the mild solution of (1.1) to the same problem for (2.6). Before proceeding further, we need additional spaces and notations. We introduce the Hilbert space $H^\gamma = \mathcal{D}(A^\frac{\gamma}{2})$ for $\gamma \in \mathbb{R}$ [23], equipped with the inner product
\[
\langle u, \omega \rangle_{H^\gamma} := \langle A^\frac{\gamma}{2} u, A^\frac{\gamma}{2} \omega \rangle_U = \sum_{i=1}^{\infty} \lambda_i^\gamma \langle u, e_i \rangle_U \langle \omega, e_i \rangle_U, \quad \gamma \in \mathbb{R},
\] and the corresponding norm $\|u\|_{H^\gamma}^2 = \langle u, u \rangle_{H^\gamma}$. Then $H^0 = U$ and $H^\alpha \subset H^\beta$ if $\alpha \geq \beta$. Furthermore, we introduce the product space
\[
H^\gamma := H^\gamma \times H^{-1}, \quad \gamma \in \mathbb{R},
\] (2.9)
endowed with the inner product
\[
\langle X_1, X_2 \rangle_{H^\gamma} = \langle u_1, u_2 \rangle_{H^\gamma} + \langle v_1, v_2 \rangle_{H^{-1}}, \quad X_1 = (u_1, v_1)^T, \quad X_2 = (u_2, v_2)^T
\] (2.10)
and the corresponding norm
\[
\|X\|_{H^\gamma}^2 = \langle X, X \rangle_{H^\gamma} = \|u\|_{H^\gamma}^2 + \|v\|_{H^{-1}}^2, \quad \gamma \in \mathbb{R}, \quad X = (u,v)^T.
\] (2.11)

For the special case $\gamma = 0$, we denote $H := H^0 = H^0 \times H^{-1}$ and $(H, \langle \cdot, \cdot \rangle_H, \|\cdot\|_H)$ is a separable Hilbert space. Throughout this work we regard $\Lambda$ as an operator from $H^1$ to $H^{-1}$, defined by $(\Lambda u)(\varphi) = \langle \nabla u, \nabla \varphi \rangle$ for $u, \varphi \in H^1$, and define
\[
\mathcal{D}(A) = \left\{ X = (u,v)^T \in H : AX = \begin{bmatrix} v \\ -\Lambda u \end{bmatrix} \in H = H^0 \times H^{-1} \right\} = H^1 = H^1 \times H^0.
\] (2.12)

In this setting one can rigorously check that $A$ is closed and densely defined in $H$ and the resolvent set of $A$ contains all non-zero real numbers and $\|\lambda (\mathcal{T} - A)^{-1}\|_{\mathcal{L}(H)} \leq \frac{1}{\lambda}$ for any $\lambda \in \mathbb{R}$ (see, e.g., Section 5.3) for more details). Here and below, by $\mathcal{T}$ we mean the identity operator in $H$. As a consequence, the operator $\Lambda$ is an infinitesimal generator of a $C_0$-group $E(t) = e^{tA}, t \in \mathbb{R}$ on $H$. In order to see the exact form of $E(t)$, note that $X(t) = E(t)X_0$ is the solution of the deterministic linear equation
\[
X = AX; \quad X(0) = X_0 = (u_0, v_0)^T.
\] (2.13)

We solve it by using an eigenfunction expansion:
\[
X(t) = e^{tA}X_0 = \sum_{i=1}^{\infty} \exp \left( \begin{bmatrix} 0 & I \\ -\lambda_i & 0 \end{bmatrix} t \right) \begin{bmatrix} \langle u_0, e_i \rangle_U e_i \\ \langle v_0, e_i \rangle_U e_i \end{bmatrix} = \sum_{i=1}^{\infty} \begin{bmatrix} \cos(\sqrt{\lambda_i} t) & \frac{1}{\sqrt{\lambda_i}} \sin(\sqrt{\lambda_i} t) \\ -\sqrt{\lambda_i} \sin(\sqrt{\lambda_i} t) & \cos(\sqrt{\lambda_i} t) \end{bmatrix} \begin{bmatrix} \langle u_0, e_i \rangle_U e_i \\ \langle v_0, e_i \rangle_U e_i \end{bmatrix}.
\]

Hence the first component of $X = (u,v)^T$ is given by
\[
u(t) = \sum_{i=1}^{\infty} \cos(\sqrt{\lambda_i} t) \langle u_0, e_i \rangle_U e_i + \frac{1}{\sqrt{\lambda_i}} \sin(\sqrt{\lambda_i} t) \langle v_0, e_i \rangle_U e_i,
\] (2.14)
\[
= \cos(tA^\frac{1}{2}) u_0 + A^{-\frac{1}{2}} \sin(tA^\frac{1}{2}) v_0,
\]
and the second component of $X = (u, v)^T$ by

$$v(t) = \dot{u}(t) = - \Lambda^{\frac{1}{2}} \sin(t \Lambda^{\frac{1}{2}}) u_0 + \cos(t \Lambda^{\frac{1}{2}}) v_0. \quad (2.15)$$

Here we introduced the cosine and sine operators $\cos(t \Lambda^{\frac{1}{2}})$ and $\sin(t \Lambda^{\frac{1}{2}}).$ Accordingly, the semigroup $E(t) = e^{t\Lambda}$ should explicitly take the form as

$$E(t) = e^{t\Lambda} = \begin{bmatrix} C(t) & \Lambda^{-\frac{1}{2}} S(t) \\ -\Lambda^{\frac{1}{2}} S(t) & C(t) \end{bmatrix}, \quad (2.16)$$

where we further write $C(t) = \cos(t \Lambda^{\frac{1}{2}})$ and $S(t) = \sin(t \Lambda^{\frac{1}{2}})$ for brevity. Obviously, the cosine and sine operators satisfy a trigonometric identity in the sense that $\|S(t)u\|_E^2 + \|C(t)u\|_E^2 = \|u\|_E^2$ for $u \in U.$ Using the trigonometric identity gives

$$\|E(t)\|_{\mathcal{L}(H)} \leq 1, \quad t \in \mathbb{R}. \quad (2.17)$$

Also, we will frequently use the fact that $S(t), C(t), t \in \mathbb{R}$ and $\Lambda^\gamma, \gamma \in \mathbb{R}$ commute in the following estimates. In the above setting, the scheme (1.4) can be rewritten as a recurrence equation in $H$:

$$X_{m+1} = E(t) \left(X_m + \tau F(X_m) + G(X_m) \Delta W_m\right). \quad (2.18)$$

For the convergence analysis, it is convenient to work with continuous processes. Hence we define a continuous extension of (2.18), $\tilde{X}(t) = (\tilde{u}(t), \tilde{v}(t))^T,$ by

$$\tilde{X}(t) = E(t-t_m) \left(X_m + \int_{t_m}^t X_m + F(X_m)(t-t_m) + G(X_m)(W(t) - W(t_m)) \right) \quad (2.19)$$

for $t \in [t_m, t_{m+1}].$ It is obvious that $\tilde{X}(t)$ coincides with $X_m$ at the grid-points $t_m \in \mathcal{T}_M.$ Note that (2.18) implies

$$X_m = E(t_m)X_0 + \sum_{k=0}^{m-1} E(t_m-t_k)F(X_k) + \sum_{k=0}^{m-1} E(t_m-t_k)G(X_k)\Delta W_k. \quad (2.20)$$

Similarly, for arbitrary $t \in [0, T], \tilde{X}(t) = (\tilde{u}(t), \tilde{v}(t))^T$ defined by (2.19) can be expressed as

$$\tilde{X}(t) = E(t)X_0 + \int_0^t E(t-[s])F(X_{[s/\tau]}) \, ds + \int_0^t E(t-[s])G(X_{[s/\tau]}) \, dW(s), \quad (2.21)$$

where we define $[s/\tau]$ as an integer number not bigger than $s/\tau$ and $[s]_\tau = [s/\tau] \cdot \tau.$

Next, we show the result on the existence and uniqueness of the mild solution of (2.6).

**Theorem 2.1 (Existence and uniqueness of mild solution)** Suppose all conditions in Assumption 1.1 are fulfilled, let $W(t), t \in [0, T]$ be a cylindrical $Q$-Wiener process on the stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{0 \leq t \leq T}),$ given by (2.5), and let $X_0 = (u_0, v_0)^T$ be an $\mathcal{F}_0$-measurable $H$-valued random variable such that $\|X_0\|_{L^p(\Omega, H)} < \infty$ for some $p \in [2, \infty).$ Then (2.6) has a unique mild solution

$$X(t) = E(t)X_0 + \int_0^t E(t-s)F(X(s)) \, ds + \int_0^t E(t-s)G(X(s)) \, dW(s), \quad a.s. \quad (2.22)$$

for $t \in [0, T].$ Moreover there exists a constant $C_{p, T} \in [0, \infty)$ depending on $p, T$ such that

$$\sup_{t \in [0, T]} \|X(t)\|_{L^p(\Omega, H)} \leq C_{p, T} \left(\|X_0\|_{L^p(\Omega, H)} + 1\right). \quad (2.23)$$
Proof of Theorem 2.1. Owing to (1.5), (1.6) and the definition of $\| \cdot \|_H$, we infer that
\begin{equation}
\|F(X_1) - F(X_2)\|_H = \|\Lambda^{-\frac{1}{2}} ((F(u_1) - F(u_2))\|_U
\leq \|\Lambda^{-\frac{1}{2}} (u_1 - u_2)\|_U \leq L\|X_1 - X_2\|_H, \tag{2.24}
\end{equation}
for arbitrary $X = (u,v)^T, X_1 = (u_1,v_1)^T, X_2 = (u_2,v_2)^T$. Here $\tilde{L} = \|\Lambda^{-\frac{1}{2}} (u_1 - u_2)\|_U \leq \tilde{L} \|X_1 - X_2\|_H$. Similarly, one can obtain that
\begin{equation}
\|G(X_1) - G(X_2)\|_{\mathcal{Z}_2(U_0,H)} = \|\Lambda^{-\frac{1}{2}} ((G(u_1) - G(u_2))\|_{\mathcal{Z}_2(U_0,U)}
\leq \|\Lambda^{-\frac{1}{2}} (u_1 - u_2)\|_{\mathcal{Z}_2(U_0,U)} \leq \tilde{L}\|X_1 - X_2\|_H, \tag{2.26}
\end{equation}
and
\begin{equation}
\|G(X)\|_{\mathcal{Z}_2(U_0,H)} = \|\Lambda^{-\frac{1}{2}} G(u)\|_{\mathcal{Z}_2(U_0,U)} \leq \tilde{L}\|u\|_U + 1 \leq \tilde{L}\|X\|_H + 1 \tag{2.27}
\end{equation}
with $\tilde{L} = L/\lambda_1^{\frac{3}{2}}$. In view of Proposition 7.1 and Theorem 7.4 in [9], the existence and uniqueness of the mild solution (2.22) follow straightforwardly and (2.23) holds. \hfill \Box

In component-wise manner, the mild solution (2.22) takes the form
\begin{equation}
\begin{cases}
\frac{d}{dt} u(t) = C(t) u_0 + \int_t^\infty \Lambda^{-\frac{1}{2}} S(t-s) F(u(s)) \, ds \\
\frac{d}{dt} v(t) = -\Lambda^{-\frac{1}{2}} S(t) u_0 + \int_t^\infty \Lambda^{-\frac{1}{2}} S(t-s) G(u(s)) \, dW(s)
\end{cases} \tag{2.28}
\end{equation}
a.s. for $t \in [0,T]$. Therefore, $u(t)$ in (2.28) serves as the unique mild solution of (1.1).

3 Strong convergence

This section focuses on the strong convergence of the scheme (1.4). We begin by presenting a slightly modified version of the Burkholder-Davis-Gundy type inequality ([9, Lemma 7.2]).

Lemma 3.1 Let $(V, \langle \cdot, \cdot \rangle_V, \| \cdot \|_V)$ be a separable Hilbert space and let $\Psi : [0,T] \times \Omega \rightarrow \mathcal{L}_2(U_0,V)$ be a predictable stochastic process. Then for $t \in [0,T]$ and $p \in [2,\infty]$ there exists a constant $C_p$ such that
\begin{equation}
\left\| \sup_{s \in [0,t]} \int_0^s \Psi(r) \, dW(r) \right\|_V \leq C_p \left( \int_0^t \|\Psi(r)\|_{\mathcal{L}_2(U_0,V)}^2 \, dr \right)^{\frac{1}{2}}. \tag{3.1}
\end{equation}

The following result is on further spatial regularity of the mild solution $X(t)$ and its numerical approximations.

