

WEAK CONVERGENCE OF FULLY DISCRETE FINITE ELEMENT APPROXIMATIONS OF SEMILINEAR HYPERBOLIC SPDE WITH ADDITIVE NOISE

MIHÁLY KOVÁCS¹, ANNIKA LANG² AND ANDREAS PETERSSON^{2,*}

Abstract. The numerical approximation of the mild solution to a semilinear stochastic wave equation driven by additive noise is considered. A standard finite element method is employed for the spatial approximation and a rational approximation of the exponential function for the temporal approximation. First, strong convergence of this approximation in both positive and negative order norms is proven. With the help of Malliavin calculus techniques this result is then used to deduce weak convergence rates for the class of twice continuously differentiable test functions with polynomially bounded derivatives. Under appropriate assumptions on the parameters of the equation, the weak rate is found to be essentially twice the strong rate. This extends earlier work by one of the authors to the semilinear setting. Numerical simulations illustrate the theoretical results.

Mathematics Subject Classification. 60H15, 65M12, 60H35, 65C30, 65M60, 60H07.

Received October 29, 2019. Accepted February 17, 2020.

1. INTRODUCTION

The stochastic wave equation is an evolutionary equation that can be used to model various time dependent phenomena influenced by random forces. One example (see [13]) is the vertical displacement $u : [0, T] \times \mathcal{D} \rightarrow \mathbb{R}$ of a DNA string suspended in a liquid,

$$du(t) - \Delta u(t) dt = -Qu(t) dt + dW(t) \quad (1.1)$$

for $t \in (0, T]$, $T < \infty$, where Δ is the Laplacian with suitable boundary conditions on a convex domain $\mathcal{D} \subset \mathbb{R}^d$, $d = 1, 2, 3$. The first term on the right hand side of (1.1) models friction due to viscosity of the fluid, while the Gaussian noise term $dW(t)$ corresponds to random bombardment of the DNA string by the fluid's molecules. This noise is white in time with spatial correlation described by the linear operator Q on $L^2(\mathcal{D}) = L^2(\mathcal{D}, \mathbb{R})$, the same operator as in the friction term. Thus (1.1) can be treated as a stochastic partial differential equation in the Itô sense, driven by a Wiener process W in $L^2(\mathcal{D})$.

Keywords and phrases. Stochastic partial differential equations, stochastic wave equations, stochastic hyperbolic equations, weak convergence, finite element methods, Galerkin methods, rational approximations of semigroups, Crank–Nicolson method, Malliavin calculus.

¹ Faculty of Information Technology and Bionics, Pázmány Péter Catholic University, PO Box 278, H-1444 Budapest, Hungary.

² Department of Mathematical Sciences, Chalmers University of Technology & University of Gothenburg, S-412 96 Gothenburg, Sweden.

*Corresponding author: andreas.petersson@chalmers.se

In this paper, we are concerned with the more general setting that the friction due to viscosity may depend non-linearly on the displacement of the DNA string and that the intensity of the molecular bombardment may vary in time. We thus consider the equation

$$d\dot{u}(t) - \Delta u(t) dt = F(t, u(t)) dt + G(t) dW(t), \quad (1.2)$$

with the goal of analyzing errors stemming from the approximation of this equation by finite elements and a rational approximation of the exponential function. The Laplacian Δ is assumed to satisfy zero Dirichlet boundary conditions, *i.e.*, $u(t) = 0$ on $\partial\mathcal{D}$ for all times $t \in (0, T]$, and the equation has initial conditions $u(0) = u_0$ and $\dot{u}(0) = v_0$.

In general, (1.2) cannot be solved analytically. The question of how to find an approximation \hat{u} of u and how to evaluate the quality of such an approximation *a priori* is therefore of great importance if one wants to use this equation in practice. In order to implement an approximation on a computer, the equation is typically discretized both in the spatial and temporal domain, in which case the resulting approximation \hat{u} is said to be fully discrete. In the literature, the quality of \hat{u} is in general evaluated by analyzing the rate of decay of the strong error $\mathbb{E}[\|u - \hat{u}\|_{L^2(\mathcal{D})}^2]^{1/2}$ (see [3, 7–10, 17–20, 25, 27–29]). Comparatively few results (see [10, 16–20, 28]) exist on the rate for the weak error $|\mathbb{E}[\phi(u) - \phi(\hat{u})]|$, where $\phi: L^2(\mathcal{D}) \rightarrow \mathbb{R}$ is a sufficiently smooth real-valued test function. Of the results cited, only [16] provides a weak convergence result for a fully discrete approximation of a semilinear stochastic wave equation. If ϕ is (locally) Lipschitz, the weak error can be bounded by the strong error, but in analysis the rate of decay of the weak error as one considers finer and finer approximations is often found to be twice the rate of the strong error.

The outline of our paper is the following. In Section 2, we analyze (1.2) in a more general, abstract, Hilbert space setting and show spatial and temporal regularity results under mild assumptions on F .

In Section 3 we deduce strong and weak error rates for the approximation of the so called mild solution u of (1.2) by means of a finite element approximation (by piecewise linear or quadratic functions) in space and a rational approximation of the exponential function in time, generalizing the result of [19] to the semilinear setting. This approach sets the paper apart from several recent works (*e.g.*, [3, 8–10, 28, 29]) on the stochastic wave equation that consider trigonometric integrators for the temporal approximation. There are situations when such integrators could be better suited such as highly oscillatory data but for complicated domain geometries the algorithms in the present article could be more advantageous from an implementation point of view, since they do not require any knowledge of the eigenfunctions of Δ or its discrete counterpart.

For the analysis we take a similar approach as the author of [28], by using negative norm strong convergence rates in our analysis of the weak error. However, instead of using Kolmogorov's equation and the Itô formula, we complete the analysis by means of Malliavin calculus. Our results are applicable under slightly more general assumptions on $F(t, \cdot)$ compared to [28], specifically when $F(t, \cdot)$ is a Nemytskij operator, *i.e.*, when $F(t, u)(x) = f(t, u(x))$ for $u \in L^2(\mathcal{D})$ and almost every $x \in \mathcal{D}$. Here $f(t, \cdot)$ is a real-valued function of at most linear growth, with bounded and Lipschitz-continuous first derivative. The test function ϕ is assumed to be twice Gâteaux differentiable with polynomially bounded derivatives.

Section 4 finishes the main part of the paper with examples in which it is noted that when F is a sufficiently smooth Nemytskij operator, the derived weak convergence rates are essentially twice as big as the strong convergence rates, provided that the initial value is smooth, for $d = 1, 2$ when the covariance operator Q of W is of trace-class, and $d = 1$ when $Q = I$. Numerical simulations in $d = 1$ illustrate our theoretical results.

In Appendix A, which completes the paper, it is shown that a sufficiently smooth Nemytskij operator fulfills the assumptions of Section 3.

Throughout the paper, we adopt the notion of generic constants, which is to say that the symbol C is used to denote a positive and finite number which may vary from occurrence to occurrence and is independent of any parameter of interest, such as spatial and temporal step sizes in a numerical method. We use the expression $a \lesssim b$ to denote the existence of a generic constant C such that $a \leq Cb$.

2. THE STOCHASTIC WAVE EQUATION

In this section the stochastic wave equation is presented along with necessary background material from probability theory and functional analysis. We use the semigroup approach of [12] and refer to this monograph for more details on the material covered here. The equation is treated in an abstract Hilbert space setting, while in the next section we restrict ourselves to the setting in which the solution takes values in the Hilbert space $L^2(\mathcal{D})$ where $\mathcal{D} \subset \mathbb{R}^d$, $d = 1, 2, 3$, denotes the underlying domain.

Let $(H, \langle \cdot, \cdot \rangle_H, \|\cdot\|_H)$ and $(U, \langle \cdot, \cdot \rangle_U, \|\cdot\|_U)$ be real separable Hilbert spaces. We denote by $(\mathcal{L}(H, U), \|\cdot\|_{\mathcal{L}(H, U)})$ the space of bounded linear operators from H to U equipped with the usual operator norm and by $(\mathcal{L}_1(H, U), \|\cdot\|_{\mathcal{L}_1(H, U)})$ and $(\mathcal{L}_2(H, U), \langle \cdot, \cdot \rangle_{\mathcal{L}_2(H, U)}, \|\cdot\|_{\mathcal{L}_2(H, U)})$ the subsets of trace-class and Hilbert–Schmidt operators, respectively. We use the shorthand notations $\mathcal{L}(H) = \mathcal{L}(H, H)$, $\mathcal{L}_1(H) = \mathcal{L}_1(H, H)$ and $\mathcal{L}_2(H) = \mathcal{L}_2(H, H)$. Note that if $(V, \langle \cdot, \cdot \rangle_V, \|\cdot\|_V)$ is another Hilbert space and if $\Gamma_2 \in \mathcal{L}(U, V)$, $\Gamma_1 \in \mathcal{L}_i(H, U)$, $i \in \{1, 2\}$, then $\Gamma_2\Gamma_1 \in \mathcal{L}_i(H, V)$ and

$$\|\Gamma_2\Gamma_1\|_{\mathcal{L}_i(H, V)} \leq \|\Gamma_2\|_{\mathcal{L}(U, V)} \|\Gamma_1\|_{\mathcal{L}_i(H, U)}. \quad (2.1)$$

Similarly, if $\Gamma_1 \in \mathcal{L}_i(V, H)$, $i \in \{1, 2\}$, then $\Gamma_1\Gamma_2 \in \mathcal{L}_i(U, H)$ and

$$\|\Gamma_1\Gamma_2\|_{\mathcal{L}_i(U, H)} \leq \|\Gamma_1\|_{\mathcal{L}_i(V, H)} \|\Gamma_2\|_{\mathcal{L}(U, V)}. \quad (2.2)$$

The trace of $\Gamma \in \mathcal{L}_1(H)$ is, for an orthonormal basis $(e_j)_{j=1}^\infty$ of H , defined by $\text{Tr}(\Gamma) = \sum_{j=1}^\infty \langle \Gamma e_j, e_j \rangle_H$ and is independent of the choice of basis. If $\Gamma_1 \in \mathcal{L}_1(H, U)$ and $\Gamma_2 \in \mathcal{L}(U, H)$ then

$$\text{Tr}(\Gamma_1\Gamma_2) = \text{Tr}(\Gamma_2\Gamma_1). \quad (2.3)$$

We will have reason to use spaces of Gâteaux differentiable mappings, which we define in the same way as the authors of [2]. By $\mathcal{C}(H, U)$ we denote the space of continuous mappings from H to U and by $\mathcal{G}^1(H, U) \subset \mathcal{C}(H, U)$ the space of Gâteaux differentiable mappings with strongly continuous derivatives, *i.e.*, the space of all continuous mappings $\varphi: H \rightarrow U$ such that

$$\varphi'(u)v = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (\varphi(u + \epsilon v) - \varphi(u))$$

exists as a limit in U for all $u, v \in H$, that $\varphi'(u) \in \mathcal{L}(H, U)$ for all $u \in H$ and that the mapping $H \ni u \mapsto \varphi'(u)v$ is continuous for all $v \in H$. If in addition $\varphi' \in \mathcal{C}(H, \mathcal{L}(H, U))$, then $\varphi \in \mathcal{C}^1(H, U)$, the space of Fréchet differentiable mappings. By $\mathcal{G}^2(H, U) \subset \mathcal{G}^1(H, U)$ we denote the space of all mappings $\varphi \in \mathcal{G}^1(H, U)$ such that

$$\varphi''(u)(v, w) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (\varphi'(u + \epsilon w)v - \varphi'(u)v)$$

exists as a limit in U for all $u, v, w \in H$, that $\varphi''(u): H \times H \rightarrow U \in \mathcal{L}^{[2]}(H, U)$, the space of all bounded bilinear mappings, for all $u \in H$, that $\varphi''(u)$ is symmetric for all $u \in H$, and that the mapping $H \ni u \mapsto \varphi''(u)(v, w)$ is continuous for all $v, w \in H$. For $n = 1, 2$, we denote by $\mathcal{G}_b^n(H, U)$ and $\mathcal{G}_p^n(H, U)$ the sets of all $\varphi \in \mathcal{G}^n(H, U)$ such that all derivatives of φ (but not necessarily φ itself) are bounded and polynomially bounded, respectively, with $\mathcal{C}_b^1(H, U)$ and $\mathcal{C}_p^1(H, U)$ defined analogously. We use the shorthand notations $\mathcal{G}^n(H) = \mathcal{G}^n(H, H)$, $\mathcal{G}_b^n(H) = \mathcal{G}_b^n(H, H)$ and $\mathcal{G}_p^n(H) = \mathcal{G}_p^n(H, H)$, and similarly for the spaces of Fréchet differentiable mappings. For $\varphi \in \mathcal{G}_p^1(H, U)$ and $u, v \in H$ the mean value theorem holds in U , *i.e.*,

$$\varphi(u) - \varphi(v) = \int_0^1 \varphi'(v + s(u - v))(u - v) ds.$$

For $0 < T < \infty$, let $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t \in [0, T]}, P)$ be a complete filtered probability space satisfying the usual conditions, which is to say that \mathcal{F}_0 contains all P -null sets and $\mathcal{F}_t = \cap_{s > t} \mathcal{F}_s$ for all $t \in [0, T]$. By $L^p(\Omega, H)$, $p \in [1, \infty)$, we denote the space of all H -valued random variables X with norm $\|X\|_{L^p(\Omega, H)} = (\mathbb{E}[\|X\|_H^p])^{1/p}$.

Let $W: \Omega \times [0, T] \rightarrow H$ be a Wiener process with a covariance operator $Q \in \mathcal{L}(H)$ that is positive semidefinite and self-adjoint, but not necessarily of trace-class.

As is usual in this setting, we write $H_0 = Q^{\frac{1}{2}}(H)$, which is a Hilbert space when equipped with the inner product $\langle \cdot, \cdot \rangle_{H_0} = \left\langle Q^{-\frac{1}{2}} \cdot, Q^{-\frac{1}{2}} \cdot \right\rangle_H$, where $Q^{-\frac{1}{2}}$ is the pseudo-inverse of $Q^{\frac{1}{2}}$. Note that for $\Gamma_1, \Gamma_2 \in \mathcal{L}_2^0 = \mathcal{L}_2(H_0, H)$,

$$\left| \langle \Gamma_1, \Gamma_2 \rangle_{\mathcal{L}_2^0} \right| = |\text{Tr}(\Gamma_1 Q \Gamma_2^*)| \leq \|\Gamma_1 Q \Gamma_2^*\|_{\text{Tr}} \quad (2.4)$$

whenever the right hand side is finite. Here and below the shorthand notation $\|\cdot\|_{\text{Tr}} = \|\cdot\|_{\mathcal{L}_1(H)}$ is used. The Wiener process allows us to handle Itô integrals $\int_0^T \Phi(t) dW(t)$, for predictable stochastic processes $\Phi: [0, T] \times \Omega \rightarrow \mathcal{L}_2^0$. The following Burkholder–Davis–Gundy type inequality turns out to be useful.

Lemma 2.1 ([11], Lem. 7.2). *For any $p \in [1, \infty)$, there exists a constant $C > 0$, such that for any predictable stochastic process $\Phi: [0, T] \times \Omega \rightarrow \mathcal{L}_2^0$ with $\|\Phi\|_{L^p(\Omega, L^2([0, T], \mathcal{L}_2^0))} < \infty$,*

$$\left\| \int_0^T \Phi(t) dW(t) \right\|_{L^p(\Omega, H)} \leq C \|\Phi\|_{L^p(\Omega, L^2([0, T], \mathcal{L}_2^0))}.$$

We are now ready to introduce the equation studied in this paper,

$$\begin{cases} d\dot{u}(t) + \Lambda u(t) dt = F(t, u(t)) dt + G(t) dW(t), & t \in (0, T], \\ u(0) = u_0, \dot{u}(0) = v_0. \end{cases} \quad (2.5)$$

Here the solution process u and the Wiener process W take values in the Hilbert space H , \dot{u} denotes the time derivative of u and F and G are deterministic mappings. The operator Λ is a densely defined, linear, unbounded positive self-adjoint operator with compact inverse, implying that it has an orthonormal eigenbasis $(e_j)_{j=1}^\infty$ spanning H with an increasing sequence $(\lambda_j)_{j=1}^\infty$ of strictly positive eigenvalues, which are used to define fractional powers $\Lambda^{\frac{\alpha}{2}}$, $\alpha \in \mathbb{R}$ (see [21], Appendix B). We adopt the notation \dot{H}^α for the Hilbert space $D(\Lambda^{\frac{\alpha}{2}})$ and remark that $\dot{H}^{-\alpha} \simeq (\dot{H}^\alpha)^*$ for $\alpha \geq 0$, where $(\dot{H}^\alpha)^*$ is the dual of \dot{H}^α with respect to $\langle \cdot, \cdot \rangle_{\dot{H}^0}$ and \dot{H}^0 is identified with $(\dot{H}^0)^*$ by the Riesz representation theorem. We have that $\dot{H}^0 = H$ and that $\dot{H}^\zeta \subset \dot{H}^\alpha$ for $\alpha \leq \zeta \in \mathbb{R}$, where the embedding is dense and continuous. By Lemma 2.1 of [6], for every $\zeta \in \mathbb{R}$, $\Lambda^{\frac{\alpha}{2}}$ can be uniquely extended to an operator in $\mathcal{L}(\dot{H}^\zeta, \dot{H}^{\zeta-\alpha})$. We make no notational distinction between $\Lambda^{\frac{\alpha}{2}}$ and its extension.

In order to treat (2.5) in a semigroup framework, we define for $\alpha \in \mathbb{R}$ the Hilbert space $\mathcal{H}^\alpha = \dot{H}^\alpha \oplus \dot{H}^{\alpha-1}$ with inner product $\langle v, w \rangle_{\mathcal{H}^\alpha} = \langle v_1, w_1 \rangle_{\dot{H}^\alpha} + \langle v_2, w_2 \rangle_{\dot{H}^{\alpha-1}}$ for $v = [v_1, v_2]^\top, w = [w_1, w_2]^\top \in \mathcal{H}^\alpha$. Writing $\mathcal{H} = \mathcal{H}^0$, let $A: D(A) = \mathcal{H}^1 \rightarrow \mathcal{H}$, $B: \dot{H}^{-1} \rightarrow \mathcal{H}$ and $\Theta^{\frac{\alpha}{2}}: \mathcal{H}^\alpha \rightarrow \mathcal{H}, \alpha \in \mathbb{R}$, be given by

$$A = \begin{bmatrix} 0 & -I \\ \Lambda & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ I \end{bmatrix} \quad \text{and} \quad \Theta^{\frac{\alpha}{2}} = \begin{bmatrix} \Lambda^{\frac{\alpha}{2}} & 0 \\ 0 & \Lambda^{\frac{\alpha}{2}} \end{bmatrix}.$$

The third operator is used to relate the norms of \mathcal{H}^α and \mathcal{H} via $\|\cdot\|_{\mathcal{H}^\alpha} = \|\Theta^{\frac{\alpha}{2}} \cdot\|_{\mathcal{H}}$. We also consider P^1 , the projection onto the first coordinate of \mathcal{H} , i.e., $P^1 v = v_1$ for $v = [v_1, v_2]^\top \in \mathcal{H}$. Note that $\Theta^{\frac{\alpha}{2}} B = B \Lambda^{\frac{\alpha}{2}}$ and that therefore, the identities

$$\|\Theta^{\frac{\alpha}{2}} B v\|_{\mathcal{H}} = \|\Lambda^{\frac{\alpha-1}{2}} v\|_{\dot{H}^0} = \|v\|_{\dot{H}^{\alpha-1}}, \quad (2.6)$$

with $v \in \dot{H}^{\alpha-1}$, and

$$\|B\|_{\mathcal{L}(\dot{H}^{-1}, \mathcal{H})} = \|B \Lambda^{\frac{1}{2}}\|_{\mathcal{L}(\dot{H}^0, \mathcal{H})} = \|\Theta^{\frac{1}{2}} B\|_{\mathcal{L}(\dot{H}^0, \mathcal{H})} = 1, \quad (2.7)$$

hold.

The operator $-A$ is the generator of a C_0 -semigroup (actually a group, see [22]) which, for $t \in \mathbb{R}$, can be written as

$$E(t) = \begin{bmatrix} C(t) & \Lambda^{-\frac{1}{2}} S(t) \\ -\Lambda^{\frac{1}{2}} S(t) & C(t) \end{bmatrix} = \begin{bmatrix} \cos(t\Lambda^{\frac{1}{2}}) & \Lambda^{-\frac{1}{2}} \sin(t\Lambda^{\frac{1}{2}}) \\ -\Lambda^{\frac{1}{2}} \sin(t\Lambda^{\frac{1}{2}}) & \cos(t\Lambda^{\frac{1}{2}}) \end{bmatrix}.$$

It fulfills

$$\|E(t)\|_{\mathcal{L}(\mathcal{H})} \leq 1 \quad (2.8)$$

uniformly in $t \in \mathbb{R}$. We note the commutative properties, with $\alpha \in \mathbb{R}$,

$$\Theta^{\frac{\alpha}{2}} E(\cdot) = E(\cdot) \Theta^{\frac{\alpha}{2}}$$

and

$$\Lambda^{\frac{\alpha}{2}} P^1 E(\cdot) B = P^1 E(\cdot) B \Lambda^{\frac{\alpha}{2}}, \quad (2.9)$$

so that, for all $\alpha \in [0, 1]$ there exists by Lemma 4.4 of [19] a constant $C > 0$ such that, for all $t, s \in [0, T]$,

$$\|\Theta^{-\frac{\alpha}{2}} (E(t) - E(s))\|_{\mathcal{L}(\mathcal{H})} = \|(E(t) - E(s)) \Theta^{-\frac{\alpha}{2}}\|_{\mathcal{L}(\mathcal{H})} \leq C|t - s|^\alpha \quad (2.10)$$

and by (2.7) and an argument similar to (4.1) of [9], we have

$$\|\Lambda^{\frac{\alpha}{2}} P^1 E(t - s) B\|_{\mathcal{L}(\dot{H}^0)} = \|P^1 E(t - s) B \Lambda^{\frac{\alpha}{2}}\|_{\mathcal{L}(\dot{H}^0)} = \|S(t - s) \Lambda^{\frac{\alpha-1}{2}}\|_{\mathcal{L}(\dot{H}^0)} \leq C|t - s|^{1-\alpha}. \quad (2.11)$$

If we write $X(t) = [X_1(t), X_2(t)]^\top = [u(t), \dot{u}(t)]^\top$ for $t \in [0, T]$, then (2.5) can be written in the abstract Itô form

$$dX(t) + AX(t) dt = BF(t, X_1(t)) dt + BG(t) dW(t) \quad (2.12)$$

with initial condition $X(0) = x_0 = [u_0, v_0]^\top$. Under the following assumption, (2.12) has a mild solution given by

$$X(t) = E(t)x_0 + \int_0^t E(t-s)BF(s, X_1(s)) ds + \int_0^t E(t-s)BG(s) dW(s) \quad (2.13)$$

for $t \in [0, T]$, the existence of which we show below.

Assumption 2.2. *There exist parameters $\beta, \eta, \delta \geq 0$ and $\theta \leq \min(\beta, \delta, 1)$ and a constant $C > 0$ such that the data in (2.12) fulfills the following requirements.*

(i) *The mapping $G: [0, T] \rightarrow \mathcal{L}_2(H_0, \dot{H}^{\beta-1})$ satisfies*

$$\|\Lambda^{\frac{\beta-1}{2}} (G(t_1) - G(t_2))\|_{\mathcal{L}_2^0} \leq C|t_1 - t_2|^\eta$$

for all $t_1, t_2 \in [0, T]$ and $\|\Lambda^{\frac{\beta-1}{2}} G(t)\|_{\mathcal{L}_2^0} \leq C$ for some $t \in [0, T]$.

(ii) *The function $F: [0, T] \times \dot{H}^0 \rightarrow \dot{H}^0$ satisfies*

$$\|\Lambda^{-\frac{1}{2}} (F(t, u) - F(t, v))\|_{\dot{H}^0} \leq C\|u - v\|_{\dot{H}^0}$$

for all $t \in [0, T]$ and $u, v \in \dot{H}^0$,

$$\|\Lambda^{\frac{\alpha}{2}} F(t, u)\|_{\dot{H}^0} \leq C(1 + \|\Lambda^{\frac{\alpha}{2}} u\|_{\dot{H}^0}),$$

for all $t \in [0, T]$, $u \in \dot{H}^\alpha$ and $\alpha \in \{0, \theta\}$ and

$$\|\Lambda^{-\frac{1}{2}} (F(t, u) - F(s, u))\|_{\dot{H}^0} \leq C(1 + \|u\|_{\dot{H}^0})|t - s|^\eta$$

for all $s, t \in [0, T]$ and $u \in \dot{H}^0$.

(iii) *The initial value $x_0 \in \mathcal{H}^\delta$ is deterministic.*

The following theorem is very similar to, e.g., [28], but since the mappings F and G depend on t , and the assumptions on F are slightly different than those in [28], we include a proof of our own.

Theorem 2.3. *Let Assumption 2.2 be satisfied. Then (2.12) has a unique mild solution given by (2.13) and for any $r \leq \min(\beta, \delta, 1 + \theta)$, $p \in [1, \infty)$,*

$$\sup_{t \in [0, T]} \|X(t)\|_{L^p(\Omega, \mathcal{H}^r)} < \infty. \quad (2.14)$$

Proof. Let $t \in [0, T]$ be fixed. Using the fact that $\Lambda^{-\alpha} \in \mathcal{L}(\dot{H}^0)$ for any $\alpha \geq 0$, we have by Assumption 2.2(ii) and (2.6) that for any $x = [x_1, x_2]^\top, y = [y_1, y_2]^\top \in \mathcal{H}$,

$$\|BF(t, x_1) - BF(t, y_1)\|_{\mathcal{H}} = \|\Lambda^{-\frac{1}{2}}(F(t, x_1) - F(t, y_1))\|_{\dot{H}^0} \lesssim \|x - y\|_{\mathcal{H}}.$$

Similarly, recalling also (2.1),

$$\begin{aligned} \|BF(t, x_1)\|_{\mathcal{H}} + \|BG(t)\|_{\mathcal{L}_2(H_0, \mathcal{H})} &= \|\Lambda^{-\frac{1}{2}}F(t, x_1)\|_{\dot{H}^0} + \|\Lambda^{-\frac{1}{2}}G(t)\|_{\mathcal{L}_2^0} \\ &\lesssim \|F(t, x_1)\|_{\dot{H}^0} + \|\Lambda^{\frac{\beta-1}{2}}G(t)\|_{\mathcal{L}_2^0} \lesssim 1 + \|x_1\|_{\dot{H}^0} \leq 1 + \|x\|_{\mathcal{H}}. \end{aligned}$$

The existence and uniqueness of the mild solution (2.13) now follows from Theorem 7.2 of [12] (for $p \geq 2$ and clearly also for $p \in [1, 2)$ since $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t \in [0, T]}, P)$ is a probability space), which also guarantees that (2.14) holds for $r = 0$. The case $r < 0$ follows immediately. To show (2.14) for $0 < r \leq \min(\beta, \delta, 1)$, we first note that

$$\begin{aligned} \|X(t)\|_{L^p(\Omega, \mathcal{H}^r)} &= \|\Theta^{\frac{r}{2}}X(t)\|_{L^p(\Omega, \mathcal{H})} \\ &\leq \|E(t)\Theta^{\frac{r}{2}}x_0\|_{\mathcal{H}} + \int_0^t \|E(t-s)\Theta^{\frac{r}{2}}BF(s, X_1(s))\|_{L^p(\Omega, \mathcal{H})} ds \\ &\quad + \left\| \int_0^t E(t-s)\Theta^{\frac{r}{2}}BG(s) dW(s) \right\|_{L^p(\Omega, \mathcal{H})}. \end{aligned}$$

For the first term, (2.8) and Assumption 2.2(iii) imply

$$\|E(t)\Theta^{\frac{r}{2}}x_0\|_{\mathcal{H}} \leq \|x_0\|_{\mathcal{H}^r} \lesssim 1.$$

Next, we first note that since $r \leq 1$, by (2.8), (2.6), Assumption 2.2(ii) and (2.14) with $r = 0$,

$$\begin{aligned} \int_0^t \|E(t-s)\Theta^{\frac{r}{2}}BF(s, X_1(s))\|_{L^p(\Omega, \mathcal{H})} ds &\leq t \sup_{s \in [0, T]} \|\Theta^{\frac{r}{2}}BF(s, X_1(s))\|_{L^p(\Omega, \mathcal{H})} \\ &= t \sup_{s \in [0, T]} \|\Lambda^{\frac{r-1}{2}}F(s, X_1(s))\|_{L^p(\Omega, \dot{H}^0)} \\ &\lesssim \sup_{s \in [0, T]} \|F(s, X_1(s))\|_{L^p(\Omega, \dot{H}^0)} \\ &\lesssim 1 + \sup_{s \in [0, T]} \|X_1(s)\|_{L^p(\Omega, \dot{H}^0)} \\ &\leq 1 + \sup_{s \in [0, T]} \|X(s)\|_{L^p(\Omega, \mathcal{H})} \lesssim 1. \end{aligned} \quad (2.15)$$