Proposition 3.1 Assume all conditions in Theorem 2.1 are fulfilled and $X_0 \in L^p(\Omega, H^{\alpha})$ for some $\alpha \in [0,\infty)$ and $p \in [2,\infty)$. Then it holds for $\gamma \in [0, \min(\alpha, \delta, 1)]$ that
\begin{equation}
\left\| \sup_{t \in [0,T]} \|X(t)\|_{H^\gamma} \right\|_{L^p(\Omega,R)} \leq C(\|X_0\|_{L^p(\Omega,H^{\alpha})} + 1), \tag{3.2}
\end{equation}
and
\begin{equation}
\left\| \sup_{t \in [0,T]} \|\tilde{X}(t)\|_{H^\gamma} \right\|_{L^p(\Omega,R)} \leq C(\|X_0\|_{L^p(\Omega,H^{\alpha})} + 1). \tag{3.3}
\end{equation}
**Proof of Proposition 3.1.** For the first step, we prove (3.2). From (2.22) we deduce that

\[
\sup_{t \in [0,T]} \|X(t)\|_{\mathcal{H}^T} \leq \sup_{t \in [0,T]} \|E(t)X_0\|_{\mathcal{H}^T} + \sup_{t \in [0,T]} \|E(t - T)\|_{\mathcal{Z}^T} \cdot \sup_{t \in [0,T]} \left( T \int_0^T \left\| E(T - s)F(X(s)) ds \right\|_{\mathcal{H}^T} \right) + \sup_{t \in [0,T]} \|E(t - T)\|_{\mathcal{Z}^T} \cdot \sup_{t \in [0,T]} \left( T \int_0^T \left\| E(T - s)G(X(s)) dW(s) \right\|_{\mathcal{H}^T} \right).
\]

Using (2.47) and elementary inequalities therefore results in

\[
\left\| \sup_{t \in [0,T]} \|X(t)\|_{\mathcal{H}^T} \right\|_{L^p(\Omega, \mathbb{R})} \leq \|X_0\|_{L^p(\Omega, \mathbb{R})} + T \sup_{t \in [0,T]} \left( T \int_0^T \|E(t - s)F(X(s))\|_{L^p(\Omega, \mathbb{R})} ds \right) + \sup_{t \in [0,T]} \left( T \int_0^T \|E(t - s)G(X(s)) dW(s)\|_{L^p(\Omega, \mathbb{R})} ds \right) = \|X_0\|_{L^p(\Omega, \mathbb{R})} + I_1 + I_2.
\]

Subsequently we will estimate \(I_1\) and \(I_2\) separately. First, (1.5), (2.23), the fact that \(\gamma \leq \min(\alpha, \delta, 1)\), the trigonometric identity and elementary inequalities enable us to get

\[
I_1 = \int_0^T \left( \left\| A^{-\frac{1}{2}}S(T - s)F(u(s)) \right\|_{L^2}^2 + \left\| C(T - s)F(u(s)) \right\|_{L^{p-1}}^{2} \right)^{1/2} ds
\]

\[
= \int_0^T \left\| A^{-\frac{1}{2}}F(u(s)) \right\|_{L^p(\Omega, \mathbb{R})} ds \leq \lambda_{1}^{-\frac{1}{2}} L \int_0^T \left( \|u(s)\|_{L^p(\Omega, \mathbb{R})} + 1 \right) ds \leq C(\|X_0\|_{L^p(\Omega, \mathbb{R})} + 1).
\]

Next, we note that

\[
\sup_{t \in [0,T]} \left\| \int_0^T E(T - s)G(X(s)) dW(s) \right\|_{\mathcal{H}^T} = \sup_{t \in [0,T]} \left( \left\| \int_0^T A^{-\frac{1}{2}}S(T - s)G(u(s)) dW(s) \right\|_{y}^2 + \left\| \int_0^T C(T - s)G(u(s)) dW(s) \right\|_{L^{p-1}}^2 \right)^{1/2}
\]

\[
\leq \sup_{t \in [0,T]} \left\| \int_0^T A^{-\frac{1}{2}}S(T - s)G(u(s)) dW(s) \right\|_{y} + \sup_{t \in [0,T]} \left\| \int_0^T A^{-\frac{1}{2}}C(T - s)G(u(s)) dW(s) \right\|_{L^{p-1}}.
\]
Finally, we use the Burkholder-Davis-Gundy type inequality \((3.1)\) and \((3.3)\) to arrive at

\[
|I_1|^2 \leq 2C_p^2 \int_0^T \left\| e^{T} S(T - s) G(u(s)) \right\|^2_{L^p(\Omega, \mathbb{R})} \, ds \\
+ 2C_p^2 \int_0^T \left\| e^{T} C(T - s) G(u(s)) \right\|^2_{L^p(\Omega, \mathbb{R})} \, ds \\
\leq 4C_p^2 \int_0^T \left\| e^{T} G(u(s)) \right\|^2_{L^p(\Omega, \mathbb{R})} \, ds \\
\leq 4C_p^2 \int_0^T \left\| e^{T} G(u(s)) \right\|^2_{L^p(\Omega, \mathbb{R})} \, ds \leq 4C_p^2 \int_0^T \left\| e^{T} G(u(s)) \right\|^2_{L^p(\Omega, \mathbb{R})} \, ds
\]

(3.6)

Here the stability properties of \( S(t), C(t), e^{T} \) for \( t \in [0, T] \), \( \gamma \leq \min(\alpha, \delta, 1) \) and some elementary inequalities were also used. Inserting the above estimates of \( I_1 \) and \( I_2 \) into \((3.4)\) and taking the assumption \( X_0 \in L^p(\Omega, H^m) \) into account yield \((3.2)\). To obtain \((3.3)\), we first derive from \((2.20)\) that

\[
\| X_m \|^2_{L^p(\Omega, H)} \leq 3 \| E(t_m) X_0 \|^2_{L^p(\Omega, H)} + 3m \tau \sum_{k=0}^{m-1} \| E(t_m - t_k) F(X_k) \|^2_{L^p(\Omega, H)} \\
+ 3C_p^2 \sum_{k=0}^{m-1} \| E(t_m - t_k) G(X_k) \|^2_{L^p(\Omega, \mathbb{R})} \, ds \\
\leq 3 \| X_0 \|^2_{L^p(\Omega, H)} + 3 \tau \sum_{k=0}^{m-1} \| F(X_k) \|^2_{L^p(\Omega, H)} + 3C_p^2 \tau \sum_{k=0}^{m-1} \| G(X_k) \|^2_{L^p(\Omega, \mathbb{R})} \\
\leq \mathcal{C}_{p,T} (\| X_0 \|^2_{L^p(\Omega, H)} + 1) + C_{p,T} \tau \sum_{k=0}^{m-1} \| X_k \|^2_{L^p(\Omega, H)},
\]

(3.7)

where \((2.11)\), \((2.25)\) and \((2.27)\) were employed. The discrete version of Gronwall’s inequality applied to \((3.7)\) and taking square roots show for all \( m \in \{0, 1, \cdots, M\} \) that

\[
\| X_m \|^2_{L^p(\Omega, H)} \leq C_{p,T} (\| X_0 \|^2_{L^p(\Omega, H)} + 1).
\]

(3.8)

With this and \((2.21)\), the proof of \((3.3)\) goes along the same way as before. □

Assumption 3.1 and Proposition 3.1 together immediately imply two corollaries as follows.

**Corollary 3.1** Assume all conditions in Theorem 2.7 are fulfilled and \( X_0 \in L^p(\Omega, H) \) for \( p \in [2, \infty) \). Let \( u(t) \) be given by \((2.28)\). Then there exists a constant \( C \), depending on \( p, L, T, \| X_0 \|^2_{L^p(\Omega, H)} \), such that

\[
\sup_{0 \leq t \leq T} \| F(u(t)) \|^2_{L^p(\Omega, U)} \leq C, \quad \text{and} \quad \sup_{0 \leq t \leq T} \| \Lambda^{\frac{\gamma}{2}} G(u(t)) \|^2_{L^p(\Omega, \mathbb{R}^d)} \leq C.
\]

(3.9)

**Corollary 3.2** Assume all conditions in Theorem 2.7 are fulfilled and \( X_0 \in L^p(\Omega, H) \) for \( p \in [2, \infty) \). Let \( u_m \) be given by \((1.3)\). Then there exists a constant \( C \), depending on \( p, L, T, \| X_0 \|^2_{L^p(\Omega, H)} \), such that

\[
\sup_{0 \leq m \leq M} \| F(u_m) \|^2_{L^p(\Omega, U)} \leq C, \quad \text{and} \quad \sup_{0 \leq m \leq M} \| \Lambda^{\frac{\gamma}{2}} G(u_m) \|^2_{L^p(\Omega, \mathbb{R}^d)} \leq C.
\]

(3.10)
In what follows we present a useful lemma, which can be found in [3,23]. The following Hölder continuity of \(E(t)\) will put the ultimate limit 1 on the convergence order one can achieve for the time-stepping scheme [1,4].

**Lemma 3.2** Assume that \(S(t)\) and \(C(t)\) are the sine and cosine operators and \(E(t)\) is the group as defined above. Then for all \(\gamma \in [0,1]\) there exists some constant \(c_\gamma\) such that

\[
\| (S(t) - S(s)) A^{-\frac{\gamma}{2}} \|_{L^p(U)} \leq c_\gamma (t-s)^\gamma, \quad \| (C(t) - C(s)) A^{-\frac{\gamma}{2}} \|_{L^p(U)} \leq c_\gamma (t-s)^\gamma
\]

(3.11)

and

\[
\| (E(t) - E(s)) X \|_H \leq c_\gamma (t-s)^\gamma \|X\|_H^\gamma
\]

(3.12)

for all \(t \geq s \geq 0\).

Equipped with this lemma, one can investigate the Hölder regularity in time of the mild solution (2.28), which is crucial in analyzing the approximation error of the time-discretization.

**Lemma 3.3** Assume conditions in Theorem [2,7] are all fulfilled and \(X_0 \in L^p(\Omega,H^\alpha)\) for some \(\alpha \in [0,\infty)\) and \(p \in [2,\infty)\). Then it holds for \(0 \leq s \leq t \leq T\) that

\[
\|u(t) - u(s)\|_{L^p(\Omega, U)} \leq C(t - s)^{\min(\alpha, \delta, 1)}.\]

(3.13)

**Proof of Lemma 3.3** To get (3.13), we first write

\[
X(t) - X(s) = (E(t-s) - \mathcal{P})X(s) + \int_s^t E(t-r)F(X(r))\,dr + \int_s^t E(t-r)G(X(r))\,dW(r),
\]

which admits

\[
u(t) - u(s) = (C(t-s) - I)u(s) + A^{-\frac{\delta}{2}} S(t-s)v(s)
+ \int_s^t A^{-\frac{\delta}{2}} S(t-r)F(u(r))\,dr + \int_s^t A^{-\frac{\delta}{2}} S(t-r)G(u(r))\,dW(r).
\]

(3.14)

In the course of the proof, we assign \(\rho = \min(\alpha, \delta, 1)\) for simplicity of presentation. The Burkholder-Davis-Gundy type inequality applied to the previous equality shows that

\[
\|u(t) - u(s)\|_{L^p(\Omega, U)} \leq \| (C(t-s) - I)u(s) \|_{L^p(\Omega, U)} + \| A^{-\frac{\delta}{2}} S(t-s)v(s) \|_{L^p(\Omega, U)}
+ \int_s^t \| A^{-\frac{\delta}{2}} S(t-r)F(u(r)) \|_{L^p(\Omega, U)}\,dr
+ C_p \left( \int_s^t \| A^{-\frac{\delta}{2}} S(t-r)G(u(r)) \|_{L^p(\Omega, \mathbb{R})}^2 \,dr \right)^{\frac{1}{2}}
\]

\[
\leq \| (C(t-s) - I)A^{-\frac{\delta}{2}} \|_{L^p(\Omega, U)} \cdot \| u(s) \|_{L^p(\Omega, H^\rho)}
+ \| A^{-\frac{\delta}{2}} S(t-s) \|_{L^p(\Omega, U)} \cdot \| v(s) \|_{L^p(\Omega, H^{\rho-1})}
+ \int_s^t \| A^{-\frac{\delta}{2}} S(t-r) \|_{L^p(\Omega, U)} \cdot \| F(u(r)) \|_{L^p(\Omega, U)}\,dr
+ C_p \left( \int_s^t \| A^{-\frac{\delta}{2}} S(t-r) \|_{L^p(\Omega, \mathbb{R})}^2 \,dr \right)^{\frac{1}{2}}
\]

(3.15)

for \(0 \leq s \leq t \leq T\) and \(p \in [2,\infty)\). Note that Proposition 3.1 ensures \(\|u(s)\|_{L^p(\Omega,H^\rho)} \leq \|X(s)\|_{L^p(\Omega,H^\rho)} < \infty\) and \(\|v(s)\|_{L^p(\Omega,H^{\rho-1})} \leq \|X(s)\|_{L^p(\Omega,H^\rho)} < \infty\) for any \(s \in [0,T]\). Combining these bounds with Corollary 3.1
and Lemma 3.2, one can easily deduce from (3.15) that
\[ \|u(t) - u(s)\|_{L^p(\Omega, U)} \leq C|t-s|^p + C \int_s^t (t-r) \|F(u(r))\|_{L^p(\Omega, U)} \, dr \]
\[ + C \left( \int_s^t (t-r)^{2\min(\delta, 1)} \|A^{\frac{\delta+1}{2}} G(u(r))\|_{L^p(\Omega, U)}^2 \, dr \right)^{\frac{1}{2}} \]
(3.16)
for \(0 \leq s \leq t \leq T\) and \(p \in [2, \infty)\). This finishes the proof of Lemma 3.3.