For the third term, by analogous arguments, Assumption 2.2(i) and Lemma 2.1 (note that the integrand below is deterministic),

$$\begin{aligned} \left\| \int_0^t E(t-s)\Theta^{\frac{r}{2}}BG(s) dW(s) \right\|_{L^p(\Omega, \mathcal{H})} &\lesssim \left(\int_0^t \|E(t-s)\Theta^{\frac{r}{2}}BG(s)\|_{\mathcal{L}_2(H_0, \mathcal{H})}^2 ds \right)^{\frac{1}{2}} \\ &\lesssim t^{\frac{1}{2}} \sup_{s \in [0, T]} \|E(t-s)\Theta^{\frac{r}{2}}BG(s)\|_{\mathcal{L}_2(H_0, \mathcal{H})} \lesssim \sup_{s \in [0, T]} \|\Lambda^{\frac{r-1}{2}}G(s)\|_{\mathcal{L}_2^0} \lesssim 1. \end{aligned}$$

Altogether, this shows (2.14) for $0 \leq r \leq \min(\beta, \delta, 1)$. Finally, for the case $r \in (1, \min(\beta, \delta, 1 + \theta)]$ we repeat the arguments above, replacing the calculation in (2.15) with

$$\begin{aligned} \sup_{s \in [0, T]} \|\Lambda^{\frac{r-1}{2}} F(s, X_1(s))\|_{L^p(\Omega, \dot{H}^0)} &\lesssim \sup_{s \in [0, T]} \|\Lambda^{\frac{\theta}{2}} F(s, X_1(s))\|_{L^p(\Omega, \dot{H}^0)} \\ &\lesssim 1 + \sup_{s \in [0, T]} \|\Lambda^{\frac{\theta}{2}} X_1(s)\|_{L^p(\Omega, \dot{H}^0)} \\ &\leq 1 + \sup_{s \in [0, T]} \|\Theta^{\frac{\min(\beta, \delta, 1)}{2}} X(s)\|_{L^p(\Omega, \mathcal{H})}, \end{aligned}$$

which is finite since we have shown that (2.14) holds with $r = \min(\beta, \delta, 1)$ and by assumption $\theta \leq \min(\beta, \delta, 1)$. \square

From here on we denote by $r = \min(\beta, \delta, 1 + \theta)$ the maximum spatial regularity of the solution to (2.12). A temporal regularity result finishes this section of the paper.

Theorem 2.4. *Let Assumption 2.2 be satisfied and let $r = \min(\beta, \delta, 1 + \theta)$. Then, for all $\alpha \leq r$, $p \geq 1$, there exists a positive constant C such that for all $s, t \in [0, T]$,*

$$\|X(t) - X(s)\|_{L^p(\Omega, \mathcal{H}^\alpha)} \leq C|t - s|^{\min(r-\alpha, \frac{1}{2})}. \quad (2.16)$$

and

$$\|X_1(t) - X_1(s)\|_{L^p(\Omega, \dot{H}^\alpha)} \leq C|t - s|^{\min(r-\alpha, 1)}. \quad (2.17)$$

Proof. Fix $0 \leq s \leq t \leq T$. We first note that

$$\Theta^{\frac{\alpha}{2}}(X(t) - X(s)) = \Theta^{\frac{\alpha}{2}}(E(t-s) - I)X(s) + \int_s^t \Theta^{\frac{\alpha}{2}} E(t-r)BF(r, X_1(r))dr + \int_s^t \Theta^{\frac{\alpha}{2}} E(t-r)BG(r)dW(r),$$

so that therefore

$$\begin{aligned} \|X(t) - X(s)\|_{L^p(\Omega, \mathcal{H}^\alpha)} &= \|\Theta^{\frac{\alpha}{2}}(X(t) - X(s))\|_{L^p(\Omega, \mathcal{H})} \\ &\leq \|\Theta^{\frac{\alpha}{2}}(E(t-s) - I)X(s)\|_{L^p(\Omega, \mathcal{H})} + \left\| \int_s^t \Theta^{\frac{\alpha}{2}} E(t-r)BF(r, X_1(r))dr \right\|_{L^p(\Omega, \mathcal{H})} \\ &\quad + \left\| \int_s^t \Theta^{\frac{\alpha}{2}} E(t-r)BG(r)dW(r) \right\|_{L^p(\Omega, \mathcal{H})}. \end{aligned} \quad (2.18)$$

By (2.10) and Theorem 2.3 the first term on the right hand side of (2.18) is bounded by

$$\|\Theta^{\frac{\alpha-r}{2}}(E(t-s) - I)\|_{\mathcal{L}(\mathcal{H})} \|\Theta^{\frac{r}{2}} X(s)\|_{L^p(\Omega, \mathcal{H})} \lesssim (t-s)^{\min(r-\alpha, 1)}.$$

For the second term, we have, since $\alpha \leq r \leq 1 + \theta$, by (2.6), (2.8), Assumption 2.2(ii) and Theorem 2.3,

$$\begin{aligned} \left\| \int_s^t \Theta^{\frac{\alpha}{2}} E(t-r)BF(r, X_1(r))dr \right\|_{L^p(\Omega, \mathcal{H})} &\leq \int_s^t \|E(t-r)\Theta^{\frac{\alpha}{2}} BF(r, X_1(r))\|_{L^p(\Omega, \mathcal{H})} dr \\ &\leq \int_s^t \|E(t-r)\|_{\mathcal{L}(\mathcal{H})} \|\Lambda^{\frac{\alpha-1}{2}} F(r, X_1(r))\|_{L^p(\Omega, \dot{H}^0)} dr \\ &\lesssim (t-s) \sup_{r \in [0, T]} \|\Lambda^{\frac{\theta}{2}} F(r, X_1(r))\|_{L^p(\Omega, \dot{H}^0)} \\ &\lesssim \left(1 + \sup_{r \in [0, T]} \|\Lambda^{\frac{\theta}{2}} X_1(r)\|_{L^p(\Omega, \dot{H}^0)} \right) (t-s). \end{aligned}$$

Similarly, Lemma 2.1 yields that the third term on the right hand side of (2.18) is bounded by a constant times

$$\begin{aligned} \left(\int_s^t \|E(t-r)\Theta^{\frac{\alpha}{2}}BG(r)\|_{\mathcal{L}_2(H_0, \mathcal{H})}^2 dr \right)^{\frac{1}{2}} &\leq (t-s)^{\frac{1}{2}} \sup_{r \in [0, T]} \|\Theta^{\frac{\alpha}{2}}BG(r)\|_{\mathcal{L}_2(H_0, \mathcal{H})} \\ &= (t-s)^{\frac{1}{2}} \sup_{r \in [0, T]} \|\Lambda^{\frac{\alpha-1}{2}}G(r)\|_{\mathcal{L}_2^0} \\ &\lesssim (t-s)^{\frac{1}{2}} \sup_{r \in [0, T]} \|\Lambda^{\frac{\beta-1}{2}}G(r)\|_{\mathcal{L}_2^0} \lesssim (t-s)^{\frac{1}{2}}, \end{aligned}$$

which completes the proof of (2.16). The proof of (2.17) is entirely similar, except for the analysis of the stochastic term. By (2.11), it satisfies

$$\begin{aligned} \left(\int_s^t \|P^1 E(t-r)B\Lambda^{\frac{\alpha}{2}}G(r)\|_{\mathcal{L}_2^0}^2 dr \right)^{\frac{1}{2}} &= \left(\int_s^t \|P^1 E(t-r)B\Lambda^{\frac{1+\alpha-\beta}{2}}\Lambda^{\frac{\beta-1}{2}}G(r)\|_{\mathcal{L}_2^0}^2 dr \right)^{\frac{1}{2}} \\ &\lesssim \sup_{\tau \in [0, T]} \|\Lambda^{\frac{\beta-1}{2}}G(\tau)\|_{\mathcal{L}_2^0} \left(\int_s^t (t-r)^{2\min(\beta-\alpha, 1)} dr \right)^{\frac{1}{2}} \\ &\lesssim (t-s)^{\min(\beta-\alpha, 1) + \frac{1}{2}} \end{aligned}$$

which combined with the previous estimates proves (2.17) and therefore finishes the proof of the theorem. \square

3. APPROXIMATION AND CONVERGENCE

We now consider a more concrete setting by taking $H = \dot{H}^0 = L^2(\mathcal{D})$, where \mathcal{D} denotes a convex polygonal bounded domain in \mathbb{R}^d . Let $\Lambda = -\Delta$ be the Laplace operator on H with zero Dirichlet boundary conditions. With this, the spaces $(\dot{H}^\alpha)_{\alpha \in \mathbb{R}}$ are related to classical Sobolev spaces by $\dot{H}^0 = L^2(\mathcal{D})$, $\dot{H}^1 = H_0^1(\mathcal{D})$ and $\dot{H}^2 = D(\Lambda) = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$, where $H^n(\mathcal{D})$ denotes the Sobolev space of order $n \in \mathbb{N}$ on \mathcal{D} and $H_0^1(\mathcal{D})$ is the subspace of functions in $H^1(\mathcal{D})$ that are zero on the boundary of \mathcal{D} (see also Appendix A). Next, we introduce our fully discrete approximation of the solution to (2.12). For the spatial discretization, a standard continuous finite element method is employed and for the temporal discretization, a rational approximation of the semigroup. This is the same approach as in Section 5 of [20] to which the reader is referred for further details, but see also [18, 19]. We then show a strong and a weak convergence result for this approximation.

To be precise, we take $(V_h^\kappa)_{h \in (0, 1]} \subset \dot{H}^1$, $\kappa \in \{2, 3\}$, to be a standard family of finite element function spaces consisting of continuous piecewise polynomials of degree $\kappa - 1$, with respect to a regular family of triangulations of \mathcal{D} with maximal mesh size h , that are zero on the boundary of \mathcal{D} . They are equipped with the inner product $\langle \cdot, \cdot \rangle_{V_h^\kappa} = \langle \cdot, \cdot \rangle_{\dot{H}^0}$. On this space, let a discrete counterpart $\Lambda_h: V_h^\kappa \rightarrow V_h^\kappa$ to Λ be defined by

$$\langle \Lambda_h v_h, u_h \rangle_{\dot{H}^0} = \left\langle \Lambda^{\frac{1}{2}} v_h, \Lambda^{\frac{1}{2}} u_h \right\rangle_{\dot{H}^0} = \langle v_h, u_h \rangle_{\dot{H}^1}$$

for all $v_h, u_h \in V_h^\kappa$. Fractional powers of Λ_h are defined in the same way as for Λ . We define the generalized orthogonal projector $P_h: \dot{H}^{-1} \rightarrow V_h^\kappa$ by $\langle P_h v, v_h \rangle = \dot{H}^{-1} \langle v, v_h \rangle_{\dot{H}^1}$ for all $v \in \dot{H}^{-1}$ and $v_h \in V_h^\kappa$, where $\dot{H}^{-1} \langle \cdot, \cdot \rangle_{\dot{H}^1}$ denotes the dual pairing with respect to \dot{H}^0 . Note that P_h coincides with the usual orthogonal projector when restricted to \dot{H}^0 . For our convergence results, we need the following assumption on $(V_h^\kappa)_{h \in (0, 1]}$.

Assumption 3.1. *There exists a constant $C > 0$ such that, for all $h \in (0, 1]$, the operators Λ_h and P_h satisfy*

$$\|\Lambda_h^{\frac{\alpha}{2}} P_h\|_{\mathcal{L}(\dot{H}^0)} \leq Ch^{-\alpha}, \quad \alpha \in [0, 2], \quad (3.1)$$

$$\|\Lambda_h^{\frac{\alpha}{2}} P_h v\|_{\dot{H}^0} \leq C \|\Lambda^{\frac{\alpha}{2}} v\|_{\dot{H}^0}, \quad v \in \dot{H}^\alpha, \alpha \in [-1, 1]. \quad (3.2)$$

In our setting, this assumption is fulfilled if the mesh underlying V_h^κ is quasi-uniform, see (3.28) of [26] and (3.17) of [21]. The counterpart to this assumption, that there exists a constant $C > 0$ such that

$$\|\Lambda^{\frac{\alpha}{2}} v_h\|_{\dot{H}^0} \leq C \|\Lambda_h^{\frac{\alpha}{2}} v_h\|_{\dot{H}^0}, \quad v_h \in V_h^\kappa, \alpha \in [-1, 1], \quad (3.3)$$

holds without the assumption of quasi-uniformity. A combination of (3.1) and (3.2) yields, for $v \in \dot{H}^0$ and $\alpha \in [0, 1]$,

$$\|P_h \Lambda^{\frac{\alpha}{2}} v\|_{\dot{H}^0} = \|\Lambda_h^{\frac{\alpha}{2}} \Lambda_h^{-\frac{\alpha}{2}} P_h \Lambda^{\frac{\alpha}{2}} v\|_{\dot{H}^0} \leq C h^{-\alpha} \|\Lambda_h^{-\frac{\alpha}{2}} P_h \Lambda^{\frac{\alpha}{2}} v\|_{\dot{H}^0} \lesssim h^{-\alpha} \|v\|_{\dot{H}^0}, \quad (3.4)$$

where we have used the fact that $\Lambda^{\frac{\alpha}{2}} v \in \dot{H}^{-\alpha}$.

Let

$$A_h = \begin{bmatrix} 0 & -I \\ \Lambda_h & 0 \end{bmatrix}$$

be a discrete counterpart to A on the product space $\mathcal{V}_h^\kappa = V_h^\kappa \oplus V_h^\kappa$, equipped with the same inner product as \mathcal{H} . With some abuse of notation, by the expression $P_h v$, $v = [v_1, v_2]^\top \in \mathcal{H}$, we denote the element $[P_h v_1, P_h v_2]^\top \in \mathcal{V}_h^\kappa$. The operator $-A_h$, similarly to $-A$, generates a C_0 -(semi)group $E_h: \mathbb{R} \rightarrow \mathcal{L}(\mathcal{V}_h^\kappa)$.

Let

$$\Theta_h^{\frac{\alpha}{2}} = \begin{bmatrix} \Lambda_h^{\frac{\alpha}{2}} & 0 \\ 0 & \Lambda_h^{\frac{\alpha}{2}} \end{bmatrix} \in \mathcal{L}(\mathcal{V}_h^\kappa)$$

and note that, as a straightforward consequence of (3.2) and (3.3), for every $\alpha \in [0, 1]$ there exists a constant $C > 0$ such that for all $h \in (0, 1]$, $v_h \in \mathcal{V}_h^\kappa$ and $v \in \mathcal{H}^\alpha$,

$$\begin{aligned} \|\Theta_h^{\frac{\alpha}{2}} v_h\|_{\mathcal{H}} &\leq C \|\Theta_h^{\frac{\alpha}{2}} v_h\|_{\mathcal{H}}, \\ \|\Theta_h^{\frac{\alpha}{2}} P_h v\|_{\mathcal{H}} &\leq C \|\Theta_h^{\frac{\alpha}{2}} v\|_{\mathcal{H}}, \end{aligned} \quad (3.5)$$

and, using (3.4), one shows that there exists a constant $C > 0$ such that for all $h \in (0, 1]$ and $v \in \mathcal{H}$,

$$\|P_h \Theta_h^{\frac{\alpha}{2}} v\|_{\mathcal{H}} \leq C h^{-\alpha} \|v\|_{\mathcal{H}}. \quad (3.6)$$

For the temporal discretization, consider a uniform time grid $t_j = j\Delta t = j(T/N_{\Delta t})$, $j = 0, \dots, N_{\Delta t}$. Let $R: \mathbb{C} \rightarrow \mathbb{C}$ be a rational function such that $|R(iy)| \leq 1$ for all $y \in \mathbb{R}$ and, for some $\rho \in \mathbb{N}$ and $C, b > 0$, $|R(iy) - e^{-iy}| \leq C|y|^{\rho+1}$ for all $y \in \mathbb{R}$ with $|y| \leq b$, where $i = \sqrt{-1}$. Several classes of such functions are described in Section 4 of [4]. They include certain Padé approximations of the exponential function, for example, the one corresponding to the backward Euler scheme ($R(z) = 1/(1+z)$) with order $\rho = 1$ and the one corresponding to the Crank–Nicolson scheme ($R(z) = (1-z/2)/(1+z/2)$) with order $\rho = 2$. We write $E_{h,\Delta t}^n = R(\Delta t A_h)^n$ for the rational approximation of the operator $E_h(t_n)$. The Crank–Nicolson approximation given by

$$R(\Delta t A_h) = \frac{1 - \frac{1}{2}\Delta t A_h}{1 + \frac{1}{2}\Delta t A_h}$$

is of particular importance for the wave equation as it preserves the energy. We define the interpolation $\tilde{E}: [0, T] \rightarrow \mathcal{L}(\mathcal{H})$ of the approximation by the step function

$$\tilde{E}(t) = \chi_{\{0\}}(t) P_h + \sum_{j=1}^{N_{\Delta t}} \chi_{(t_{j-1}, t_j]}(t) E_{h,\Delta t}^j P_h, \quad (3.7)$$

for $t \in [0, T]$, where χ denotes the indicator function. For this interpolation, the stability result

$$\|\tilde{E}(t)\|_{\mathcal{L}(\mathcal{H})} \leq 1 \quad (3.8)$$

for all $t \in [0, T]$ holds uniformly in $h, \Delta t \in (0, 1]$, see, e.g., [19]. Moreover, the following error estimate with respect to its first component holds.

Lemma 3.2 ([20], Lem. 5.2). *Let $\alpha \geq 0$ and assume that \tilde{E} is given by (3.7) for a rational approximation $E_{h,\Delta t}$ of order $\rho \in \mathbb{N}$ of $E_h(\Delta t)$ for $\Delta t \leq 1$. Then there exists a constant $C > 0$ such that, for all $h, \Delta t \in (0, 1]$,*

$$\begin{aligned} \sup_{t \in [0, T]} \left(\|P^1(\tilde{E}(t) - E(t))\Theta^{-\frac{\alpha}{2}}\|_{\mathcal{L}(\mathcal{H}, \dot{H}^0)} + \|\Lambda^{-\frac{\alpha}{4}}P^1(\tilde{E}(t) - E(t))B\Lambda^{\frac{1}{2}-\frac{\alpha}{4}}\|_{\mathcal{L}(\dot{H}^0)} \right) \\ \leq C \left(h^{\min(\alpha \frac{\kappa}{\kappa+1}, \kappa)} + \Delta t^{\min(\alpha \frac{\rho}{\rho+1}, 1)} \right). \end{aligned}$$

We now define the fully discrete approximation $\left(X_{h,\Delta t}^j\right)_{j=0}^{N_{\Delta t}}$ by the recursion scheme

$$X_{h,\Delta t}^j = E_{h,\Delta t} \left(X_{h,\Delta t}^{j-1} + P_h BF(t_{j-1}, P^1 X_{h,\Delta t}^{j-1}) \Delta t + P_h BG(t_{j-1}) \Delta W^j \right) \quad (3.9)$$

for $j = 1, 2, \dots, N_{\Delta t}$, where $\Delta W^j = W(t_j) - W(t_{j-1})$, with $X_{h,\Delta t}^0 = P_h x_0$. In closed form it is given by the discrete mild solution formulation

$$X_{h,\Delta t}^n = E_{h,\Delta t}^n x_0 + \Delta t \sum_{j=0}^{n-1} E_{h,\Delta t}^{n-j} P_h BF \left(t_j, P^1 X_{h,\Delta t}^j \right) + \sum_{j=0}^{n-1} E_{h,\Delta t}^{n-j} P_h BG(t_j) \Delta W^j$$

for $n = 0, 1, \dots, N_{\Delta t}$. This we extend to a continuous time process $\tilde{X}: \Omega \times [0, T] \rightarrow \mathcal{V}_h^\kappa$ by

$$\tilde{X}(t) = \tilde{E}(t)x_0 + \int_0^t \tilde{E}(t-s)BF \left(\lfloor s \rfloor_{\Delta t}, \tilde{X}_1(\lfloor s \rfloor_{\Delta t}) \right) ds + \int_0^t \tilde{E}(t-s)BG(\lfloor s \rfloor_{\Delta t}) dW(s), \quad (3.10)$$

for $t \in [0, T]$. Here $\tilde{X}_1 = P^1 \tilde{X}$ and $\lfloor \cdot \rfloor_{\Delta t} = \lfloor \cdot / \Delta t \rfloor \Delta t$, where $\lfloor \cdot \rfloor$ denotes the floor function. It is straightforward to see that $\tilde{X}(t_j) = X_{h,\Delta t}^j$ P -a.s. for all $j = 0, 1, \dots, N_{\Delta t}$.

3.1. Strong convergence

Under Assumptions 2.2 and 3.1, we now deduce a strong convergence result, *i.e.*, convergence measured in $\|\cdot\|_{L^2(\Omega, \dot{H}^0)}$. For the weak convergence we shall also need strong convergence in a negative norm, for which we need an additional assumption.

Assumption 3.3. *In addition to the requirements of Assumption 2.2, there exist parameters $\mu \in [0, 2]$, $\nu \in [\max(\mu - 1, 0), \min(r, 1)]$ and a constant $C > 0$ such that for every $t \in [0, T]$, $F(t, \cdot) \in \mathcal{G}_p^1(\dot{H}^0, \dot{H}^{-\min(\mu, 1)})$ and*

$$\|\Lambda^{-\frac{\mu}{2}}F'(t, u)v\|_{\dot{H}^0} \leq C \left(1 + \|\Lambda^{\frac{\mu}{2}}u\|_{\dot{H}^0} \right) \|\Lambda^{-\frac{\nu}{2}}v\|_{\dot{H}^0} \quad (3.11)$$

for all $u \in \dot{H}^\nu$ and $v \in \dot{H}^{-\nu}$.

Note that, as a consequence of the Lipschitz condition of Assumption 2.2(ii) and the fact that $\|\cdot\|_{\dot{H}^{-1}}$ is continuous on $\dot{H}^{-\min(\mu, 1)}$, we obtain that for all $t \in [0, T]$ and $u, v \in \dot{H}^0$,

$$\begin{aligned} \|\Lambda^{-\frac{1}{2}}F'(t, u)v\|_{\dot{H}^0} &= \left\| \Lambda^{-\frac{1}{2}} \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (F(t, u + \epsilon v) - F(t, u)) \right\|_{\dot{H}^0} \\ &= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \left\| \Lambda^{-\frac{1}{2}} (F(t, u + \epsilon v) - F(t, u)) \right\|_{\dot{H}^0} \\ &\lesssim \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \epsilon \|v\|_{\dot{H}^0} = \|v\|_{\dot{H}^0}, \end{aligned} \quad (3.12)$$

which, since $\mathcal{G}_p^1(\dot{H}^0, \dot{H}^{-\min(\mu, 1)}) \subset \mathcal{G}_p^1(\dot{H}^0, \dot{H}^{-1})$, implies that $F(t, \cdot) \in \mathcal{G}_b^1(\dot{H}^0, \dot{H}^{-1})$.

We also need the following version of Gronwall's lemma, see 2.2 (9) of [15].

Lemma 3.4 (Gronwall's lemma). *Let $C > 0$ and let $(a_j)_{j=1}^\infty, (b_j)_{j=1}^\infty$ be nonnegative sequences. If*

$$a_n \leq C + \sum_{j=1}^{n-1} b_j a_j$$

for all $n \geq 1$ then

$$a_n \leq C \exp \left(\sum_{j=1}^{n-1} b_j \right)$$

for all $n \geq 1$.

Note that we, for a real-valued sequence $(a_j)_{j=1}^\infty$, use the convention $\sum_{j=1}^0 a_j = 0$. With this lemma in place, we are ready to deduce a strong convergence result.

Theorem 3.5 (Strong convergence). *Let X be the mild solution of the stochastic wave equation given by (2.13), let \tilde{X} be the fully discrete approximation given by (3.10), let Assumptions 2.2 and 3.1 hold. Then, for all $\alpha \in [0, \min(r, 1)]$ and any $p \in [1, \infty)$,*

$$\sup_{\substack{n \in \{0, 1, \dots, N_{\Delta t}\} \\ h, \Delta t \in (0, 1]}} \|\tilde{X}(t_n)\|_{L^p(\Omega, \mathcal{H}^\alpha)} < \infty, \quad (3.13)$$

and there exists a constant $C > 0$ such that for all $h, \Delta t \in (0, 1]$

$$\sup_{n \in \{0, 1, \dots, N_{\Delta t}\}} \|\tilde{X}_1(t_n) - X_1(t_n)\|_{L^p(\Omega, \dot{H}^0)} \leq C(h^{r \frac{\kappa}{\kappa+1}} + \Delta t^{\min(r \frac{\rho}{\rho+1}, \eta, 1)}). \quad (3.14)$$

If, in addition to this, Assumption 3.3 also holds, then for any $p \in [1, \infty)$, there exists a constant $C > 0$ such that for all $h, \Delta t \in (0, 1]$

$$\sup_{n \in \{0, 1, \dots, N_{\Delta t}\}} \|\tilde{X}_1(t_n) - X_1(t_n)\|_{L^p(\Omega, \dot{H}^{-\nu})} \leq C(h^{r' \frac{\kappa}{\kappa+1}} + \Delta t^{\min(r' \frac{\rho}{\rho+1}, \eta, 1)}), \quad (3.15)$$

where $r' = \min(\max(2\nu, \beta), \max(2\nu, 1 + \theta), \delta)$.

Proof. We start by showing (3.13) using the representation (3.10). Take $\alpha \in [0, \min(r, 1)]$. First note that since $\Theta_h^{\frac{\alpha}{2}}$ commutes with $\tilde{E}(t_n)$ for all $n \in \{0, 1, \dots, N_{\Delta t}\}$ and since $\alpha \leq 1$, by (3.5) and (3.8),

$$\|\Theta_h^{\frac{\alpha}{2}} \tilde{E}(t_n) v\|_{\mathcal{H}} \lesssim \|\Theta_h^{\frac{\alpha}{2}} \tilde{E}(t_n) v\|_{\mathcal{H}} = \|\tilde{E}(t_n) \Theta_h^{\frac{\alpha}{2}} P_h v\|_{\mathcal{H}} \lesssim \|\Theta_h^{\frac{\alpha}{2}} P_h v\|_{\mathcal{H}} \lesssim \|\Theta_h^{\frac{\alpha}{2}} v\|_{\mathcal{H}}$$

for all $n \in \{0, 1, \dots, N_{\Delta t}\}$ and $v \in \mathcal{H}^\alpha$. Using this along with Lemma 2.1, (2.6) and finally Assumption 2.2, one obtains for $n = 1, 2, \dots, N_{\Delta t}$ that

$$\begin{aligned} \|\tilde{X}(t_n)\|_{L^p(\Omega, \mathcal{H}^\alpha)} &= \|\Theta_h^{\frac{\alpha}{2}} X_{h, \Delta t}^n\|_{L^p(\Omega, \mathcal{H})} \\ &\lesssim \|\Theta_h^{\frac{\alpha}{2}} \tilde{E}(t_n) x_0\|_{\mathcal{H}} + \int_0^{t_n} \|\Theta_h^{\frac{\alpha}{2}} \tilde{E}(t_n - s) BF(\lfloor s \rfloor_{\Delta t}, \tilde{X}_1(\lfloor s \rfloor_{\Delta t}))\|_{L^p(\Omega, \mathcal{H})} ds \\ &\quad + \left(\int_0^{t_n} \|\Theta_h^{\frac{\alpha}{2}} \tilde{E}(t_n - s) BG(\lfloor s \rfloor_{\Delta t})\|_{\mathcal{L}_2(H_0, \mathcal{H})}^2 ds \right)^{\frac{1}{2}} \\ &\lesssim \|\Theta_h^{\frac{\alpha}{2}} x_0\|_{\mathcal{H}} + \sum_{j=0}^{n-1} \Delta t \|\Theta_h^{\frac{\alpha}{2}} BF(t_j, P^1 X_{h, \Delta t}^j)\|_{L^p(\Omega, \mathcal{H})} + \left(\sum_{j=0}^{n-1} \Delta t \|\Theta_h^{\frac{\alpha}{2}} BG(t_j)\|_{\mathcal{L}_2(H_0, \mathcal{H})}^2 \right)^{\frac{1}{2}} \end{aligned}$$

$$\begin{aligned}
&= \|x_0\|_{\mathcal{H}^\alpha} + \sum_{j=0}^{n-1} \Delta t \|\Lambda^{\frac{\alpha-1}{2}} F(t_j, P^1 X_{h,\Delta t}^j)\|_{L^p(\Omega, \dot{H}^0)} + \left(\sum_{j=0}^{n-1} \Delta t \|\Lambda^{\frac{\alpha-1}{2}} G(t_j)\|_{\mathcal{L}_2^0} \right)^{\frac{1}{2}} \\
&\lesssim \|x_0\|_{\mathcal{H}^\delta} + \Delta t \sum_{j=0}^{n-1} \left(1 + \|\Lambda^{\frac{\alpha-1}{2}} P^1 X_{h,\Delta t}^j\|_{L^p(\Omega, \dot{H}^0)} \right) + \sqrt{T} \lesssim 1 + \Delta t \sum_{j=0}^{n-1} \|\tilde{X}(t_j)\|_{L^p(\Omega, \mathcal{H}^\alpha)}.
\end{aligned}$$

An application of Lemma 3.4 now yields (3.13).