It is worthwhile to remark that one can similarly work with Corollary 3.3

As an immediate consequence, we have the following strong convergence result.

Proof of Theorem 3.1. Suppose that all conditions in Theorem 2.7 are fulfilled and assume \(X_0 \in L^p(\Omega, H^1)\) for some \(p \in [2, \infty)\). Then it holds that
\[ \left\| \sup_{t \in [0,T]} \|X(t) - \bar{X}(t)\|_H \right\|_{L^p(\Omega, R)} \leq C \tau^{\min(\delta, 1)}, \]
(3.17)
where \(X(t)\) is the mild solution of (2.6) and \(\bar{X}(t)\) is a continuous-time extension of \(X_m\), given by (2.19).

As an immediate consequence, we have the following strong convergence result.

Corollary 3.3. Suppose that all conditions in Theorem 3.7 are fulfilled. Then
\[ \left\| \sup_{t \in [0,T]} \|u(t) - \bar{u}(t)\|_U \right\|_{L^p(\Omega, R)} \leq C \tau^{\min(\delta, 1)}, \]
(3.18)
where \(u(t)\) and \(\bar{u}(t)\) are the first components of \(X(t)\) and \(\bar{X}(t)\), respectively.

The above results reveal that the order of strong convergence is essentially governed by the spatial regularity of the noise term and is in accordance with the exponents of the Hölder regularity in time as stated in Lemma 3.3. In addition, one can easily observe that the upper limit on strong order can be achieved only when \(\delta \in [1, \infty)\). Furthermore, as indicated in [8] for the linear SWE, the exponential scheme (1.4) for nonlinear SWE also allows for higher strong order than the implicit Euler and Crank-Nicolson time discretizations do (compare Corollary 3.3 with [23] Theorem 4.6).

Proof of Theorem 3.7. Combining (2.21) and (2.22) gives
\[ X(s) - \bar{X}(s) = \int_0^s \left( E(s-r)F(X(r)) - E(s-\lfloor r \rfloor \tau)F(X_{\lfloor r \rfloor \tau}) \right) \, dr \]
\[ + \int_0^s \left( E(s-r)G(X(r)) - E(s-\lfloor r \rfloor \tau)G(X_{\lfloor r \rfloor \tau}) \right) \, dW(r). \]
(3.19)

Thanks to the same arguments as used in (3.4), one can similarly get
\[ \left\| \sup_{s \in [0,T]} \|X(s) - \bar{X}(s)\|_H \right\|_{L^p(\Omega, R)} \]
\[ \leq \int_0^T \left\| E(t-r)F(X(r)) - E(t-\lfloor r \rfloor \tau)F(X_{\lfloor r \rfloor \tau}) \right\|_{L^p(\Omega, H)} \, dr \]
\[ + \left\| \sup_{s \in [0,T]} \left( E(t-r)G(X(r)) - E(t-\lfloor r \rfloor \tau)G(X_{\lfloor r \rfloor \tau}) \right) \right\|_{L^p(\Omega, R)} \]
\[ = I_3 + I_4. \]
(3.20)
Noting that $\bar{X}([r]_\tau) = X_{[r]_\tau}$ and using (2.17), (2.24), (3.9), (3.13) and Lemma 3.2, we find that

$$I_3 \leq \int_0^t \left( \| E(t-r) - E(t-[r]_\tau) \|_{L^p(\Omega, H)} \right) \, dr$$

$$+ \int_0^t \left( \| E(t-[r]_\tau) \|_{L^p(\Omega, H)} \right) \, dr$$

$$+ \int_0^t \left( \| E(t-[r]_\tau) \|_{L^p(\Omega, H)} \right) \, dr$$

$$\leq c_1 \int_0^t \| F(X(r)) \|_{L^p(\Omega, H)} \, dr + \int_0^t \| F(X(r)) - \bar{F}(X([r]_\tau)) \|_{L^p(\Omega, H)} \, dr$$

$$\leq c_1 \tau + C_\tau \min(\delta, 1) + C_\tau \int_0^t \| X(r) - \bar{X}(r) \|_H \, dr$$

With regard to $I_4$, we apply the Burkholder-Davis-Gundy type inequality (3.1) to get

$$|I_4|^2 \leq C_\delta^2 \int_0^t \left( \| E(t-r) \|_{L^p(\Omega, H)} \right)^2 \, dr$$

$$\leq 3C_\delta^2 \int_0^t \left( \| E(t-r) \|_{L^p(\Omega, H)} \right)^2 \, dr$$

$$+ \int_0^t \left( \| E(t-[r]_\tau) \|_{L^p(\Omega, H)} \right)^2 \, dr$$

Further employing (2.17), (2.26), (3.9), (3.13) and Lemma 3.2 shows that

$$|I_4|^2 \leq C_\delta \tau^{2 \min(\delta, 1)} \int_0^t \left( \| \hat{X}(r) \|^{2 \min(\delta, 1)} \right) \, dr$$

$$\leq C_\delta \tau^{2 \min(\delta, 1)} \int_0^t \left( \| \hat{X}(r) \|_{L^p(\Omega, H)} \right)^2 \, dr$$

$$\leq C_\delta \tau^{2 \min(\delta, 1)} \int_0^t \left( \| \hat{X}(r) \|_{L^p(\Omega, H)} \right)^2 \, dr$$

$$\leq C_\delta \tau^{2 \min(\delta, 1)} + C_\tau \int_0^t \| X(r) - \bar{X}(r) \|_H^2 \, dr.$$
Moreover, the initial data
\[ X(0) = \bar{X} \]
and taking the estimates (3.21) and (3.23) into account we derive from (3.20) that
\[ \vartheta(t) = \left\| \sup_{s \in [0,t]} \| X(s) - \bar{X}(s) \| \right\|_{L^p(\Omega, \mathbb{R})}^2 \leq C t^{2 \min(\delta, 1)} + C \int_0^t \vartheta(s) \, ds. \] (3.24)

Proposition 3.1 guarantees the boundedness of \( \vartheta(t) \) for \( t \in [0,T] \), which enables us to apply Gronwall’s lemma (see, e.g., [29, Theorem 8.1]) to show for all \( t \in [0,T] \) that
\[ \vartheta(t) = \left\| \sup_{s \in [0,t]} \| X(s) - \bar{X}(s) \| \right\|_{L^p(\Omega, \mathbb{R})}^2 \leq C t^{2 \min(\delta, 1)}. \] (3.25)

Finally, taking square roots of (3.25) completes the proof of Theorem 3.1. \( \square \)

4 Applications to examples

The aim of this section is to include several concrete examples which fit in the abstract setting formulated above. To this end, let \( \mathcal{O} \subset \mathbb{R}^d \), \( d = 1, 2, 3 \), be a bounded open set with smooth boundary \( \partial \mathcal{O} \). A nonlinear white noise driven stochastic wave equation (SWE) with Dirichlet boundary condition is usually described by
\[
\begin{cases}
\frac{\partial^2 u}{\partial t^2} = \Delta u + f(u, \xi, u) \mathcal{W}(t), & t \in (0, T], \xi \in \mathcal{O},
\frac{\partial u}{\partial n} = 0, & t \in (0, T],
u(0, \xi) = u_0(\xi), \frac{\partial u}{\partial \xi}(0, \xi) = v_0(\xi), & \xi \in \mathcal{O},
\end{cases}
\] (4.1)

where \( T \in (0, \infty) \), \( \Delta = \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2} \) is the Laplace operator and \( f, g: \mathcal{O} \times \mathbb{R} \rightarrow \mathbb{R} \) are deterministic functions. Moreover, the initial data \( u_0, v_0: \mathcal{O} \times \Omega \rightarrow \mathbb{R} \) are random variables and \( \mathcal{W} \) is a noise process, which will be specified later. Numerics of such equation has been considered in [6, 15, 34, 36, 38].

Let \( u := L^2(\mathcal{O}, \mathbb{R}) \) be the separable Hilbert space of real-valued square integrable functions from \( \mathcal{O} \) to \( \mathbb{R} \), with the scalar product and the norm
\[ \langle u_1, u_2 \rangle_U = \int_{\mathcal{O}} u_1(\xi) u_2(\xi) d\xi, \quad \| u \|_U = \left( \int_{\mathcal{O}} |u(\xi)|^2 d\xi \right)^{1/2} \]
for all \( u, u_1, u_2 \in U \). Moreover, let \( \{\mathcal{W}(t)\}_{t \in [0,T]} \) be a cylindrical \( \mathcal{Q} \)-Wiener process on a stochastic basis \( (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t \in [0,T]}) \), given by (2.5). It is a classical result that the covariance operator \( \mathcal{Q} = \mathcal{Q}^\perp \circ \mathcal{Q}^\perp \in \mathcal{L}(U) \) is a nonnegative, symmetric operator so that
\[ \mathcal{Q} \phi_i = q_i \phi_i, q_i \geq 0, i \in \mathbb{N}. \] (4.2)

Then one can rewrite (4.1) in an abstract Itô form as (1.1), namely,
\[
\begin{cases}
\frac{d\hat{u}}{dt} = -\Delta \hat{u} \, dt + F(\hat{u}) \, dt + G(\hat{u}) \, d\mathcal{W}(t), & t \in (0, T],
\hat{u}(0) = u_0, \hat{u}(0) = v_0,
\end{cases}
\]

where \( -\Delta: \mathcal{D}(\Delta) \subset U \rightarrow U \) denotes the Laplacian with homogeneous Dirichlet boundary condition, and where \( F: U \rightarrow U, G(\hat{u}): U \rightarrow U \) are the Nemyskij operators given by
\[ F(u)(\xi) = f(\xi, u(\xi)), \quad (G(u)(\phi))(\xi) = g(\xi, u(\xi)) \cdot \phi(\xi), \quad \xi \in \mathcal{O}. \] (4.3)

With the above setting, we take a close look at conditions in Assumption 1.1 To do this we consider separately several cases as follows.
4.1 Additive noise

For the first case, we consider SWEs driven by additive noise with \( g(\xi, u) \equiv 1 \) for all \( \xi \in \mathcal{O}, u \in \mathbb{R} \). Furthermore, assume for some \( \beta \in (0, \infty) \) that

\[
\| \Lambda^{\frac{\beta-1}{2}} Q^\frac{1}{2} \|_{\mathcal{L}_2(U)} = \left( \sum_{i=1}^m q_i \| \Lambda^{\frac{\beta-1}{2}} \phi_i \|_U^2 \right)^{\frac{1}{2}} < \infty. \tag{4.4}
\]