We prove (3.14) and (3.15) in tandem, with $\alpha \in \{0, \nu\}$, by first making the split

$$\begin{aligned}
\|\tilde{X}_1(t_n) - X_1(t_n)\|_{L^p(\Omega, \dot{H}^{-\alpha})} &= \|\Lambda^{-\frac{\alpha}{2}} P^1 (\tilde{X}(t_n) - X(t_n))\|_{L^p(\Omega, \dot{H}^0)} \\
&\leq \|\Lambda^{-\frac{\alpha}{2}} P^1 (\tilde{E}(t_n) - E(t_n)) x_0\|_{L^p(\Omega, \dot{H}^0)} \\
&\quad + \int_0^{t_n} \left\| \Lambda^{-\frac{\alpha}{2}} P^1 \left(\tilde{E}(t_n - s) BF(\lfloor s \rfloor_{\Delta t}, \tilde{X}_1(\lfloor s \rfloor_{\Delta t})) \right. \right. \\
&\quad \left. \left. - E(t_n - s) BF(s, X_1(s)) \right) \right\|_{L^p(\Omega, \dot{H}^0)} ds \\
&\quad + \left\| \int_0^{t_n} \Lambda^{-\frac{\alpha}{2}} P^1 \left(\tilde{E}(t_n - s) BG(\lfloor s \rfloor_{\Delta t}) - E(t_n - s) BG(s) \right) dW(s) \right\|_{L^p(\Omega, \dot{H}^0)} \\
&= \text{I} + \text{II} + \text{III},
\end{aligned}$$

for arbitrary $n = 1, 2, \dots, N_{\Delta t}$. For the first term, by Lemma 3.2 and Assumption 2.2(iii),

$$\begin{aligned}
\text{I} &= \|\Lambda^{-\frac{\alpha}{2}} P^1 (\tilde{E}(t_n) - E(t_n)) \Theta^{-\frac{\delta}{2}} \Theta^{\frac{\delta}{2}} x_0\|_{\dot{H}^0} \\
&\leq \|\Lambda^{-\frac{\alpha}{2}}\|_{\mathcal{L}(\dot{H}^0)} \|P^1 (\tilde{E}(t_n) - E(t_n)) \Theta^{-\frac{\delta}{2}}\|_{\mathcal{L}(\mathcal{H}, \dot{H}^0)} \|\Theta^{\frac{\delta}{2}} x_0\|_{\mathcal{H}} \\
&\lesssim h^{\min(\delta \frac{\kappa}{\kappa+1}, \kappa)} + \Delta t^{\min(\delta \frac{\rho}{\rho+1}, 1)}.
\end{aligned}$$

Before treating II, we consider the last term. Lemma 2.1 yields

$$\begin{aligned}
\text{III}^2 &\lesssim \int_0^{t_n} \|\Lambda^{-\frac{\alpha}{2}} P^1 (\tilde{E}(t_n - s) BG(\lfloor s \rfloor_{\Delta t}) - E(t_n - s) BG(s))\|_{\mathcal{L}_2^0}^2 ds \\
&= \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \|\Lambda^{-\frac{\alpha}{2}} P^1 (\tilde{E}(t_n - s) BG(t_j) - E(t_n - s) BG(s))\|_{\mathcal{L}_2^0}^2 ds.
\end{aligned}$$

For the last integrand, we make the split

$$\begin{aligned}
\|\Lambda^{-\frac{\alpha}{2}} P^1 (\tilde{E}(t_n - s) BG(t_j) - E(t_n - s) BG(s))\|_{\mathcal{L}_2^0} &\leq \|\Lambda^{-\frac{\alpha}{2}} P^1 (\tilde{E}(t_n - s) - E(t_n - s)) BG(t_j)\|_{\mathcal{L}_2^0} \\
&\quad + \|\Lambda^{-\frac{\alpha}{2}} P^1 (E(t_n - s) B(G(t_j) - G(s)))\|_{\mathcal{L}_2^0} \\
&= \text{IV} + \text{V}.
\end{aligned}$$

For the first of these terms, by (2.1), (2.7) and Lemma 3.2, and since $\Lambda^{-\frac{\alpha}{2}}$ is a bounded operator, we get

$$\begin{aligned}
\text{IV} &\leq \|\Lambda^{-\frac{\alpha}{2}}\|_{\mathcal{L}(\dot{H}^0)} \|P^1 (\tilde{E}(t_n - s) - E(t_n - s)) B \Lambda^{\frac{1-\beta}{2}}\|_{\mathcal{L}(\dot{H}^0)} \|\Lambda^{\frac{\beta-1}{2}} G(t_j)\|_{\mathcal{L}_2^0} \\
&\lesssim \|P^1 (\tilde{E}(t_n - s) - E(t_n - s)) B \Lambda^{\frac{1-\beta}{2}}\|_{\mathcal{L}(\dot{H}^0)} \\
&= \|P^1 (\tilde{E}(t_n - s) - E(t_n - s)) \Theta^{-\frac{\beta}{2}} B \Lambda^{\frac{1}{2}}\|_{\mathcal{L}(\dot{H}^0)} \\
&\lesssim \|P^1 (\tilde{E}(t_n - s) - E(t_n - s)) \Theta^{-\frac{\beta}{2}}\|_{\mathcal{L}(\dot{H}^0)} \leq h^{\min(\beta \frac{\kappa}{\kappa+1}, \kappa)} + \Delta t^{\min(\beta \frac{\rho}{\rho+1}, 1)}.
\end{aligned}$$

Furthermore, Lemma 3.2 also implies that

$$\begin{aligned} \|\Lambda^{-\frac{\alpha}{2}} P^1 \left(\tilde{E}(t_n - s) - E(t_n - s) \right) B \Lambda^{\frac{1-\beta}{2}} \|_{\mathcal{L}(\dot{H}^0)} &= \|\Lambda^{-\frac{\alpha}{2}} P^1 \left(\tilde{E}(t_n - s) - E(t_n - s) \right) B \Lambda^{\frac{1-\alpha}{2}} \|_{\mathcal{L}(\dot{H}^0)} \\ &\quad \times \|\Lambda^{\frac{\alpha-\beta}{2}} \|_{\mathcal{L}(\dot{H}^0)} \\ &\lesssim h^{2\alpha \frac{\kappa}{\kappa+1}} + \Delta t^{\min(2\alpha \frac{\rho}{\rho+1}, 1)}, \end{aligned}$$

using also the fact that $\alpha \leq r \leq \beta$ and that $\alpha \leq 2 \leq \kappa$ in the last inequality. Therefore

$$\text{IV} \lesssim h^{\min(\max(2\alpha, \beta) \frac{\kappa}{\kappa+1}, \kappa)} + \Delta t^{\min(\max(2\alpha, \beta) \frac{\rho}{\rho+1}, 1)}.$$

Next, for V, by (2.1), (2.8), Assumption 2.2 and (2.7), we obtain

$$\begin{aligned} \text{V} &= \|\Lambda^{-\frac{\alpha+\beta}{2}} P^1 (E(t_n - s) B \Lambda^{\frac{1}{2}} \Lambda^{\frac{\beta-1}{2}} (G(t_j) - G(s))) \|_{\mathcal{L}_2^0} \\ &\leq \|\Lambda^{-\frac{\alpha+\beta}{2}} \|_{\mathcal{L}(\dot{H}^0)} \|E(t_n - s)\|_{\mathcal{L}(\mathcal{H})} \|B \Lambda^{\frac{1}{2}} \|_{\mathcal{L}(\dot{H}^0, \mathcal{H})} \|\Lambda^{\frac{\beta-1}{2}} (G(t_j) - G(s)) \|_{\mathcal{L}_2^0} \\ &\lesssim |t_j - s|^\eta. \end{aligned}$$

As a consequence, we arrive at

$$\text{III} \lesssim \left(h^{\min(\max(2\alpha, \beta) \frac{\kappa}{\kappa+1}, \kappa)} + \Delta t^{\min(\max(2\alpha, \beta) \frac{\rho}{\rho+1}, \eta, 1)} \right).$$

We now continue with II and note that

$$\text{II} = \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \|\Lambda^{-\frac{\alpha}{2}} P^1 \left(\tilde{E}(t_n - s) B F(t_j, \tilde{X}_1(t_j)) - E(t_n - s) B F(s, X_1(s)) \right) \|_{L^p(\Omega, \dot{H}^0)} ds.$$

We split the integrand with

$$\begin{aligned} &\|\Lambda^{-\frac{\alpha}{2}} P^1 \left(\tilde{E}(t_n - s) B F(t_j, \tilde{X}_1(t_j)) - E(t_n - s) B F(s, X_1(s)) \right) \|_{L^p(\Omega, \dot{H}^0)} \\ &\leq \|\Lambda^{-\frac{\alpha}{2}} P^1 \left(\tilde{E}(t_n - s) - E(t_n - s) \right) B F(t_j, \tilde{X}_1(t_j)) \|_{L^p(\Omega, \dot{H}^0)} \\ &\quad + \|\Lambda^{-\frac{\alpha}{2}} P^1 E(t_n - s) B \left(F(t_j, \tilde{X}_1(t_j)) - F(t_j, X_1(t_j)) \right) \|_{L^p(\Omega, \dot{H}^0)} \\ &\quad + \|\Lambda^{-\frac{\alpha}{2}} P^1 E(t_n - s) B \left(F(t_j, X_1(t_j)) - F(t_j, X_1(s)) \right) \|_{L^p(\Omega, \dot{H}^0)} \\ &\quad + \|\Lambda^{-\frac{\alpha}{2}} P^1 E(t_n - s) B \left(F(t_j, X_1(s)) - F(s, X_1(s)) \right) \|_{L^p(\Omega, \dot{H}^0)} = \text{VI} + \text{VII} + \text{VIII} + \text{IX}. \end{aligned}$$

For the first of these terms, under Assumption 2.2, by Lemma 3.2 and (2.7), using the assumption that $\theta \leq \min(\beta, \delta, 1) \leq \min(r, 1)$, we get

$$\begin{aligned} \text{VI} &\leq \|\Lambda^{-\frac{\alpha}{2}} \|_{\mathcal{L}(\dot{H}^0)} \|P^1 \left(\tilde{E}(t_n - s) - E(t_n - s) \right) B \Lambda^{\frac{\theta}{2}} \|_{\mathcal{L}(\dot{H}^0)} \|\Lambda^{\frac{\theta}{2}} F(t_j, \tilde{X}_1(t_j)) \|_{L^p(\Omega, \dot{H}^0)} \\ &\lesssim \|P^1 \left(\tilde{E}(t_n - s) - E(t_n - s) \right) \Theta^{-\frac{1+\theta}{2}} \|_{\mathcal{L}(\mathcal{H}, \dot{H}^0)} \|B \Lambda^{\frac{1}{2}} \|_{\mathcal{L}(\dot{H}^0, \mathcal{H})} \left(1 + \|\Lambda^{\frac{\theta}{2}} \tilde{X}_1(t_j) \|_{L^p(\Omega, \dot{H}^0)} \right) \\ &\lesssim \left(h^{(1+\theta) \frac{\kappa}{\kappa+1}} + \Delta t^{\min((1+\theta) \frac{\rho}{\rho+1}, 1)} \right). \end{aligned}$$

Since $\alpha \leq r \leq 1 + \theta \leq 2$, similarly to term IV, Lemma 3.2 also implies that

$$\begin{aligned}
\text{VI} &\lesssim \|\Lambda^{-\frac{\alpha}{2}} P^1 \left(\tilde{E}(t_n - s) - E(t_n - s) \right) B \Lambda^{\frac{1-\alpha}{2}} \|_{\mathcal{L}(\dot{H}^0)} \|\Lambda^{\frac{\alpha-1}{2}} F(t_j, \tilde{X}_1(t_j))\|_{L^p(\Omega, \dot{H}^0)} \\
&\lesssim \|\Lambda^{-\frac{\alpha}{2}} P^1 \left(\tilde{E}(t_n - s) - E(t_n - s) \right) B \Lambda^{\frac{1-\alpha}{2}} \|_{\mathcal{L}(\dot{H}^0)} \|\Lambda^{\frac{\theta}{2}} F(t_j, \tilde{X}_1(t_j))\|_{L^p(\Omega, \dot{H}^0)} \\
&\lesssim \left(h^{2\alpha \frac{\kappa}{\kappa+1}} + \Delta t^{\min(2\alpha \frac{\rho}{\rho+1}, 1)} \right) \left(1 + \|\Lambda^{\frac{\theta}{2}} \tilde{X}_1(t_j)\|_{L^p(\Omega, \dot{H}^0)} \right),
\end{aligned}$$

whence

$$\text{VI} \lesssim \left(h^{\max(2\alpha, 1+\theta) \frac{\kappa}{\kappa+1}} + \Delta t^{\min(\max(2\alpha, 1+\theta) \frac{\rho}{\rho+1}, 1)} \right).$$

Term VII can be bounded using (2.9), (2.8) and (2.7), as

$$\begin{aligned}
\text{VII} &= \|P^1 E(t_n - s) B \Lambda^{\frac{1}{2}} \Lambda^{-\frac{1+\alpha}{2}} \left(F(t_j, \tilde{X}_1(t_j)) - F(t_j, X_1(t_j)) \right)\|_{L^p(\Omega, \dot{H}^0)} \\
&\lesssim \|\Lambda^{-\frac{1+\alpha}{2}} \left(F(t_j, \tilde{X}_1(t_j)) - F(t_j, X_1(t_j)) \right)\|_{L^p(\Omega, \dot{H}^0)}.
\end{aligned}$$

If only Assumption 2.2 holds, we directly obtain

$$\text{VII} \lesssim \|\tilde{X}_1(t_j) - X_1(t_j)\|_{L^p(\Omega, \dot{H}^0)}.$$

If also Assumption 3.3 holds with $\alpha = \nu \geq \mu - 1$, then, since $\mathcal{G}_p^1(\dot{H}^0, \dot{H}^{-(1+\nu)}) \subset \mathcal{G}_p^1(\dot{H}^0, \dot{H}^{-\mu})$, we may use the mean value theorem along with (3.11) to deduce that

$$\begin{aligned}
\text{VII} &\leq \int_0^1 \|\Lambda^{-\frac{1+\nu}{2}} F' \left(t_j, s\tilde{X}_1(t_j) + (1-s)X_1(t_j) \right) \left(\tilde{X}_1(t_j) - X_1(t_j) \right)\|_{L^p(\Omega, \dot{H}^0)} ds \\
&\lesssim \int_0^1 \|\Lambda^{-\frac{\mu}{2}} F' \left(t_j, s\tilde{X}_1(t_j) + (1-s)X_1(t_j) \right) \left(\tilde{X}_1(t_j) - X_1(t_j) \right)\|_{L^p(\Omega, \dot{H}^0)} ds \\
&\lesssim \left(1 + \|\Lambda^{\frac{\nu}{2}} \tilde{X}_1(t_j)\|_{L^p(\Omega, \dot{H}^0)} + \|\Lambda^{\frac{\nu}{2}} X_1(t_j)\|_{L^p(\Omega, \dot{H}^0)} \right) \|\Lambda^{-\frac{\nu}{2}} \left(\tilde{X}_1(t_j) - X_1(t_j) \right)\|_{L^p(\Omega, \dot{H}^0)} \\
&\lesssim \left(1 + \|\Lambda^{\frac{\min(r,1)}{2}} \tilde{X}_1(t_j)\|_{L^p(\Omega, \dot{H}^0)} + \|\Lambda^{\frac{\min(r,1)}{2}} X_1(t_j)\|_{L^p(\Omega, \dot{H}^0)} \right) \\
&\quad \times \|\Lambda^{-\frac{\nu}{2}} \left(\tilde{X}_1(t_j) - X_1(t_j) \right)\|_{L^p(\Omega, \dot{H}^0)} \lesssim \|\Lambda^{-\frac{\nu}{2}} \left(\tilde{X}_1(t_j) - X_1(t_j) \right)\|_{L^p(\Omega, \dot{H}^0)},
\end{aligned}$$

where we also applied Theorem 2.3 and (3.13) in the last step. Similarly, we have

$$\text{VIII} \lesssim \|X_1(t_j) - X_1(s)\|_{L^p(\Omega, \dot{H}^0)} \lesssim |t_j - s|^{\min(r,1)}$$

if we only consider Assumption 2.2, while if also Assumption 3.3 holds, then

$$\text{VIII} \lesssim \|\Lambda^{-\frac{\nu}{2}} (X_1(t_j) - X_1(s))\|_{L^p(\Omega, \dot{H}^0)} \lesssim |t_j - s|^{\min(2\nu, 1)},$$

using Theorems 2.3 and 2.4 in the last step. For the last term, by (2.9), (2.8) and (2.7),

$$\text{IX} \lesssim \|\Lambda^{-\frac{1}{2}} (F(t_j, X_1(s)) - F(s, X_1(s)))\|_{L^p(\Omega, \dot{H}^0)} \lesssim \left(1 + \|X_1(s)\|_{L^p(\Omega, \dot{H}^0)} \right) |t_j - s|^\eta \lesssim |t_j - s|^\eta,$$

where we have applied Theorem 2.3 and Assumption 2.2(ii) in the last inequality.

Collecting the bounds on terms VI, VII, VIII and IX, we have shown that under Assumption 2.2, with $\alpha = 0$,

$$\text{II} \lesssim h^{(1+\theta) \frac{\kappa}{\kappa+1}} + \Delta t^{\min((1+\theta) \frac{\rho}{\rho+1}, r, 1)} + \Delta t \sum_{j=0}^{n-1} \|\tilde{X}_1(t_j) - X_1(t_j)\|_{L^p(\Omega, \dot{H}^0)}.$$

Taking also into account the bound on term I and that

$$\text{III} \lesssim \left(h^{\min(\beta \frac{\kappa}{\kappa+1}, \kappa)} + \Delta t^{\min(\beta \frac{\rho}{\rho+1}, \eta, 1)} \right),$$

we obtain (3.14) by Lemma 3.4.

On the other hand, if Assumption 3.3 also holds, the bound on term I remains the same while with $\alpha = \nu$,

$$\text{II} \lesssim h^{\min(\max(2\nu, 1+\theta) \frac{\kappa}{\kappa+1}, \kappa)} + \Delta t^{\min(\max(2\nu, 1+\theta) \frac{\rho}{\rho+1}, \max(2\nu, r), \eta, 1)} + \Delta t \sum_{j=0}^{n-1} \|\Lambda^{-\frac{\nu}{2}}(\tilde{X}_1(t_j) - X_1(t_j))\|_{L^p(\Omega, \dot{H}^0)}.$$

Taking also into account that

$$\text{III} \lesssim \left(h^{\min(\max(2\nu, \beta) \frac{\kappa}{\kappa+1}, \kappa)} + \Delta t^{\min(\max(2\nu, \beta) \frac{\rho}{\rho+1}, \eta, 1)} \right),$$

by another application of Lemma 3.4, we obtain

$$\sup_{n \in \{0, 1, \dots, N_{\Delta t}\}} \|\tilde{X}_1(t_n) - X_1(t_n)\|_{L^p(\Omega, \dot{H}^{-\nu})} \lesssim h^{r' \frac{\kappa}{\kappa+1}} + \Delta t^{\min(\max(2\nu, \beta) \frac{\rho}{\rho+1}, \max(2\nu, 1+\theta) \frac{\rho}{\rho+1}, \delta \frac{\rho}{\rho+1}, \max(2\nu, r), \eta, 1)}. \quad (3.16)$$

We now note that

$$\min(\max(2\nu, \beta), \max(2\nu, 1+\theta), \delta) \leq \max(2\nu, r),$$

so, since $\rho/(\rho+1) < 1$, it holds that either $\max(2\nu, r) > \max(2\nu, \beta)\rho/(\rho+1)$ or $\max(2\nu, r) > \max(2\nu, 1+\theta)\rho/(\rho+1)$ or $\max(2\nu, r) > \delta\rho/(\rho+1)$. No matter which of these three cases occur, the minimum in the exponent of Δt in (3.16) is not attained at $\max(2\nu, r)$. In other words,

$$\begin{aligned} & \min \left(\max(2\nu, \beta) \frac{\rho}{\rho+1}, \max(2\nu, 1+\theta) \frac{\rho}{\rho+1}, \delta \frac{\rho}{\rho+1}, \max(2\nu, r), \eta, 1 \right) \\ &= \min \left(\max(2\nu, \beta) \frac{\rho}{\rho+1}, \max(2\nu, 1+\theta) \frac{\rho}{\rho+1}, \delta \frac{\rho}{\rho+1}, \eta, 1 \right) = \min \left(r' \frac{\rho}{\rho+1}, \eta, 1 \right). \end{aligned}$$

With this, the proof is completed. \square

3.2. Weak convergence and Malliavin calculus

To deduce a result on the weak convergence of the approximation (3.10) to the mild solution given by (2.13), we use Malliavin calculus for which we briefly review some definitions and results from [2] (the authors therein consider so called *refined* Sobolev–Malliavin spaces whereas we only need classical ones, hence the difference in notation below). To avoid technicalities we assume from here on that the filtration $(\mathcal{F}_t)_{t \in [0, T]}$ is generated by the Wiener process W . Let $I(\varphi) = \int_0^T \varphi(t) dW(t)$ for $\varphi \in L^2([0, T], \mathcal{L}_2(H_0, \mathbb{R})) \simeq L^2([0, T], H_0)$ and let \mathcal{S} denote the set of all cylindrical random variables of the form $F = f(I(\varphi_1), \dots, I(\varphi_N))$ for $f \in \mathcal{C}_b^1(\mathbb{R}^N, \mathbb{R})$ and a sequence $(\varphi_j)_{j=1}^N \subset L^2([0, T], H_0)$, $N \in \mathbb{N}$. The definition of the Malliavin derivative of $F \in \mathcal{S}$ is given by $D.F = \sum_{j=1}^N \partial_j f(I(\varphi_1), \dots, I(\varphi_N)) \varphi_j(\cdot) \in L^2([0, T] \times \Omega, H_0)$. For a given real separable Hilbert space U we let $\mathcal{S}(U)$ be the space of all U -valued random variables of the form $Y = \sum_{j=1}^M F_j v_j$ with $F_j \in \mathcal{S}$ and $v_j \in U$ for $j = 1, 2, \dots, M$, and define $DY = \sum_{j=1}^M v_j \otimes DF_j \in U \otimes L^2([0, T] \times \Omega, H_0)$. Note that $U \otimes L^2([0, T] \times \Omega, H_0) \simeq L^2([0, T] \times \Omega, \mathcal{L}_2(H_0, U)) \simeq L^2(\Omega, L^2([0, T], \mathcal{L}_2(H_0, U)))$. This definition of D does not depend on the specific representation of Y . The operator $D: \mathcal{S}(U) \subset L^p(\Omega, U) \rightarrow L^p(\Omega, L^2([0, T], \mathcal{L}_2(H_0, U)))$ is closable for any $p > 1$ and we write $\mathbb{D}^{1,p}(U)$ for the closure of $\mathcal{S}(U)$ in $L^p(\Omega, U)$ with respect to the norm

$$\|F\|_{\mathbb{D}^{1,p}(U)} = \left(\mathbb{E}[\|F\|_U^p] + \mathbb{E}[\|DF\|_{L^2([0, T], \mathcal{L}_2(H_0, U))}^p] \right)^{\frac{1}{p}}.$$

Next, we recall some of the basic properties of the Malliavin derivative. First of all, any deterministic element is Malliavin differentiable with Malliavin derivative zero. For predictable processes $\Phi \in L^2([0, T] \times \Omega, \mathcal{L}_2(H_0, U))$ and any $F \in \mathbb{D}^{1,2}(U)$, the following equality holds, sometimes referred to as Malliavin integration by parts,

$$\left\langle F, \int_0^T \Phi(t) dW(t) \right\rangle_{L^2(\Omega, U)} = \int_0^T \langle D_r F, \Phi(r) \rangle_{L^2(\Omega, \mathcal{L}_2(H_0, U))} dr. \quad (3.17)$$

We also need to know how the Malliavin derivative acts on stochastic integrals, but restrict ourselves to the case that the integrand $\Phi \in L^2([0, T], \mathcal{L}_2(H_0, U))$ is deterministic. Then, for all $t \in [0, T]$, it follows that $\int_0^t \Phi(r) dW(r) \in \mathbb{D}^{1,p}(U)$ for all $p > 1$ and

$$D \int_0^t \Phi(r) dW(r) = \chi_{(0,t]}(\cdot) \Phi(\cdot). \quad (3.18)$$

For Lebesgue integrals of stochastic processes on the other hand, Malliavin differentiation and integration simply commute (see [21], Prop. 4.8). If $\Phi: [0, T] \times \Omega \rightarrow U$ fulfill $\Phi \in \mathbb{D}^{1,p}(L^2([0, T], U))$ for some $p \geq 2$, then $\int_0^T \Phi(t) dt \in \mathbb{D}^{1,p}(U)$ and

$$D \int_0^T \Phi(t) dt = \int_0^T D\Phi(t) dt. \quad (3.19)$$

It is known (cf. [23], Sect. 1.3) that a sufficient condition for $\Phi \in \mathbb{D}^{1,p}(L^2([0, T], U))$ is that $\Phi(t) \in \mathbb{D}^{1,p}(U)$ for all $t \in [0, T]$ and that $\sup_{t \in [0, T]} \|D\Phi(t)\|_{L^2(\Omega \times [0, T], \mathcal{L}_2(H_0, U))} < \infty$. It is also worth mentioning that D commutes with any bounded linear operator between Hilbert spaces. For nonlinear mappings $\varphi \in \mathcal{G}_p^1(U, V)$, where V is another arbitrary real separable Hilbert space, a chain rule holds instead. If there is a $q \geq 1$ and a constant $C > 0$ such that $\|\varphi(x)\|_V \leq C(1 + \|x\|_U^q)$ and $\|\varphi'(x)\|_{\mathcal{L}(U, V)} \leq C(1 + \|x\|_U^{q-1})$ for all $x \in U$, then for all $p > 1$ and $F \in \mathbb{D}^{1,pq}(U)$, it holds that $\varphi(F) \in \mathbb{D}^{1,p}(V)$ and

$$D\varphi(F) = \varphi'(F)DF. \quad (3.20)$$

With these results in place, to be able to deduce a weak convergence result, we need to impose a stronger condition on G in (2.12) and on the covariance operator Q of the Wiener process. We also take this opportunity to specify our assumptions on the test function $\phi: \dot{H}^0 \rightarrow \mathbb{R}$. We identify $\mathcal{L}(\dot{H}^0, \mathbb{R})$ with \dot{H}^0 so that, for every $u \in \dot{H}^0$, $\phi'(u) \in \dot{H}^0$.

Assumption 3.6. *The following conditions hold:*

- (i) $\phi \in \mathcal{G}_p^2(\dot{H}^0, \mathbb{R})$,
- (ii) $G(t) = I$ for all $t \in [0, T]$, and either
- (iii) $\|\Lambda^{\beta-\frac{1}{2}} Q \Lambda^{-\frac{1}{2}}\|_{\text{Tr}} < \infty$ or
- (iv) $\text{Tr}(Q) < \infty$ and $\Lambda^{-1/2} \phi''(u) = \phi''(u) \Lambda^{-1/2}$ for all $u \in \dot{H}^0$.

The condition Assumption 3.6(iii), $\|\Lambda^{\beta-\frac{1}{2}} Q \Lambda^{-\frac{1}{2}}\|_{\text{Tr}} < \infty$, implies the corresponding condition $\|\Lambda^{\frac{\beta-1}{2}}\|_{\mathcal{L}_2^0} < \infty$ of Assumption 2.2(i). They are equivalent in the important case that $Q = I$ or more generally when $Q\Lambda = \Lambda Q$, see Theorem 2.1 of [18]. We also have the following simple but useful lemma.