Concerning \( f \), we assume \( f: \mathcal{O} \times \mathbb{R} \to \mathbb{R} \) in (4.1) satisfies

\[
|f(\xi, u)| \leq L(|u| + 1), \quad |f(\xi, u_1) - f(\xi, u_2)| \leq L|u_1 - u_2| \tag{4.5}
\]

for all \( \xi \in \mathcal{O}, u, u_1, u_2 \in \mathbb{R} \). Then the Nemytskij operator \( F \) satisfies

\[
||F(u)||_0^2 = \int_\mathcal{O} |f(\xi, u(\xi))|^2 d\xi \leq 2L^2 \int_\mathcal{O} \left( |u(\xi)|^2 + 1 \right) d\xi = 2L^2 \left( \|u\|_U^2 + \rho(\mathcal{O}) \right), \tag{4.6}
\]

where \( \rho(\mathcal{O}) \), the measure of the set \( \mathcal{O} \), is bounded by assumption. In a similar way,

\[
||F(u_1) - F(u_2)||_{\mathcal{L}_2} \leq L||u_1 - u_2||_U. \tag{4.7}
\]

Moreover, \( G(u) \equiv I \) in this setting and thus \( G(u_1) - G(u_2) \equiv 0 \) and

\[
||\Lambda^{\frac{\beta-1}{2}} G(u)\|_{\mathcal{L}_2(\mathcal{O})} = ||\Lambda^{\frac{\beta-1}{2}} Q^\frac{1}{2} \|_{\mathcal{L}_2(U)} < \infty. \tag{4.8}
\]

Therefore, Assumption 1.1 is fulfilled with \( \delta = \beta \). An immediate consequence of Corollary 3.3 reads:

**Corollary 4.1** Suppose that \( g(\xi, u) \equiv 1 \) for all \( \xi \in \mathcal{O}, u \in \mathbb{R} \) and \( f: \mathcal{O} \times \mathbb{R} \to \mathbb{R} \) satisfies (4.5). Let \( W(t) \) be a cylindrical \( Q \)-Wiener process on the stochastic basis \( (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{0 \leq t \leq T}) \) with (4.4) fulfilled. Additionally, we assume \( u_0, v_0 \in \mathcal{F}_0 \) and \( u_0 \in L^p(\Omega, H^1), v_0 \in L^p(\Omega, U) \) for some \( p \in [2, \infty) \). Then the problem (4.1) has a unique mild solution. Moreover, there exists a generic constant \( C \in [0, \infty) \) depending on \( T, L, p, \beta \) and the initial data, such that

\[
\left\| \sup_{t_n \in \mathcal{M}} \| u(t_n) - u_m \|_{L^p(\Omega, \mathbb{R})} \right\| \leq C N^\min(\beta, 1), \tag{4.9}
\]

where \( u(t) \) is the mild solution of (4.1) and \( u_m \) is the numerical solution produced by (1.4).

4.2 Multiplicative trace class noise

Assume \( f: \mathcal{O} \times \mathbb{R} \to \mathbb{R} \) in (4.1) satisfies (4.5) and \( g: \mathcal{O} \times \mathbb{R} \to \mathbb{R} \) in (4.1) satisfies

\[
|g(\xi, u)| \leq L(|u| + 1), \quad |g(\xi, u_1) - g(\xi, u_2)| \leq L|u_1 - u_2| \tag{4.10}
\]

for all \( \xi \in \mathcal{O}, u, u_1, u_2 \in \mathbb{R} \). Furthermore, we assume in this subsection

\[
\text{Tr}(Q) = \sum_{i \in \mathbb{N}} q_i < \infty, \quad \sup_{\xi \in \mathcal{O}} \sup_{\xi \in \mathcal{O}} |\phi(\xi)| < \infty. \tag{4.11}
\]

Then it is not difficult to see that

\[
\|G(u)\|_{\mathcal{L}_2(\mathcal{O})}^2 = \|G(u)Q^\frac{1}{2}\|_{\mathcal{L}_2(U)}^2 = \sum_{i=1}^m \|G(u)Q^\frac{1}{2}\phi_i\|_U^2 \leq \sum_{i=1}^m q_i \|G(u)\phi_i\|_U^2 = \sum_{i=1}^m q_i \int_\mathcal{O} |g(\xi, u(\xi))\phi_i(\xi)|^2 d\xi \leq \sup_{\xi \in \mathcal{O}} \sup_{\xi \in \mathcal{O}} |\phi(\xi)|^2 \cdot \text{Tr}(Q) \cdot 2L^2 \left( \|u\|_U^2 + \rho(\mathcal{O}) \right). \tag{4.12}
\]
In the same way, one can obtain
\[
\left\| (G(u_1) - G(u_2))Q^2 \right\|_{\mathcal{L}(U)}^2 \leq \sup_{i \in \mathbb{N}} \sup_{\xi \in \Theta} |\phi_i(\xi)|^2 \cdot \text{Tr}(Q) \cdot L^2 \| u_1 - u_2 \|^2_U. \tag{4.13}
\]

Consequently, Assumption 1.1 is fulfilled with \( \delta = 1 \) and the following corollary follows.

**Corollary 4.2** Suppose that \( f, g : \Theta \times \mathbb{R} \to \mathbb{R} \) satisfy (4.5) and (4.10). Let \( W(t) \) be a standard \( U \)-valued Q-Wiener process on a stochastic basis \( (\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{0 \leq t \leq T}) \) with (4.11) fulfilled. Additionally, we assume \( u_0, v_0 \in \mathcal{F}_0 \) and \( u_0 \in L^p(\Omega, H^1) \), \( v_0 \in L^p(\Omega, U) \) for some \( p \in [2, \infty) \). Then the problem (4.1) has a unique mild solution and it holds that
\[
\left\| \sup_{t_n \in [0,T]} \| u(t_n) - u_m \|_U \right\|_{L^p(\Omega, \mathbb{R})} \leq C \tau, \tag{4.14}
\]
where \( u(t) \) is the mild solution of (4.1) and \( u_m \) is produced by (1.4).

### 4.3 Multiplicative space-time white noise

In this subsection, let \( d = 1, \Theta = (0, 1) \) and let \( Q = I \). Then \( W(t) \) becomes a cylindrical \( I \)-Wiener process, which can be given by
\[
W(t) = \sum_{i \in \mathbb{N}} \beta_i(t)e_i, \quad t \in [0, T],
\tag{4.15}
\]
where \( \{\beta_i(t)\}_{i \in \mathbb{N}} \) are a family of mutually independent real Brownian motions and \( \{e_i = \sqrt{2}\sin(\pi i x), x \in (0, 1)\}_{i \in \mathbb{N}} \) form an orthonormal basis of \( U \) consisting of eigenfunctions of \( \Lambda \) with \( \Lambda e_i = \lambda_i e_i, \lambda_i = \pi^2 i^2, i \in \mathbb{N} \). It is then obvious for any \( \varepsilon > 0 \) that
\[
\left\| \Lambda^{-\frac{\varepsilon+1}{2}} \right\|_{\mathcal{L}(U)}^2 = \pi^{-(\varepsilon+1)} \sum_{i=1}^{\infty} i^{-(\varepsilon+1)} < \infty, \quad \text{and} \quad \sup_{i \in \mathbb{N}} \sup_{x \in [0,1]} |e_i(x)| \leq \sqrt{2}. \tag{4.16}
\]

Therefore, it follows that
\[
\left\| \Lambda^{-\frac{\varepsilon+1}{2}} G(u) \right\|_{\mathcal{L}(U)}^2 = \sum_{i=1}^{\infty} \left\| \Lambda^{-\frac{\varepsilon+1}{2}} G(u)e_i \right\|_U^2 = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \left| \langle \Lambda^{-\frac{\varepsilon+1}{2}} G(u)e_i, e_j \rangle \right|^2
\]
\[
= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \left| \int_{\Theta} g(\xi, u(\xi))e_i(\xi)e_j(\xi) d\xi \right|^2
\]
\[
\leq 2 \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \left| \int_{\Theta} g(\xi, u(\xi))e_i(\xi) d\xi \right|^2
\]
\[
= 2\pi^{-(\varepsilon+1)} \sum_{i=1}^{\infty} i^{-(\varepsilon+1)} \sup_{x \in [0,1]} \|g(\cdot, u(\cdot))\|_U^2 \leq C_\varepsilon \left( \|u\|^2_U + \rho(\Theta) \right),
\tag{4.17}
\]
for arbitrarily small \( \varepsilon > 0 \). Similarly, one can show for arbitrarily small \( \varepsilon > 0 \) that
\[
\left\| \Lambda^{-\frac{\varepsilon+1}{2}} (G(u_1) - G(u_2)) \right\|_{\mathcal{L}(U)} \leq C_\varepsilon \|u_1 - u_2\|_U. \tag{4.18}
\]

Hence, Assumption 1.1 is satisfied with \( \delta = \frac{1-\varepsilon}{2} \) for arbitrarily small \( \varepsilon > 0 \).
Corollary 4.3 Suppose that \( f, g : (0, 1) \times \mathbb{R} \to \mathbb{R} \) satisfy (4.5) and (4.10). Additionally, we assume \( u_0, v_0 \in \mathcal{F}_0 \) and \( v_0 \in L^p(\Omega, H^1), v_0 \in L^p(\Omega, U) \) for some \( p \in [2, \infty] \). Moreover, let \( W(t) \) be a cylindrical \( I \)-Wiener process given by (4.15). Then the problem (4.1) has a unique mild solution. Moreover, there exists a constant \( C \in [0, \infty) \) depending on \( T, L, p, \varepsilon \) and the initial data, such that

\[
\| u(t) - u_m \|_{L^p(\Omega, \mathbb{R})} \leq C \varepsilon^{\frac{1}{2} + \beta},
\]

where \( u(t) \) is the mild solution of (4.1) and \( u_m \) is the numerical solution produced by (1.4).

5 Weak convergence

In many applications, a key task is to approximate the quantity \( \mathbb{E}[\varphi(u(T))] \), where \( \varphi \) is a functional of the mild solution to (1.1). This leads to another important notion of convergence for a numerical scheme, the weak convergence, which is concerned with the approximation of law. In this section, let us turn our attention to this topic. As is well known that the rate of the weak error can, in some situations, be improved compared to that of the strong error. Below, we shall show this fact for nonlinear SWE driven by additive noise. Since the treatment of multiplicative noise case causes additional technical difficulties, it will be addressed elsewhere in our future work.

For the additive case when \( G(u) \equiv I \) for \( u \in U \), the abstract equation (2.6) reduces to

\[
\begin{align*}
\frac{dX(t)}{dt} &= AX(t)dt + F(X)dt + BdW(t), & t \in (0, T], \\
X(0) &= X_0,
\end{align*}
\]

(5.1)

where we adopt the same notations as in (2.7) and additionally define here \( B : U \to H \) by \( B = (0, I)^T \).

To begin with, we need the following assumption.

Assumption 5.1 Assume \( G(u) \equiv I \) for all \( u \in U \), and let \( F : U \to U \) be a twice differentiable mapping satisfying

\[
\begin{align*}
\| F(u) \|_U &\leq L(\| u \|_U + 1), \\
\| F'(u) \|_U &\leq L \| \psi_1 \|_U, \\
\| F''(u) \|_U &\leq L \| \psi_1 \|_U \| \psi_2 \|_U
\end{align*}
\]

for all \( u, \psi, \psi_1, \psi_2 \in U \). Furthermore we assume (1.8) holds for some \( \beta \in (0, 1) \).

Several comments should be added concerning the above assumption. We first mention that the essential condition (1.8) has been also used in (4.2.23), to study weak errors for the linear SWE (\( F \equiv 0 \)) with additive noise. Secondly, it should be pointed out that \( F \) in the previous assumption is not required to be twice Fréchet differentiable in \( U \), which is in general not fulfilled for the Nemitskij operators (see Example 5.1 below). Besides, recall that \( \| \Lambda^{-\frac{1}{2}} Q^2 \|_{\mathcal{L}(U)}^2 \leq \| \Lambda^{-\frac{1}{2} + i\beta} Q \Lambda^{-\frac{1}{2}} \|_{\mathcal{L}(U)} \) (see Theorem 2.1 in (4.2.23)). Accordingly, it is easy to see that Assumption (1.4) is fulfilled with \( \delta = \beta \) under Assumption 5.1. In other words, Assumption 5.1 suffices to guarantee a unique mild solution of (5.1) and the strong convergence order of \( \min(\beta, 1) \) for the underlying scheme (1.4). When \( \beta \in [1, \infty) \), Theorem 3.1 shows that the scheme possesses the strong order of 1. This trivially implies weak convergence order of 1, which can not be improved further as shown in the later weak convergence result (Theorem 5.1). Based on the above observations, we take condition (1.8) with \( \beta \in (0, 1) \) here, instead of \( \beta \in (0, \infty) \).