Lemma 3.7. *Suppose that $\|\Lambda^{\beta-\frac{1}{2}} Q \Lambda^{-\frac{1}{2}}\|_{\text{Tr}} < \infty$. Then $\mathcal{L}(\dot{H}^{-1}, \dot{H}^0) \subset \mathcal{L}_2^0$ and for $\Gamma \in \mathcal{L}(\dot{H}^{-1}, \dot{H}^0)$ there exists a constant $C > 0$ such that*

$$\|\Gamma\|_{\mathcal{L}_2^0} \leq C \|\Gamma\|_{\mathcal{L}(\dot{H}^{-1}, \dot{H}^0)} = C \|\Gamma \Lambda^{\frac{1}{2}}\|_{\mathcal{L}(\dot{H}^0)}.$$

Proof. The claim is a consequence of (2.4), (2.3) and (2.1) via the calculation

$$\begin{aligned}\|\Gamma\|_{\mathcal{L}_2^0}^2 &= |\operatorname{Tr}(\Gamma Q \Gamma^*)| = |\operatorname{Tr}(\Gamma^* \Gamma Q)| \\ &= \left| \operatorname{Tr} \left(\left(\Gamma \Lambda^{\frac{1}{2}} \right)^* \Gamma \Lambda^{\frac{1}{2}} \Lambda^{-\beta} \Lambda^{\beta-\frac{1}{2}} Q \Lambda^{-\frac{1}{2}} \right) \right| \\ &\leq \|\Gamma \Lambda^{\frac{1}{2}}\|_{\mathcal{L}(\dot{H}^0)}^2 \|\Lambda^{-\beta}\|_{\mathcal{L}(\dot{H}^0)} \|\Lambda^{\beta-\frac{1}{2}} Q \Lambda^{-\frac{1}{2}}\|_{\operatorname{Tr}} \\ &= \|\Gamma\|_{\mathcal{L}(\dot{H}^{-1}, \dot{H}^0)}^2 \|\Lambda^{-\beta}\|_{\mathcal{L}(\dot{H}^0)} \|\Lambda^{\beta-\frac{1}{2}} Q \Lambda^{-\frac{1}{2}}\|_{\operatorname{Tr}}.\end{aligned}$$

□

The reason for including the alternative Assumption 3.6(iv) is that the condition $\|\Lambda^{\beta-\frac{1}{2}} Q \Lambda^{-\frac{1}{2}}\|_{\operatorname{Tr}} < \infty$ is hard to interpret in the common (cf. [5], Cor. 4.9, where Q is trace-class and induced by a covariance kernel) case that $\Lambda Q \neq Q \Lambda$. Note that Assumption 3.6(iv) implies that $\|\Lambda^{\frac{\beta-1}{2}}\|_{\mathcal{L}_2^0} < \infty$ for all $\beta \leq 1$.

As a first step towards our weak convergence result, we need the following regularity estimate for the Malliavin derivatives of the first component of the mild solution given by (2.13) and its approximation.

Lemma 3.8. *Let Assumptions 2.2 and 3.3 hold. Let X be the mild solution given by (2.13) of the stochastic wave equation and let \tilde{X} be the fully discrete approximation given by (3.10). Under Assumption 3.6(ii) and (iii), $X_1(t) \in \mathbb{D}^{1,p}(\dot{H}^0)$ for all $p \geq 2, t \in [0, T]$, and*

$$\sup_{t \in [0, T]} \|DX_1(t) \Lambda^{\frac{1}{2}}\|_{L^2([0, T] \times \Omega, \mathcal{L}(\dot{H}^0))} < \infty. \quad (3.21)$$

Furthermore, $\tilde{X}_1(t_n) \in \mathbb{D}^{1,p}(\dot{H}^0)$ for all $p \geq 2, n \in \{0, 1, \dots, N_{\Delta t}\}$, and

$$\sup_{\substack{n \in \{0, 1, \dots, N_{\Delta t}\} \\ h, \Delta t \in (0, 1]}} \|D\tilde{X}_1(t_n) \Lambda^{\frac{1}{2}}\|_{L^2([0, T] \times \Omega, \mathcal{L}(\dot{H}^0))} < \infty. \quad (3.22)$$

If Assumption 3.6(iv) holds in place of Assumption 3.6(iii), similar statements hold with (3.21) replaced by

$$\sup_{t \in [0, T]} \|\Lambda^{\frac{1}{2}} DX_1(t)\|_{L^2([0, T] \times \Omega, \mathcal{L}(\dot{H}^0))} < \infty, \quad (3.23)$$

and (3.22) by

$$\sup_{\substack{n \in \{0, 1, \dots, N_{\Delta t}\} \\ h, \Delta t \in (0, 1]}} \|\Lambda^{\frac{1}{2}} D\tilde{X}_1(t_n)\|_{L^2([0, T] \times \Omega, \mathcal{L}(\dot{H}^0))} < \infty. \quad (3.24)$$

Proof. We start by showing (3.21). Define the sequence $(X_1^n)_{n=0}^\infty$ by $X_1^0 = 0$ and, for $t \in [0, T]$ and $n \geq 0$,

$$X_1^{n+1}(t) = P^1 E(t) x_0 + \int_0^t P^1 E(t-s) BF(s, X_1^n(s)) ds + \int_0^t P^1 E(t-s) B dW(s). \quad (3.25)$$

Since the existence result ([12], Thm. 7.2) that we cited in Theorem 2.3 is proven via a fixed point argument for this sequence, it follows that $\lim_n X_1^n(t) = X_1(t)$ in $L^p(\Omega, \dot{H}^0)$ for all $t \in [0, T]$ and $p \geq 2$. By (2.7), (2.9) and (2.8), we have

$$\sup_{t \in [0, T]} \|P^1 E(t) B\|_{L(\dot{H}^{\alpha-1}, \dot{H}^\alpha)} < \infty. \quad (3.26)$$

This implies, along with Assumption 3.3, that $P^1 E(t-s) BF(t, \cdot) \in \mathcal{G}_b^1(\dot{H}^0, \dot{H}^{1-\min(\mu, 1)}) \subset \mathcal{G}_b^1(\dot{H}^0, \dot{H}^0)$ for all $s, t \in [0, T]$. The chain rule for the Malliavin derivative is then applicable so that $P^1 E(t-s) BF(t, X_1^n(t)) \in \mathbb{D}^{1,2}(\dot{H}^0)$ for all $s, t \in [0, T]$ as long as $X_1^n(t) \in \mathbb{D}^{1,2}(\dot{H}^0)$ for all $t \in [0, T]$, and that in this case

$DP^1E(t-s)BF(t, X_1^n(t)) = P^1E(t-s)BF'(t, X_1^n(t))DX_1^n(t)$ for all $s, t \in [0, T]$. Therefore, we may apply D to both sides of (3.25) and hence, by the fact that the Malliavin derivative of a deterministic element is zero, (3.18) and (3.19), we get

$$DX_1^{n+1}(t) = \int_0^t P^1E(t-r)BF'(r, X_1^n(r))DX_1^n(r)dr + \chi_{(0,t]}(\cdot)P^1E(t-\cdot)B$$

for all $t \in [0, T]$. Our aim is now to show that the sequence $(DX_1^n)_{n=0}^\infty$ has a limit in $\mathcal{C}([0, T], L^p(\Omega \times [0, T], \mathcal{L}(\dot{H}^{-1}, \dot{H}^0)))$ for any $p \geq 2$. To this end, note first that the mapping $[0, T] \ni t \mapsto \chi_{(0,t]}(\cdot)P^1E(t-\cdot)B$ is in this space by (3.26) with $\alpha = 0$. Next, we show that there is an equivalent norm $\|\cdot\|_\sigma$ on $\mathcal{C}([0, T], L^p(\Omega \times [0, T], \mathcal{L}(\dot{H}^{-1}, \dot{H}^0)))$ such that the mapping

$$[0, T] \ni t \mapsto I(X, Y)_t = \int_0^t P^1E(t-r)BF'(r, X(r))Y(r)dr,$$

where $X \in L^2(\Omega \times [0, T], \dot{H}^0)$ and $Y \in \mathcal{C}([0, T], L^p(\Omega \times [0, T], \mathcal{L}(\dot{H}^{-1}, \dot{H}^0)))$, fulfills

$$\|I(X, Y)\|_\sigma \leq \alpha \|Y\|_\sigma \quad (3.27)$$

for some $\alpha \in [0, 1)$ and all $X \in L^2(\Omega \times [0, T], \dot{H}^0)$. We choose, for $\sigma \geq 0$ to be determined, $\|Y\|_\sigma = \sup_{t \in [0, T]} e^{-\sigma t} \|Y(t)\|_{L^p(\Omega \times [0, T], \mathcal{L}(\dot{H}^{-1}, \dot{H}^0))}$ and note that for $t \in [0, T]$, by (2.8) and (3.12),

$$\begin{aligned} \|I(X, Y)_t\|_{L^p(\Omega \times [0, T], \mathcal{L}(\dot{H}^{-1}, \dot{H}^0))} &\leq \int_0^t \|P^1E(t-r)\|_{\mathcal{L}(\mathcal{H}, \dot{H}^0)} \|BF'(r, X(r))Y(r)\|_{L^p(\Omega \times [0, T], \mathcal{L}(\dot{H}^{-1}, \mathcal{H}))} dr \\ &\lesssim \int_0^t \|Y(r)\|_{L^p(\Omega \times [0, T], \mathcal{L}(\dot{H}^{-1}, \dot{H}^0))} dr \leq \|Y\|_\sigma \int_0^t e^{\sigma r} dr = \frac{e^{\sigma t} - 1}{\sigma} \|Y\|_\sigma. \end{aligned}$$

This implies that $\|I(X, Y)\|_\sigma \lesssim \sigma^{-1} \|Y\|_\sigma$ so that (3.27) is fulfilled for sufficiently large σ . By the Banach fixed point theorem, therefore, $(DX_1^n)_{n=0}^\infty$ has a limit \hat{Y} in $\mathcal{C}([0, T], L^p(\Omega \times [0, T], \mathcal{L}(\dot{H}^{-1}, \dot{H}^0)))$. In particular, $\lim_n DX_1^n(t) = \hat{Y}(t)$ in $L^p(\Omega \times [0, T], \mathcal{L}(\dot{H}^{-1}, \dot{H}^0))$ for all $t \in [0, T]$. Thus, by Lemma 3.7, $\lim_n DX_1^n(t) = \hat{Y}(t)$ in $L^p(\Omega \times [0, T], \mathcal{L}_2^0)$ for all $t \in [0, T]$. Since D is closed and since $\lim_n X_1^n(t) = X_1(t)$ in $L^p(\Omega, \dot{H}^0)$ for all $t \in [0, T]$ this implies that $X_1(t) \in \mathbb{D}^{1,p}(\dot{H}^0)$ for all $t \in [0, T]$, and that $\hat{Y}(t) = DX_1(t)$ for all $t \in [0, T]$. With this, we have deduced (3.21).

Next, we move on to the corresponding results for the approximation. Since $P^1X_{h,\Delta t}^0 \in \mathbb{D}^{1,p}(\dot{H}^0)$ for all $p \geq 2$, a proof by induction using (3.9), (3.18) and (3.19) shows that $P^1X_{h,\Delta t}^n = \tilde{X}_1(t_n) \in \mathbb{D}^{1,p}(\dot{H}^0)$ for all $p \geq 2, n \in \{0, 1, \dots, N_{\Delta t}\}$ and we have

$$DP^1X_{h,\Delta t}^n = \Delta t \sum_{j=0}^{n-1} P^1E_{h,\Delta t}^{n-j} P_h BF(t_j, P^1X_{h,\Delta t}^j) DP^1X_{h,\Delta t}^j + \sum_{j=0}^{n-1} \chi_{(t_j, t_{j+1}]}(\cdot) P^1E_{h,\Delta t}^{n-j} P_h B.$$

Therefore, (3.22) follows by Lemma 3.4 and we omit the details.

The results (3.23) and (3.24) are deduced in the same way, by noting that if $\text{Tr}(Q) < \infty$, then $\mathcal{L}(\dot{H}^0, \dot{H}^1) \subset \mathcal{L}(\dot{H}^0) \subset \mathcal{L}_2^0$ (cf. [24], Lem. 2.3.7) and by using $\alpha = 1$ instead of $\alpha = 0$ in (3.26). \square

The following error representation (cf. [21], Thm. 5.9) is a direct consequence of the mean value theorem, (3.17), (3.20), and the facts that $D_G \phi(P^1 \cdot) = (P^1)^* \phi'(P^1 \cdot)$ and $D_G^2 \phi(P^1 \cdot) = (P^1)^* \phi''(P^1 \cdot) P^1$, where D_G denotes the Gâteaux derivative and $(P^1)^*$ the adjoint of P^1 .

Proposition 3.9. *Let Assumptions 2.2, 3.3 and 3.6 hold. Let X be the mild solution given by (2.13) of the stochastic wave equation and let \tilde{X} be the fully discrete approximation given by (3.10). Then, the weak error of the approximation satisfies*

$$\begin{aligned} \mathbb{E} [\phi(\tilde{X}_1(T)) - \phi(X_1(T))] &= \int_0^1 \mathbb{E} \left[\left\langle \phi' \left(X_1(T) + s(\tilde{X}_1(T) - X_1(T)) \right), P^1 \left(\tilde{E}(T) - E(T) \right) x_0 \right. \right. \\ &\quad \left. \left. + \int_0^T P^1 \left(\tilde{E}(T-r) BF(\lfloor r \rfloor_{\Delta t}, \tilde{X}_1(\lfloor r \rfloor_{\Delta t})) - E(T-r) BF(r, X_1(r)) \right) dr \right\rangle_{\dot{H}^0} \right] ds \\ &\quad + \int_0^1 \int_0^T \mathbb{E} \left[\left\langle \phi'' \left(X_1(T) + s(\tilde{X}_1(T) - X_1(T)) \right) \left((1-s) D_r X_1(T) + s D_r \tilde{X}_1(T) \right), \right. \right. \\ &\quad \left. \left. P^1 \left(\tilde{E}(T-r) - E(T-r) \right) B \right\rangle_{\mathcal{L}_2^0} \right] dr ds. \end{aligned}$$

We are now equipped to show a weak convergence result.