Theorem 5.1 Suppose that all conditions in Assumption 5.1 are satisfied and assume additionally that for \( \kappa = \min(\beta, \frac{1}{2} - \varepsilon) \) with arbitrarily small \( \varepsilon \in (0, \frac{1}{2}) \),

\[
\| F'(u)z \|_{-\kappa} \leq C \| z \|_{-\kappa}(\| u \|_\kappa + 1), \quad u \in H^\kappa, z \in U,
\]

(5.3)

and \( X_0 = (u_0, v_0)^T \in H^1 \) is deterministic. Then there exists a constant \( C \) depending on \( T, \beta, \varepsilon, \Phi, L \) and the initial data, such that

\[
\mathbb{E}[\Phi(X(T))] - \mathbb{E}[\Phi(X_0)] \leq C \varepsilon^{\min(2\beta, \frac{1}{2} + \beta - \varepsilon, 1)}
\]

(5.4)
for $\Phi \in C^2_b(H; \mathbb{R})$, and where $X(t), t \in [0, T]$ is the mild solution of (5.1) and $X_m, m = 0, 1, \ldots M$ are produced by the recurrence equation (2.13).

The proof of Theorem 5.1 is postponed after some preparations. Here and below, by $C_b^k(V, V')$ we denote the space of not necessarily bounded mappings from a Banach space $V$ to the other Banach space $V'$ that have continuous and bounded Fréchet derivatives up to order $k$. For the particular case when $V$ is a Hilbert space and $V' = \mathbb{R}$, one can identify the first derivative $D\Psi(X) \in \mathcal{L}(V, \mathbb{R})$ of a function $\Psi : V \rightarrow \mathbb{R}$ with an element in $V$ due to the Riesz representation theorem, i.e.,

$$D\Psi(X) Z = \langle D\Psi(X), Z \rangle_V, \quad \forall \ X, Z \in V,$$

and the second derivative $D^2\Psi(X)$ with a bounded linear operator such that

$$D^2\Psi(X)(Z_1, Z_2) = \langle D^2\Psi(X)Z_1, Z_2 \rangle_V, \quad \forall \ X, Z_1, Z_2 \in V.$$

Next, we make some comments concerning the new key condition 5.3. Such condition imposed on the derivative operator $F'(u), u \in U$ can be fulfilled in a concrete setting established in Section 4. More specifically, we set $U := L^2((0, 1), \mathbb{R})$ and let $F : U \rightarrow U$ be the Nemitskij operator given by $F(u)(\xi) = f(\xi, u(\xi)), u \in U, \xi \in (0, 1)$ with $f : \partial \times \mathbb{R} \rightarrow \mathbb{R}$ being a real-valued function. When $f$ has bounded partial derivatives up to order two as required in (5.51), the condition (5.3) can be fulfilled (see Example 5.1 below for more details on the verification). Of course, the condition (5.3) is hard to be satisfied if the associated function $f$ grows super-linearly or is non-Lipschitz.

An immediate consequence of Theorem 5.1 gives the following result.

**Corollary 5.2** Assume all conditions in Theorem 5.1 are fulfilled. Then it holds for arbitrarily small $\varepsilon \in (0, \frac{1}{2})$ that

$$|\mathbb{E}[\phi(u(T))] - \mathbb{E}[\phi(u_M)]| \leq C_\varepsilon \varepsilon^{\min(2\beta, \frac{1}{2} + \beta - \varepsilon, 1)}$$

(5.5) for $\phi \in C^2_b(U; \mathbb{R})$. Here $u(T)$ and $u_M$ are the first components of $X(T)$ and $X_M$, respectively.

Comparing the weak error bound (5.5) with the strong one (4.9), one can find that the rate of weak convergence is, as expected, improved. To see this, let us look at the special interesting case of space-time white noise ($Q = I$). In this case, (4.9) admits a strong convergence order $\beta < \frac{1}{2}$, while (5.5) gives a weak convergence order $1 - \varepsilon$ for arbitrarily small $\varepsilon > 0$. A closely related work on weak convergence of numerical schemes for SWEs can be found in [23], where only the linear SWE was considered and time discretizations were done by rational approximation to the exponential function. The corresponding weak rates in [23] are min$(\frac{1}{p+1}, \beta, 1)$, with $p \geq 1$ being method parameters. Particularly, the linear implicit Euler method ($p = 1$) admits weak rate of min$(\beta, 1)$ and the Crank-Nicolson method ($p = 2$) admits weak rate of min$(\frac{1}{2}, \beta, 1)$. Apparently, the exponential scheme attains better weak convergence rates than time discretization schemes in [23].

To carry out the weak error analysis, we first introduce a function $\mu : [0, T] \times H \rightarrow \mathbb{R}$, defined by

$$\mu(t, x) = \mathbb{E}[\Phi(X(t, x))], \quad t \in [0, T], \quad x \in H,$$

(5.6)

where $\Phi \in C^2_b(H; \mathbb{R})$ and $X(t, x)$ is the unique mild solution of (5.1) with the initial value $x \in H$. Owing to (5.2), one can readily infer that

$$\|F'(X) Z\|_H = \|A^{-\frac{1}{2}} F'(u) u\|_L \leq \lambda_1^{-\frac{1}{2}} L \|u\|_U \leq \lambda_1^{-\frac{1}{2}} L \|Z\|_H,$$

$$\|F''(X)(Z_1, Z_2)\|_H = \|A^{-\frac{1}{2}} F''(u)(u^1, u^2)\|_U \leq L \|u^1\|_U \|u^2\|_U \leq L \|Z_1\|_U \|Z_2\|_H$$

(5.7)

for all $X = (u, v)^T \in H, Z = (u^1, v^1)^T \in H, Z_1 = (u^1, v^3)^T \in H$ and $Z_2 = (u^2, v^2)^T \in H$. This straightforwardly implies $F \in C_b^2(H, H)$. By [9, Theorem 9.4], one knows that the solution $X(t, x)$ is twice differentiable with respect to the initial value $x$. More formally, the process $\zeta^h(t) = \frac{d}{dt}(X(t, x))h$ for $t \in [0, T], h \in H$ is the mild solution of the following equation

$$\left\{ \begin{array}{ll} d\zeta^h = (A\zeta^h + F'(X(t, x)) \zeta^h) dt, & t \in (0, T], \\ \zeta^h(0) = h. & \end{array} \right.$$

(5.8)
And $\eta^{h,g}(t) = \frac{\partial^2 X(t,x)}{\partial x^2}(t,x)(h,g)$ for $t \in [0,T]$, $h,g \in H$ is the mild solution of

$$
\begin{cases}
    d\eta^{h,g} = \left( A\eta^{h,g} + F'(X(t,x))\eta^{h,g} + F''(X(t,x))(\xi^h, \xi^g) \right) dt, & t \in (0,T], \\
    \eta^{h,g}(0) = 0.
\end{cases}
$$

(5.9)

Bearing these facts in mind and differentiating (5.6) with respect to $x$ yield

$$
D\mu(t,x) h = \mathbb{E}\left[ D\Phi(X(t,x)) \xi^h(t) \right], \quad t \in [0,T], \ x, h \in H.
$$

(5.10)

Differentiating (5.10) further shows

$$
D^2\mu(t,x)(h,g) = \mathbb{E}\left[ D^2\Phi(X(t,x)) (\xi^h(t), \xi^g(t)) \right] + \mathbb{E}\left[ D\Phi(X(t,x)) \eta^{h,g}(t) \right], \ t \in [0,T], \ x, h, g \in H,
$$

(5.11)

where $\xi^h(t), \xi^g(t)$ and $\eta^{h,g}(t)$ for $t \in [0,T], h, g \in H$ are processes defined as above. Then it is well-known (see [2] Theorem 9.17, [10] Theorem 5.4.2) that $\mu(t,x)$ defined by (5.6) is a unique strict solution to the following deterministic PDE:

$$
\begin{cases}
    \frac{\partial \mu}{\partial t}(t,x) = \langle Ax + F(x), D\mu(t,x) \rangle_H + \frac{1}{2} \text{Tr}\left[ D^2\mu(t,x) BQ^2 (BQ^2)^* \right], & t \in (0,T), \ x \in \mathcal{D}(A), \\
    \mu(0,x) = \Phi(x), & x \in H.
\end{cases}
$$

(5.12)

In order to eliminate the operator $A$, we introduce another process $v: \ [0,T] \times H \to \mathbb{R}$, defined by

$$
v(t,y) = \mu(t,E(-t)y), \quad t \in [0,T], \ y \in H.
$$

(5.13)

By virtue of (5.12), and also noticing that $\frac{\partial}{\partial t}(E(-t)y) = -AE(-t)y$ for $y \in \mathcal{D}(A)$, one can derive that

$$
\begin{align}
\frac{\partial v}{\partial t}(t,y) &= \frac{\partial \mu}{\partial t}(t,E(-t)y) + \langle D\mu(t,E(-t)y), -AE(-t)y \rangle_H \\
&= \langle AE(-t)y + F(E(-t)y), D\mu(t,E(-t)y) \rangle_H \\
&\quad + \frac{1}{2} \text{Tr}\left[ D^2\mu(t,E(-t)y) BQ^2 (BQ^2)^* \right] - \langle D\mu(t,E(-t)y), AE(-t)y \rangle_H \\
&= \langle F(E(-t)y), D\mu(t,E(-t)y) \rangle_H + \frac{1}{2} \text{Tr}\left[ D^2\mu(t,E(-t)y) BQ^2 (BQ^2)^* \right].
\end{align}
$$

(5.14)

Note that

$$
Dv(t,y) z = D\mu(t,E(-t)y) E(-t)z, \quad t \in [0,T], \ y, z \in H,
$$

(5.15)

and

$$
D^2v(t,y)(z_1, z_2) = D^2\mu(t,E(-t)y)(E(-t)z_1, E(-t)z_2), \quad t \in [0,T], \ y, z_1, z_2 \in H.
$$

(5.16)

The above three identities together with (5.4) immediately imply that

$$
\begin{cases}
    \frac{\partial v}{\partial t}(t,y) = \frac{1}{2} \text{Tr}\left[ D^2v(t,y) E(t) BQ^2 (E(t) BQ^2)^* \right] + \langle E(t) F(E(-t)y), Dv(t,y) \rangle_H, & t \in (0,T], \ y \in \mathcal{D}(A) \\
    v(0,y) = \Phi(y), & y \in H.
\end{cases}
$$

(5.17)

Before proceeding further, we show regularity results on $Dv(t,y)$ and $D^2v(t,y)$.

**Lemma 5.1** Assume Assumption 5.1 holds. Then $Dv(t,y): H \to \mathbb{R}$ given by (5.15) and $D^2v(t,y): H \times H \to \mathbb{R}$ given by (5.16) for any $t \in [0,T]$ and $y \in H$ satisfy

$$
\|Dv(t,y)\|_H \leq C, \quad \text{and} \quad \|D^2v(t,y)\|_{\mathcal{L}(H)} \leq C
$$

(5.18)

for all $t \in [0,T]$ and $y \in H$. 

Proof of Lemma 5.2. First of all, we shall prove for all $t \in [0, T]$ and $x \in H$ that
\[
\|D\mu(t,x)\|_H \leq C, \quad \text{and} \quad \|D^2\mu(t,x)\|_{\mathcal{L}(H)} \leq C. \quad (5.19)
\]
Then (5.18) is an easy consequence of (5.19) thanks to (5.15), (5.16) and (2.17). Hence we focus on the proof of (5.19). Recall that $\zeta^h(t) = \frac{2X(t)}{\partial s}h$ is the mild solution of (5.8), given by
\[
\zeta^h(t) = E(t)h + \int_0^t E(t-s)F'(X(s,x))\zeta^h(s)\,ds, \quad t \in [0, T], \ h \in H. \quad (5.20)
\]
Therefore, (2.17) and (5.7) guarantee that
\[
\|\zeta^h(t)\|_H \leq \|E(t)h\|_H + \lambda_1^{-\frac{1}{2}}L \int_0^t \|\zeta^h(s)\|_H \, ds, \quad t \in [0, T]. \quad (5.21)
\]
The Gronwall inequality thus shows for all $t \in [0, T]$ that
\[
\|\zeta^h(t)\|_H \leq C\|h\|_H. \quad (5.22)
\]
Since $\Phi \in \mathcal{C}_0^2(H, \mathbb{R})$ and (5.22) holds, one can derive from (5.10) that
\[
\|D\mu(t,x)\| \leq \mathbb{E}\left[\|D\Phi(X(t,x))\|_H \cdot \|\zeta^h(t)\|_H\right] \leq C\|h\|_H \quad (5.23)
\]
for all $t \in [0, T], h \in H$. Likewise, $\eta^{h,g}(t) = \frac{2X(t,x)}{\partial x} \delta(t,g)$ for $t \in [0, T], h, g \in H$ is given by
\[
\eta^{h,g}(t) = \int_0^t E(t-s) \left(F'(X(s,x))\eta^{h,g}(s) + F''(X(s,x))(\zeta^h(s), \zeta^g(s))\right)\,ds. \quad (5.24)
\]
With (5.7) and (5.22) at hand, we first obtain
\[
\|\eta^{h,g}(t)\|_H \leq C\|h\|_H \cdot \|g\|_H + \lambda_1^{-\frac{1}{2}}L \int_0^t \|\eta^{h,g}(s)\|_H \, ds. \quad (5.25)
\]
Once again, we use the Gronwall inequality to get
\[
\|\eta^{h,g}(t)\|_H \leq C\|h\|_H \cdot \|g\|_H, \quad t \in [0, T], h, g \in H, \quad (5.26)
\]
which together with (5.11), (5.22) implies
\[
\|D^2\mu(t,x,h,g)\| \leq C\|h\|_H \cdot \|g\|_H, \quad t \in [0, T], x, h, g \in H. \quad (5.27)
\]
This and (5.24) together finish the proof of (5.19). \qed

As opposed to the parabolic case [11, 14, 37], $D\mu(t,y)$ and $D^2\mu(t,y)$ defined as above only admit spatial regularity in the space $H$ and this also remains true for $D\nu(t,y)$ and $D^2\nu(t,y)$. This is due to the lack of the smoothing property of the associated group $E(t)$.

Lemma 5.2. Suppose that $u_0 \in H^1$, $v_0 \in U$ and Assumption $[\Sigma]$ is fulfilled. Then it holds for any $\gamma \in [0, 1)$ and $p \in [2, \infty)$ that
\[
\|\tilde{u}(t) - u_m\|_{L^p((\Omega, H; \gamma)} \leq C\|\gamma^{\min(\gamma, 1)} \quad (5.28)
\]
for $t \in [t_m, t_{m+1}]$, $m = 0, 1, \cdots, M - 1$, and where $\tilde{u}(t)$ is the first component of $X(t)$ given by (2.19) and $u_m$ comes from (1.4).
Proof of Lemma 5.2 Equality (2.19) shows for $t \in [t_m, t_{m+1}]$ that
\[
\tilde{X}(t) - X_m = (E(t - t_m) - \mathcal{F})X_m + E(t - t_m) \left( F(X_m)(t - t_m) + B(W(t) - W(t_m)) \right),
\]
which implies for $t \in [t_m, t_{m+1}]$ that
\[
\bar{u}(t) - u_m = (C(t - t_m) - I)u_m + \Lambda^{-\frac{1}{2}}S(t - t_m)v_m
+ \Lambda^{-\frac{1}{2}}S(t - t_m)F(u_m)(t - t_m) + \Lambda^{-\frac{1}{2}}S(t - t_m)(W(t) - W(t_m)).
\]
Further, using Lemma 3.1, Proposition 3.1, Corollary 3.2 and (3.11) yields for $t \in [t_m, t_{m+1}]$, $p \in [2, \infty)$ and $\gamma \in (0, 1)$ that
\[
\|\bar{u}(t) - u_m\|_{L^p(\Omega, H^{\gamma})} \leq 4\|\Lambda^{-\frac{1}{2}}(C(t - t_m) - I)u_m\|_{L^p(\Omega, U)}^2 + 4\|\Lambda^{-\frac{1}{2}}S(t - t_m)v_m\|_{L^p(\Omega, U)}^2
+ 4(t - t_m)^2 \cdot \|\Lambda^{-\frac{1}{2}}S(t - t_m)F(u_m)(t - t_m)\|_{L^p(\Omega, U)}^2
+ 4C_p^2(t - t_m) \cdot \|\Lambda^{-\frac{1}{2}}S(t - t_m)\|_{L^p(\Omega, U)}^2 \cdot \|\Lambda^{-\frac{1}{2}}\mathcal{Q}(t - t_m)\|_{L^p(\Omega, U)}^2
\leq C\tau^{2\min(\beta, \gamma, 1)} + C\tau^4\|F(u_m)\|_{L^p(\Omega, U)}^2
+ C|t - t_m| \cdot \|\Lambda^{-\frac{1}{2}}S(t - t_m)\|_{L^p(\Omega, U)}^2 \cdot \|\Lambda^{-\frac{1}{2}}\mathcal{Q}(t - t_m)\|_{L^p(\Omega, U)}^2
\leq C\tau^{2\min(\beta, \gamma, 1)}.
\]
The proof of Lemma 5.2 is thus completed. □

Roughly speaking, Lemma 5.2 implies that higher Hölder regularity in time of the numerical approximation process $\bar{u}(t)$ can be achieved when measured in the negative Sobolev space $H^{-\gamma} = \mathcal{D}(\Lambda^{-\frac{1}{2}})$, $\gamma > 0$. This together with the condition (5.3) is the key ingredient in the following weak error estimates and will be exploited to obtain the expected weak convergence rates (see the estimate of $K_m^{2\gamma}$ below). Now we are in a position to start the proof of Theorem 5.7.

Proof of Theorem 5.7 First, we define an auxiliary process $\tilde{Y}(t) = E(T - t)\tilde{X}(t)$, $t \in [0, T]$, by
\[
\tilde{Y}(t) = E(T - t_m)X_m + \int_{t_m}^t E(T - t_m)F(X_m)\, ds + \int_{t_m}^t E(T - t_m)B\, dW(s), \quad t \in [t_m, t_{m+1}].
\]
The above definition of $\tilde{Y}(t)$ allows for
\[
\tilde{Y}(T) = \tilde{X}(T) = X_M \quad \text{and} \quad \tilde{Y}(0) = E(T)X_0.
\]
This together with (5.6) and (5.13) ensures that
\[
\mathbb{E}\left[ \Phi(X_M) \right] - \mathbb{E}\left[ \Phi(X(T)) \right] = \mathbb{E}\left[ \Phi(\tilde{X}(T)) \right] - \mathbb{E}\left[ \mu(T, X_0) \right]
\]
\[
= \mathbb{E}\left[ \Phi(\tilde{Y}(T)) \right] - \mathbb{E}\left[ \mu(T, X_0) \right]
\]
\[
= \mathbb{E}\left[ \nu(0, \tilde{Y}(T)) \right] - \mathbb{E}\left[ \nu(T, E(T)X_0) \right]
\]
\[
= \mathbb{E}\left[ \nu(0, \tilde{Y}(0)) \right] - \mathbb{E}\left[ \nu(T, \tilde{Y}(0)) \right].
\]

Before further analysis, we define a finite dimensional subspace $U_N$ of $U$ by $U_N := \text{span}\{e_1, e_2, \cdots, e_N\}$ for $N \in \mathbb{N}$, and the projection operator $P_N : H^\alpha \rightarrow U_N$ by $P_N \Phi = \sum_{i=1}^N (\Phi, e_i) U e_i$, for $\Phi \in H^\alpha$, $\alpha \geq -1$. Then, we define $\mathcal{D}_N X = (P_N u, P_N v)^T$ for $X = (u, v)^T \in H$. It can be easily verified that $\mathcal{D}_N X \in \mathcal{D}(A)$ for $X \in H$ and
that \( \lim_{N \to \infty} \| \mathcal{P}_N X - X \|_H = 0 \) for all \( X \in H \). At the moment, we apply Itô’s formula \([9] \) Theorem 4.17\) to \( \nu(T-t, \mathcal{P}_N \tilde{Y}(t)) \) in each interval \([t_m, t_{m+1}]\) and utilize (5.17) to obtain

\[
\begin{align*}
\mathbb{E}[\nu(0, \mathcal{P}_N \tilde{Y}(T))] - \mathbb{E}[\nu(T, \mathcal{P}_N \tilde{Y}(0))] &= \sum_{m=0}^{M-1} \left( \mathbb{E}[\nu(T - t_{m+1}, \mathcal{P}_N \tilde{Y}(t_{m+1}))] - \mathbb{E}[\nu(T - t_m, \mathcal{P}_N \tilde{Y}(t_m))] \right) \\
&= \sum_{m=0}^{M-1} \left( -\int_{t_m}^{t_{m+1}} \mathbb{E}\left[ \frac{d\nu}{dt}(T-t, \mathcal{P}_N \tilde{Y}(t)) \right] dt + \frac{1}{2} \mathbb{E}\left[ \int_{t_m}^{t_{m+1}} \text{Tr} \left[ D^2 \nu(T-t, \mathcal{P}_N \tilde{Y}(t)) \mathcal{P}_N E(T-t_m)BQ^{\frac{3}{2}} \left( \mathcal{P}_N E(T-t_m)BQ^{\frac{3}{2}} \right)^* \right] dt \right] \\
&= \sum_{m=0}^{M-1} \left( -\int_{t_m}^{t_{m+1}} \mathbb{E}\left[ \left( \langle D\nu(T-t, \tilde{Y}(t)), E(T-t_m)F(X_m) - E(T-t)F(E(t-T)\tilde{Y}(t)) \rangle \right)_H \right] dt + \frac{1}{2} \mathbb{E}\left[ \int_{t_m}^{t_{m+1}} \text{Tr} \left[ D^2 \nu(T-t, \tilde{Y}(t)) \left( E(T-t_m)BQ^{\frac{3}{2}} \left( E(T-t_m)BQ^{\frac{3}{2}} \right)^* - E(T-t)BQ^{\frac{3}{2}} \left( E(T-t)BQ^{\frac{3}{2}} \right)^* \right) \right] dt \right] \\
&= \sum_{m=0}^{M-1} \left( K^1_m + K^2_m \right). \quad (5.34)
\end{align*}
\]

Note that \( \nu(t, \cdot) \in C^2_b(H; \mathbb{R}) \) for \( t \in [0, T] \) and \( \lim_{N \to \infty} \mathcal{P}_N X = X \) in \( H \). Therefore, taking limits in (5.34) as \( N \to \infty \) shows

\[
\begin{align*}
\mathbb{E}[\nu(0, \tilde{Y}(T))] - \mathbb{E}[\nu(T, \tilde{Y}(0))] &= \sum_{m=0}^{M-1} \left( \int_{t_m}^{t_{m+1}} \mathbb{E}\left[ \left( \langle D\nu(T-t, \tilde{Y}(t)), E(T-t_m)F(X_m) - E(T-t)F(E(t-T)\tilde{Y}(t)) \rangle \right)_H \right] dt + \frac{1}{2} \mathbb{E}\left[ \int_{t_m}^{t_{m+1}} \text{Tr} \left[ D^2 \nu(T-t, \tilde{Y}(t)) \left( E(T-t_m)BQ^{\frac{3}{2}} \left( E(T-t_m)BQ^{\frac{3}{2}} \right)^* - E(T-t)BQ^{\frac{3}{2}} \left( E(T-t)BQ^{\frac{3}{2}} \right)^* \right) \right] dt \right] \\
&= \sum_{m=0}^{M-1} \left( K^1_m + K^2_m \right). \quad (5.35)
\end{align*}
\]

Now, we further decompose \( K^1_m \) and \( K^2_m \) as

\[
\begin{align*}
K^1_m &= \int_{t_m}^{t_{m+1}} \mathbb{E}\left[ \left( \langle D\nu(T-t, \tilde{Y}(t)), E(T-t_m)F(X_m) - E(T-t)F(E(t-T)\tilde{Y}(t)) \rangle \right)_H \right] dt + \int_{t_m}^{t_{m+1}} \mathbb{E}\left[ \left( \langle D\nu(T-t, \tilde{Y}(t)), E(T-t)F(X_m) - F(\tilde{X}(t)) \rangle \right)_H \right] dt \\
&= K^{1,1}_m + K^{1,2}_m \quad (5.36)
\end{align*}
\]
and
\[
K_m^2 = \frac{1}{2} \mathbb{E} \int_{t_m}^{t_{m+1}} \text{Tr} \left[ D^2 \nu(T-t, \tilde{Y}(t))(E(T-t_m)-E(T-t))BQ^{\frac{1}{2}}(E(T-t_m)BQ^{\frac{1}{2}})^* \right] dt \\
+ \frac{1}{2} \mathbb{E} \int_{t_m}^{t_{m+1}} \text{Tr} \left[ D^2 \nu(T-t, \tilde{Y}(t))E(T-t)BQ^2 \left((E(T-t_m)-E(T-t))BQ^2 \right)^* \right] dt \\
= \frac{1}{2} \mathbb{E} \int_{t_m}^{t_{m+1}} \text{Tr} \left[ D^2 \nu(T-t, \tilde{Y}(t))(E(T-t_m)-E(T-t))BQ (E(T-t_m)B)^* \right] dt \\
+ \frac{1}{2} \mathbb{E} \int_{t_m}^{t_{m+1}} \text{Tr} \left[ D^2 \nu(T-t, \tilde{Y}(t))E(T-t)BQ ((E(T-t_m)-E(T-t))B)^* \right] dt \\
= K_m^{1,2} + K_m^{2,2}. 
\] (5.37)

In (5.36), the fact was used that \( E(t-T) \tilde{Y}(t) = \tilde{X}(t) \) by the previous definitions. In the sequel, we shall estimate \( K_m^{1,1}, K_m^{1,2}, K_m^{2,1}, K_m^{2,2} \) separately. With the aid of (5.10), (5.12) and (5.18), we first estimate \( K_m^{1,1} \) as follows
\[
|K_m^{1,1}| \leq \int_{t_m}^{t_{m+1}} \mathbb{E} \left[ \left\| D\nu(T-t, \tilde{Y}(t)) \right\|_{H} \cdot c_1 \tau \left\| F(X_m) \right\|_{H^1} \right] dt \\
\leq C \tau \int_{t_m}^{t_{m+1}} \mathbb{E} \left[ \left\| F(u_m) \right\|_{L^2} \right] dt \\
\leq C \tau^2. 
\] (5.38)

In order to estimate \( K_m^{1,2} \), a common choice is to invoke techniques involved with the Malliavin calculus [1, 3, 4, 14, 37]. Here we will provide an alternative way. To be precise, by means of (5.13), Hölder’s inequality, the trigonometric identity and Taylor’s formula in Banach space we obtain
\[
|K_m^{1,2}| \leq \int_{t_m}^{t_{m+1}} \mathbb{E} \left[ \left\| D\nu(T-t, \tilde{Y}(t)) \right\|_{H} \cdot \left\| E(T-t) \left( F(X_m) - F(\tilde{X}(t)) \right) \right\|_{H} \right] dt \\
\leq C \int_{t_m}^{t_{m+1}} \left( \mathbb{E} \left[ \left\| A^{-\frac{1}{2}} S(T-t) \left( F(u_m) - F(\tilde{u}(t)) \right) \right\|_{L^2} \right] \\
+ \mathbb{E} \left[ \left\| A^{-\frac{1}{2}} C(T-t) \left( F(u_m) - F(\tilde{u}(t)) \right) \right\|_{L^2} \right] \frac{1}{2} \right) dt \\
= C \int_{t_m}^{t_{m+1}} \left\| A^{-\frac{1}{2}} \left( F(\tilde{u}(t)) - F(u_m) \right) \right\|_{L^2(W, L^2)} dt \\
\leq C \int_{t_m}^{t_{m+1}} \left\| A^{-\frac{1}{2}} F'(u_m + r(\tilde{u}(t) - u_m)) \cdot (\tilde{u}(t) - u_m) \right\|_{L^2(W, L^2)} dt. 
\] (5.39)

Proposition 3.1 ensures that, for \( r \in [0, 1], p \in [2, \infty], \kappa = \min(\beta, \frac{1}{2} - \epsilon) \) with arbitrarily small \( \epsilon > 0 \),
\[
\| u_m + r(\tilde{u}(t) - u_m) \|_{L^p(W, H^k)} \leq \sup_{t \in [0, T]} \| \tilde{u}(t) \|_{L^p(W, H^k)} \leq \sup_{t \in [0, T]} \| \tilde{X}(t) \|_{L^p(W, H^k)} \leq C. 
\] (5.40)

Accordingly, by (5.3), (5.28), Hölder’s inequality and also taking the preceding estimate into account, we derive from (5.39) that
\[
|K_m^{1,2}| \leq C \int_{t_m}^{t_{m+1}} \int_0^1 \left\| A^{-\frac{1}{2}} \left( \tilde{u}(t) - u_m \right) - \kappa \cdot \left( \| u_m + r(\tilde{u}(t) - u_m) \|_{L^2(W, \mathbb{R}^d)} \right) \right\|_{L^2(W, L^2)} dt dr \\
\leq C^{1 + \min(\beta + \kappa, 1)} = C^{1 + \min(2\beta, \frac{1}{2} + \beta - \epsilon, 1)} 
\] (5.41)

with \( \kappa = \min(\beta, \frac{1}{2} - \epsilon) \) for arbitrarily small \( \epsilon > 0 \). This together with (5.38) yields
\[
|K_m^{1}| \leq C^{1 + \min(2\beta, \frac{1}{2} + \beta - \epsilon, 1)}. 
\] (5.42)
An exponential integrator scheme for time discretization of nonlinear stochastic wave equation 23

Now we turn to the term $K_m^2$. Concerning $K_m^{2,1}$, we use (2.17), (5.18) and (1.18) to get

$$\left| \text{Tr} \left[ D^2 \nu(T, t, \bar{Y}(t)) \left( E(T - t_m) - E(T - t) \right) BQ \left( E(T - t_m) B \right)^* \right] \right| \leq \left\| D^2 \nu(T, t, \bar{Y}(t)) \left( E(T - t_m) - E(T - t) \right) BQ \left( E(T - t_m) B \right)^* \right\|_{L^1(H)} \leq C \left( t - t_m \right)^{2 \min(2\beta, 1)} \cdot \left\| u \right\|_U,$$

and the trigonometric identity gives

$$\left\| (E(T - t_m) BA^{1/2} u) \right\|_H^2 = \left\| S(T - t_m) u \right\|_U^2 + \left\| C(T - t_m) u \right\|_U^2 = \left\| u \right\|_U^2.$$  

With the aid of these two estimates, we derive from (5.45) that

$$\left| \text{Tr} \left[ D^2 \nu(T, t, \bar{Y}(t)) \left( E(T - t_m) - E(T - t) \right) BQ \left( E(T - t_m) B \right)^* \right] \right| \leq C \left( t - t_m \right)^{2 \min(2\beta, 1)},$$

and hence

$$|K_m^{2,1}| \leq C \tau^{1 + \min(2\beta, 1)},$$

Following the same arguments as used in the estimate of $K_m^{2,1}$ and noticing that $\left\| A^{\frac{1}{2} + \beta} Q A^{\frac{1}{2}} \right\|_{L^1(U)} = \left\| A^{\frac{1}{2} + \beta} \right\|_{L^1(U)}$, one can similarly obtain that

$$|K_m^{2,2}| \leq C \tau^{1 + \min(2\beta, 1)}.$$  

Gathering (5.47) and (5.48) together leads us to

$$|K_m^2| \leq C \tau^{1 + \min(2\beta, 1)}.$$  

Taking (5.42), (5.49) and (5.35) into account, we conclude form (5.33) that

$$\left| \text{E} \left[ \Phi(X_M) \right] - \text{E} \left[ X(T) \right] \right| \leq \sum_{m=0}^{M-1} \left( |K_m^1| + |K_m^2| \right) \leq \sum_{m=0}^{M-1} \left( C \tau^{1 + \min(2\beta, \frac{1}{2} + \beta - \varepsilon, 1)} + C \tau^{1 + \min(2\beta, 1)} \right) \leq C \tau^{\min(2\beta, \frac{1}{2} + \beta - \varepsilon, 1)},$$

which finishes the proof of Theorem 5.1 □

To illustrate the previous assumptions, we give a concrete example as follows.

Example 5.1 Consider a class of SWEs as discussed in Subsection 4.1, where we assign here $d = 1$, $\phi = (0, 1)$, $g(\xi, u) \equiv 1$ for all $\xi \in (0, 1), u \in \mathbb{R}$, and $f: [0, 1] \times \mathbb{R} \to \mathbb{R}$ is assumed to be a smooth nonlinear function satisfying

$$|f(\xi, u)| \leq L |u| + 1, \quad \left| \frac{\partial f(\xi, u)}{\partial u} \right| \leq L, \quad \left| \frac{\partial^2 f(\xi, u)}{\partial^2 u^2} \right| \leq L, \quad \text{and} \quad \left| \frac{\partial^2 f(\xi, u)}{\partial^2 u} \right| \leq L$$  

(5.51)
By (5.56) with
\[ F'(u)(\psi)(\xi) = \frac{\partial F}{\partial u}(\xi, u(\xi)) \cdot \psi(\xi), \quad \xi \in (0, 1), \] (5.52)
\[ F''(u)(\psi_1, \psi_2)(\xi) = \frac{\partial^2 F}{\partial u^2}(\xi, u(\xi)) \cdot \psi_1(\xi) \cdot \psi_2(\xi), \quad \xi \in (0, 1) \] (5.53)
for all \( u, \psi, \psi_1, \psi_2 \in U \). Thanks to Hölder’s inequality and a Sobolev inequality: \( H^\gamma = \mathcal{D}(\Lambda^\frac{\gamma}{2}) \) is continuously embedded into \( L^\infty((0, 1), \mathbb{R}) \) for \( \gamma > \frac{1}{2} \), we get
\[
\| \Lambda^{-\frac{1}{2}} F''(u)(\psi_1, \psi_2) \|_U = \sup_{\|\psi\| \leq 1} \left| \langle \Lambda^{-\frac{1}{2}} F''(u)(\psi_1, \psi_2), \psi \rangle_U \right|
\leq \| F''(u)(\psi_1, \psi_2) \|_{L^2((0, 1), \mathbb{R})} \cdot \sup_{\|\psi\| \leq 1} \| \Lambda^{-\frac{1}{2}} \psi \|_{L^\infty((0, 1), \mathbb{R})}
\leq L \| \psi_1 \|_{L^2((0, 1), \mathbb{R})} \cdot \| \psi_2 \|_{L^2((0, 1), \mathbb{R})} \cdot \frac{C}{\| \psi \|_U}
\leq C \| \psi_1 \|_U \cdot \| \psi_2 \|_U.
\] (5.54)
Now it remains to check (5.3). Since in our example \( F \) is a Nemytskij operator, the derivative operator \( F'(u), u \in U \) is self-adjoint. Therefore, using the self-adjointness of \( \Lambda \) and \( F'(u), u \in U \) yields that
\[
\| \Lambda^{-\frac{1}{2}} F'(u) \|_U = \sup_{\|\psi\| \leq 1} \left| \langle \Lambda^{-\frac{1}{2}} F'(u), \psi \rangle_U \right|
\leq \| F'(u) \|_U \cdot \sup_{\|\psi\| \leq 1} \| \Lambda^{-\frac{1}{2}} \psi \|_U
\leq C [u]_{\gamma} \cdot \| \psi \|_U,
\] (5.55)
with \( \kappa = \min(\beta, \frac{1}{2} - \epsilon) \) for arbitrarily small \( \epsilon > 0 \). Further, the setting in Example 5.1 suffices to ensure
\[
\| F'(u) \|_U \leq C (\| u \|_{\gamma} + 1) \| \psi \|_{\sigma}
\] (5.56)
for any \( u \in H^\gamma, \psi \in H^\sigma \), with arbitrarily \( \gamma \in (0, \frac{1}{2}) \) and \( \sigma \in (\frac{1}{2}, 1) \) (see [38] Lemma 4.4 for more details). By (5.56) with \( \gamma = \kappa = \min(\beta, \frac{1}{2} - \epsilon), \sigma = 1 \) we derive from (5.55) that
\[
\| \Lambda^{-\frac{1}{2}} F'(u) \|_U \leq C \| \lambda^{-\frac{1}{2}} \|_U \cdot \sup_{\|\psi\| \leq 1} \| \psi \|_U
\leq C \| u \|_{\kappa} \cdot (\| u \|_{\kappa} + 1).
\] (5.57)
This completes the verification of the key condition (5.3). \( \square \)

6 Numerical results

In this section, numerical results are included to demonstrate the above assertions. Consider the Sine-Gordon equation subject to noise as follows
\[
\begin{aligned}
\frac{\partial^2 u}{\partial t^2} &= \frac{\partial^2 u}{\partial x^2} - \sin(u) + (\sigma_0 + \sigma_1 u) W, \quad t \in (0, 1), \quad x \in (0, 1), \\
u(0, \xi) &= u_0(\xi), \quad \frac{\partial u}{\partial t} (0, \xi) = v_0(\xi), \quad \xi \in (0, 1), \\
u(t, 0) &= u(t, 0), \quad u(t, 0) = 0, \quad t \in (0, 1),
\end{aligned}
\] (6.1)
where for simplicity we assume that the covariance operator $Q$ has the same eigenfunctions as the Laplacian with Dirichlet boundary condition, i.e., $\phi_i = e^{\sqrt{2} \sin(i \pi \xi)}, i \in \mathbb{N}$. Such equation driven by additive noise ($\sigma_1 = 0$) has been considered as a numerical example in [8,38]. Although only semi-discretization in time has been investigated in this article, spatial discretization needs to be done in order to perform the simulations on the computer. To this end we simply spatially discretize (6.1) via a spectral Galerkin method, with $N = 2^{10}$ fixed (see, for example, [38]). Since the true solutions required in the following are not available, we take numerical solutions produced by the Crank-Nicolson scheme, using very small stepsize $\tau_{\text{exact}} = 2^{-12}$, as reference solutions. Furthermore, the expectations are approximated by averages over 1000 samples in the following numerical tests.

First, let us start with tests on the strong convergence rates and examine the strong approximation error
\[
\left( \mathbb{E} \left[ \left| u(1) - u_M \right|_0^2 \right] \right)^{1/2} = \left( \mathbb{E} \int_0^1 \left| u_{M}(\xi) \right|^2 \, d\xi \right)^{1/2},
\]
which arises due to the linear implicit Euler (LIE), Crank-Nicolson (CN) and exponential Euler (EE) time discretizations. We choose a set of parameters as $\sigma_0 = 0, \sigma_1 = 1, u_0(\xi) = 0, v_0(\xi) = \cos(\xi)$. For the case of space-time white noise ($q_i = 1, i \in \mathbb{N}$) and trace-class noise ($q_i = i^{-1}, i \in \mathbb{N}$), the left and right plots of Fig. 6.1 respectively, depict the strong approximation errors against $1/M$ on a log-log scale, with $M = 2^k, k = 5, 6, 7, 8$. It turns out that the exponential Euler (EE) scheme (1.4) performs better than the linear implicit Euler (LIE) and Crank-Nicolson (CN) time discretizations and exhibits the right strong rates, i.e., order $1/2$ for the space-time white noise and order 1 for the trace-class noise.

![Fig. 6.1](image-url) Strong approximation errors for time-stepping schemes applied to (6.1) with space-time white noise (left) and trace class noise (right).

To illustrate the weak convergence results, we consider the additive space-time white noise and assign $\sigma_0 = 1, \sigma_1 = 0, u_0(\xi) = \cos(\pi(\xi - \frac{1}{2})), v_0(\xi) = 0$. In addition, we choose a particular test function $\varphi(u) = 10 \int_0^1 u(\xi) \sin(\pi \xi) \, d\xi$ and consider the weak errors (5.5) when using time-stepping schemes to approximate the quantity $\mathbb{E}[\varphi(u(1))] = 10 \mathbb{E} \left[ \int_0^1 u(1, \xi) \sin(\pi \xi) \, d\xi \right]$. In Table 6.1 numerical approximations of $\mathbb{E}[\varphi(u(1))] = -4.92169$ by various schemes are presented. Obviously, the scheme (1.4) always gives more reliable approximations than the other two time-stepping schemes. Even for a large step size $\tau = 2^{-3}$, one can observe a good behavior of the scheme (1.4). Also, weak approximation errors of the three schemes are plotted in Fig. 6.2 where one can observe an expected weak rate of 1 for the scheme (1.4). Despite larger weak approximation errors, the other two schemes (LIE and CN) exhibit almost weak order one in this particular example.
Table 6.1 Numerical approximations of $E[\varphi(u(1))] = -4.92169$ by various schemes

| $\tau$   | Linear implicit Euler | Crank-Nicolson | Scheme (1,4) |
|----------|-----------------------|----------------|--------------|
| $2^{-3}$ | -2.77375              | -5.07689       | -4.89865     |
| $2^{-4}$ | -3.62967              | -4.99096       | -4.90763     |
| $2^{-5}$ | -4.21349              | -4.95601       | -4.91620     |
| $2^{-6}$ | -4.55052              | -4.93828       | -4.91888     |

Fig. 6.2 Weak approximation errors for various time-stepping schemes applied to (6.1) with space-time white noise.

Acknowledgment

The author thanks the anonymous referee whose insightful comments and valuable suggestions are crucial to the improvement of the manuscript. The author would like to thank Professor Arnulf Jentzen for his financial support and helpful discussions during the author’s short visit to ETH Zürich in 2013. Thanks also go to Dr. Fengze Jiang for his careful reading the early version of this manuscript. This work was partially supported by National Natural Science Foundations of China under grant numbers 11301550, 11171352, China Postdoctoral Science Foundation under grant numbers 2013M531798, 2014T70779 and Research Foundation of Central South University.

References

1. Andersson, A., Larsson, S.: Weak convergence for a spatial approximation of the nonlinear stochastic heat equation. arXiv preprint [arXiv:1212.5564v3] (2012)
2. Andersson, A., Kruse, R., Larsson, S.: Duality in refined Watanabe-Sobolev spaces and weak approximations of SPDE. arXiv preprint [arXiv:1312.5893v2] (2013)
3. Brehier, C.E.: Approximation of the invariant measure with an Euler scheme for stochastic PDEs driven by space-time white noise. Potential Anal. 40, 1-40 (2014)
4. Brehier, C.E., Kopec M.: Approximation of the invariant law of SPDEs: error analysis using a Poisson equation for a full-discretization scheme. arXiv preprint [arXiv:1311.7030] (2013)
5. Cañada, E.M.: On barrier problems for the vibrating string. Prbab. Theory Rel. Fields 22, 13-24 (1972)
6. Cao, Y., Yin, L.: Spectral Galerkin method for stochastic wave equations driven by space-time white noise. Commun. Pure. Appl. Anal. 6, 607-617 (2007)
7. Cohen, D., Sigg, M.: Convergence analysis of trigonometric methods for stiff second-order stochastic differential equations. Numer. Math. 121, 1-29 (2012)
8. Cohen, D., Larsson, S., Sigg, M.: A trigonometric method for the linear stochastic wave equation. SIAM J. Numer. Anal. 51, 204-222 (2013)
9. Da Prato, G., Zabczyk, J.: Stochastic Equations in Infinite Dimensions. Cambridge University Press, Cambridge (1992)
An exponential integrator scheme for time discretization of nonlinear stochastic wave equation

10. Da Prato, G., Zabczyk, J.: Ergodicity for Infinite Dimensional Systems. Cambridge University Press, Cambridge (1996)
11. Dalang, R.C., Khoshnevisan, D., Mueller, C., Nualart, D., Xiao, Y.: A Minicourse on Stochastic Partial Differential Equations. vol. 1962 of Lecture Notes in Mathematics, Springer-Verlag, Berlin (2009)
12. De Bouard, A., Debussche, A.: Weak and strong order of convergence of a semi discrete scheme for the stochastic nonlinear Schrödinger equation. Appl. Math. Opt. 54, 369-399 (2006)
13. Debussche, A., Printems, J.: Weak order for the discretization of the stochastic heat equation. Math. Comp. 78, 845-863 (2009)
14. Debussche, A.: Weak approximation of stochastic partial differential equations: the nonlinear case. Math. Comp. 80, 89-117 (2011)
15. Hausenblas, E.: Weak approximation of the stochastic wave equation. J. Comput. Appl. Math. 235, 33-58 (2010)
16. Jentzen, A., Kloeden, P.E.: Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive space-time noise. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 465(2102), 649-667 (2009)
17. Jentzen, A., Kloeden, P.E.: The numerical approximation of stochastic partial differential equations. Milan J. Math. 77, 205-244 (2009)
18. Jentzen, A., Kloeden, P.E.: Taylor Approximations for Stochastic Partial Differential Equations. vol. 83 of CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia (2011)
19. Jentzen, A., Kloeden, P.E., Winkel, G.: Efficient simulation of nonlinear parabolic SPDEs with additive noise. Ann. Appl. Probab. 21, 908-950 (2011)
20. Kloeden, P.E., Lord, G.J., Neuenkirch, A., Shardlow T.: The exponential integrator scheme for stochastic partial differential equations: pathwise error bounds. J. Comput. Appl. Math. 235, 1245-1260 (2011)
21. Kloeden, P.E., Platen, E.: Numerical Solution of Stochastic Differential Equations. Springer, Berlin (1992)
22. Kovács, M., Larsson, S., Lindgren, F.: Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise. BIT 52, 85-108 (2012)
23. Kovács, M., Larsson, S., Lindgren, F.: Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise II. Fully discrete schemes. BIT 53, 497-525 (2013)
24. Kovács, M., Larsson, S., Saedpanah, F.: Finite element approximation of the linear stochastic wave equation with additive noise. SIAM J. Numer. Anal. 48, 408-427 (2010)
25. Kruse, R.: Strong and weak approximation of semilinear stochastic evolution equations. PhD thesis, University of Bielefeld (2012)
26. Lindgren, F.: On weak and strong convergence of numerical approximations of stochastic partial differential equations. PhD thesis, Chalmers University of Technology (2012)
27. Lord, G.J., Rougemont, J.: A numerical scheme for stochastic PDEs with Gevrey regularity. IMA J. Numer. Anal. 24, 587-604 (2004)
28. Lord, G.J., Tambue, A.: Stochastic exponential integrators for the finite element discretization of SPDEs for multiplicative and additive noise. IMA J. Numer. Anal. 33, 515-543 (2013)
29. Mao, X.: Stochastic Differential Equations and Applications. Horwood, New York (1997)
30. Martin, A., Prigarin, S.M., Winkler, G.: Exact and fast numerical algorithms for the stochastic wave equation. Int. J. Comput. Math. 80, 1535-1541 (2003)
31. Prévôt, C., Röckner, M.: A Concise Course on Stochastic Partial Differential Equations. vol. 1905 of Lecture Notes in Mathematics, Springer, Berlin (2007)
32. Printems, J.: On the discretization in time of parabolic stochastic partial differential equations, Math. Model. Numer. Anal. 35, 1055-1078 (2001)
33. Qi, R., Yang, X.: Weak convergence of finite element method for stochastic elastic equation driven by additive noise. J. Sci. Comput. 56, 450-470 (2013)
34. Quer-Sardanyons, L., Sanz-Solé, M.: Space semi-discretisations for a stochastic wave equation. Potential Anal. 24, 303-332 (2006)
35. Walsh, J.B.: An Introduction to Stochastic Partial Differential Equations. Springer, Berlin Heidelberg (1986)
36. Walsh, J.B.: On numerical solutions of the stochastic wave equation. Illinois J. Math. 50, 991-1018 (2006)
37. Wang, X., Gan, S.: Weak convergence analysis of the linear implicit Euler method for semilinear stochastic partial differential equations with additive noise. J. Math. Anal. Appl. 398, 151-169 (2013)
38. Wang, X., Gan, S., Tang, J.: Higher order strong approximations of semilinear stochastic wave equation with additive space-time white noise. SIAM J. Sci. Comput. 36(6), A2611-A2632 (2014)