Theorem 3.10 (Weak convergence). *Let X be the mild solution given by (2.13) of the stochastic wave equation and let \tilde{X} be the fully discrete approximation given by (3.10). Suppose that Assumptions 2.2, 3.1, 3.3 and 3.6 all hold and let $r' = \min(\max(2\nu, \beta), 1 + \theta, \delta)$. Then, for $\mu \leq 1$, there exists a constant $C > 0$ such that, for all $h, \Delta t \in (0, 1]$,*

$$\left| \mathbb{E} [\phi(X_1(T)) - \phi(\tilde{X}_1(T))] \right| \leq C(h^{r' \frac{\kappa}{\kappa+1}} + \Delta t^{\min(r' \frac{\rho}{\rho+1}, \eta, 1)}).$$

If, on the other hand, $1 < \mu \leq 2$, then there exists a constant $C > 0$ such that, for all $h, \Delta t \in (0, 1]$,

$$\left| \mathbb{E} [\phi(X_1(T)) - \phi(\tilde{X}_1(T))] \right| \leq C(h^{r' \frac{\kappa}{\kappa+1} + 1 - \mu} + h^{1-\mu} \Delta t^{\min(r' \frac{\rho}{\rho+1}, \eta, 1)}).$$

Proof. We first prove the theorem under Assumption 3.6(iii). Writing

$$\hat{\phi}_1(s) = \phi' \left(X_1(T) + s(\tilde{X}_1(T) - X_1(T)) \right)$$

and

$$\hat{\phi}_2(s) = \phi'' \left(X_1(T) + s(\tilde{X}_1(T) - X_1(T)) \right),$$

we use Proposition 3.9 to split the weak error

$$\begin{aligned} \left| \mathbb{E} [\phi(\tilde{X}_1(T)) - \phi(X_1(T))] \right| &\leq \int_0^1 \mathbb{E} \left[\left| \left\langle \hat{\phi}_1(s), P^1 \left(\tilde{E}(T) - E(T) \right) x_0 \right\rangle_{\dot{H}^0} \right| \right] ds \\ &\quad + \int_0^1 \mathbb{E} \left[\left| \left\langle \hat{\phi}_1(s), \int_0^T P^1 \tilde{E}(T-r) BF(\lfloor r \rfloor_{\Delta t}, \tilde{X}_1(\lfloor r \rfloor_{\Delta t})) \right. \right. \right. \\ &\quad \left. \left. - P^1 E(T-r) BF(r, X_1(r)) dr \right\rangle_{\dot{H}^0} \right| \right] ds \\ &\quad + \int_0^1 \int_0^T \mathbb{E} \left[\left| \left\langle \hat{\phi}_2(s) (1-s) D_r X_1(T), P^1 \left(\tilde{E}(T-r) - E(T-r) \right) B \right\rangle_{\mathcal{L}_2^0} \right| \right] dr ds \\ &\quad + \int_0^1 \int_0^T \mathbb{E} \left[\left| \left\langle \hat{\phi}_2(s) s D_r \tilde{X}_1(T), P^1 \left(\tilde{E}(T-r) - E(T-r) \right) B \right\rangle_{\mathcal{L}_2^0} \right| \right] dr ds \\ &= \text{I} + \text{II} + \text{III} + \text{IV}. \end{aligned}$$

First we note that as a consequence of (2.14), (3.13) and Assumption 3.6(i),

$$\sup_{\substack{s \in [0,1] \\ h, \Delta t \in (0,1]}} \|\hat{\phi}_1(s)\|_{L^2(\Omega, \dot{H}^0)} < \infty.$$

Therefore, by Hölder's inequality, Lemma 3.2 and Assumption 2.2(iii),

$$I \leq \sup_{\substack{s \in [0,1] \\ h, \Delta t \in (0,1)}} \|\hat{\phi}_1(s)\|_{L^2(\Omega, \dot{H}^0)} \|P^1(\tilde{E}(T) - E(T))\Theta^{-\frac{\delta}{2}}\|_{\mathcal{L}(\mathcal{H}, \dot{H}^0)} \|\Theta^{\frac{\delta}{2}}x_0\|_{\mathcal{H}} \lesssim h^{\min(\delta \frac{\kappa}{\kappa+1}, \kappa)} + \Delta t^{\min(\delta \frac{\rho}{\rho+1}, 1)}.$$

By the same arguments, we obtain

$$\begin{aligned} \Pi &\lesssim \int_0^T \|P^1 \tilde{E}(T-r)BF(\lfloor r \rfloor_{\Delta t}, \tilde{X}_1(\lfloor r \rfloor_{\Delta t})) - P^1 E(T-r)BF(r, X_1(r))\|_{L^2(\Omega, \dot{H}^0)} dr \\ &= \sum_{j=0}^{N_{\Delta t}-1} \int_{t_j}^{t_{j+1}} \|P^1 (\tilde{E}(T-r)BF(t_j, \tilde{X}_1(t_j)) - E(T-r)BF(r, X_1(r)))\|_{L^p(\Omega, \dot{H}^0)} dr, \end{aligned}$$

and we split the integrand as follows:

$$\begin{aligned} &\|P^1 (\tilde{E}(T-r)BF(t_j, \tilde{X}_1(t_j)) - E(T-r)BF(r, X_1(r)))\|_{L^p(\Omega, \dot{H}^0)} \\ &\leq \|P^1 \tilde{E}(T-r)B(F(t_j, \tilde{X}_1(t_j)) - F(t_j, X_1(t_j)))\|_{L^p(\Omega, \dot{H}^0)} \\ &\quad + \|P^1 \tilde{E}(T-r)B(F(t_j, X_1(t_j)) - F(r, X_1(t_j)))\|_{L^p(\Omega, \dot{H}^0)} \\ &\quad + \|P^1 \tilde{E}(T-r)B(F(r, X_1(t_j)) - F(r, X_1(r)))\|_{L^p(\Omega, \dot{H}^0)} \\ &\quad + \|P^1 (\tilde{E}(T-r) - E(T-r))BF(r, X_1(r))\|_{L^p(\Omega, \dot{H}^0)} = \text{V} + \text{VI} + \text{VII} + \text{VIII}. \end{aligned}$$

Next, by (3.8), the mean value theorem and (3.11), since $\nu \leq \min(r, 1)$, it follows that

$$\begin{aligned} \text{V} &= \|P^1 \tilde{E}(T-r)P_h B(F(t_j, \tilde{X}_1(t_j)) - F(t_j, X_1(t_j)))\|_{L^p(\Omega, \dot{H}^0)} \\ &\leq \|P^1 \tilde{E}(T-r)\|_{\mathcal{L}(\mathcal{H}, \dot{H}^0)} \|P_h B \Lambda^{\frac{\mu}{2}}\|_{\mathcal{L}(\dot{H}^0, \mathcal{H})} \|\Lambda^{-\frac{\mu}{2}}(F(t_j, \tilde{X}_1(t_j)) - F(t_j, X_1(t_j)))\|_{L^p(\Omega, \dot{H}^0)} \\ &\lesssim \|P_h B \Lambda^{\frac{\mu}{2}}\|_{\mathcal{L}(\dot{H}^0, \mathcal{H})} \int_0^1 \|\Lambda^{-\frac{\mu}{2}} F'(t_j, \tau \tilde{X}_1(t_j) + (1-\tau)X_1(t_j))(\tilde{X}_1(t_j) - X_1(t_j))\|_{L^p(\Omega, \dot{H}^0)} d\tau \\ &\lesssim \|P_h B \Lambda^{\frac{\mu}{2}}\|_{\mathcal{L}(\dot{H}^0, \mathcal{H})} \|\Lambda^{-\frac{\mu}{2}}(\tilde{X}_1(t_j) - X_1(t_j))\|_{L^p(\Omega, \dot{H}^0)}, \end{aligned}$$

where we have also used Theorems 2.3 and 3.5. If $\mu \leq 1$, then, by (2.7),

$$\|P_h B \Lambda^{\frac{\mu}{2}}\|_{\mathcal{L}(\dot{H}^0, \mathcal{H})} \leq \|B \Lambda^{\frac{1}{2}}\|_{\mathcal{L}(\dot{H}^0, \mathcal{H})} \|\Lambda^{\frac{\mu-1}{2}}\|_{\mathcal{L}(\dot{H}^0)} \lesssim 1,$$

while if $1 < \mu \leq 2$, by (3.6) and (2.7), it follows that

$$\|P_h B \Lambda^{\frac{\mu}{2}}\|_{\mathcal{L}(\dot{H}^0, \mathcal{H})} = \|P_h \Theta^{\frac{\mu-1}{2}} B \Lambda^{\frac{1}{2}}\|_{\mathcal{L}(\dot{H}^0, \mathcal{H})} \leq \|P_h \Theta^{\frac{\mu-1}{2}}\|_{\mathcal{L}(\mathcal{H})} \|B \Lambda^{\frac{1}{2}}\|_{\mathcal{L}(\dot{H}^0, \mathcal{H})} \lesssim h^{1-\mu}.$$

Next, we similarly derive by (3.8) and (2.7), that

$$\text{VI} \lesssim \|\Lambda^{-\frac{1}{2}}(F(t_j, X_1(t_j)) - F(r, X_1(t_j)))\|_{L^p(\Omega, \dot{H}^0)} \lesssim (1 + \|X_1(t_j)\|_{L^p(\Omega, \dot{H}^0)}) |t_j - r|^\eta \lesssim |t_j - r|^\eta,$$

where we applied Theorem 2.3 and Assumption 2.2(ii) in the last inequality. Term VII is treated like V, and thus, if $\mu \leq 1$,

$$\text{VII} \lesssim (1 + \|\Lambda^{\frac{\mu}{2}} X_1(t_j)\|_{L^p(\Omega, \dot{H}^0)} + \|\Lambda^{\frac{\mu}{2}} X_1(r)\|_{L^p(\Omega, \dot{H}^0)}) \|\Lambda^{-\frac{\mu}{2}}(X_1(t_j) - X_1(r))\|_{L^p(\Omega, \dot{H}^0)} \lesssim |t_j - r|^{\min(2\mu, 1)},$$

using Theorem 2.3 with $\nu \leq r$ and Theorem 2.4 in the last step. On the other hand, if $1 < \mu \leq 2$,

$$\text{VII} \lesssim h^{1-\mu} |t_j - r|^{\min(2\nu, 1)}.$$

Term VIII is handled by (2.7), Assumption 2.2(ii), Lemma 3.2 and Theorem 2.3 yielding the estimate, since $\theta \leq r$,

$$\begin{aligned} \text{VIII} &\leq \|P^1 \left(\tilde{E}(T-r) - E(T-r) \right) B \Lambda^{-\frac{\theta}{2}}\|_{\mathcal{L}(\dot{H}^0)} \|\Lambda^{\frac{\theta}{2}} F(t_j, X_1(r))\|_{L^p(\Omega, \dot{H}^0)} \\ &\leq \|P^1 \left(\tilde{E}(T-r) - E(T-r) \right) \Theta^{-\frac{1+\theta}{2}}\|_{\mathcal{L}(\mathcal{H}, \dot{H}^0)} \left(1 + \|\Lambda^{\frac{\theta}{2}} X_1(r)\|_{L^p(\Omega, \dot{H}^0)} \right) \\ &\lesssim h^{(1+\theta)\frac{\kappa}{\kappa+1}} + \Delta t^{\min((1+\theta)\frac{\rho}{\rho+1}, 1)}. \end{aligned}$$

In summary, if $\mu \leq 1$, we get for II, using also Theorem 3.5,

$$\begin{aligned} \text{II} &\lesssim h^{(1+\theta)\frac{\kappa}{\kappa+1}} + \Delta t^{\min((1+\theta)\frac{\rho}{\rho+1}, 2\nu, \eta, 1)} + \Delta t \sum_{j=0}^{N_{\Delta t}-1} \|\tilde{X}_1(t_j) - X_1(t_j)\|_{L^2(\Omega, \dot{H}^{-\nu})} \\ &\lesssim h^{r'\frac{\kappa}{\kappa+1}} + \Delta t^{\min(r'\frac{\rho}{\rho+1}, \eta, 1)}. \end{aligned}$$

For $1 < \mu \leq 2$, we instead obtain

$$\text{II} \lesssim h^{r'\frac{\kappa}{\kappa+1}+1-\mu} + h^{1-\mu} \Delta t^{\min(r'\frac{\rho}{\rho+1}, \eta, 1)}.$$

We now continue with term III, which by (2.4), (2.3), (2.1) and (2.2) satisfies

$$\begin{aligned} \text{III} &= \int_0^1 \int_0^T \mathbb{E} \left[\left| \text{Tr} \left((1-s) P^1 \left(\tilde{E}(T-r) - E(T-r) \right) B Q \left(\hat{\phi}_2(s) D_r X_1(T) \right)^* \right) \right| \right] dr ds \\ &= \int_0^1 \int_0^T (1-s) \mathbb{E} \left[\left| \text{Tr} \left(\hat{\phi}_2(s)^* P^1 \left(\tilde{E}(T-r) - E(T-r) \right) B Q \left(D_r X_1(T) \right)^* \right) \right| \right] dr ds \\ &\leq \int_0^1 \int_0^T \mathbb{E} \left[\left\| \hat{\phi}_2(s)^* P^1 \left(\tilde{E}(T-r) - E(T-r) \right) B Q \left(D_r X_1(T) \right)^* \right\|_{\text{Tr}} \right] dr ds \\ &= \int_0^1 \int_0^T \mathbb{E} \left[\left\| \hat{\phi}_2(s)^* P^1 \left(\tilde{E}(T-r) - E(T-r) \right) \Theta^{-\beta} B \Lambda^{\frac{1}{2}} \Lambda^{\beta-\frac{1}{2}} Q \Lambda^{-\frac{1}{2}} \left(D_r X_1(T) \Lambda^{\frac{1}{2}} \right)^* \right\|_{\text{Tr}} \right] dr ds. \end{aligned}$$

Tonelli's theorem, Hölder's inequality and Jensen's inequality imply

$$\begin{aligned} \text{III} &\leq \int_0^1 \int_0^T \mathbb{E} \left[\left\| \hat{\phi}_2(s)^* \right\|_{\mathcal{L}(\dot{H}^0)} \left\| D_r X_1(T) \Lambda^{\frac{1}{2}} \right\|_{\mathcal{L}(\dot{H}^0)} \right] \left\| P^1 \left(\tilde{E}(T-r) - E(T-r) \right) \Theta^{-\beta} B \Lambda^{\frac{1}{2}} \Lambda^{\beta-\frac{1}{2}} Q \Lambda^{-\frac{1}{2}} \right\|_{\text{Tr}} dr ds \\ &\leq \sup_{\tau \in [0, T]} \left\| P^1 \left(\tilde{E}(T-\tau) - E(T-\tau) \right) \Theta^{-\beta} B \Lambda^{\frac{1}{2}} \Lambda^{\beta-\frac{1}{2}} Q \Lambda^{-\frac{1}{2}} \right\|_{\text{Tr}} \\ &\quad \times \mathbb{E} \left[\int_0^1 \left\| \hat{\phi}_2(s)^* \right\|_{\mathcal{L}(\dot{H}^0)} ds \int_0^T \left\| D_r X_1(T) \Lambda^{\frac{1}{2}} \right\|_{\mathcal{L}(\dot{H}^0)} dr \right] \\ &\lesssim \sup_{\tau \in [0, T]} \left\| P^1 \left(\tilde{E}(T-\tau) - E(T-\tau) \right) \Theta^{-\beta} B \Lambda^{\frac{1}{2}} \Lambda^{\beta-\frac{1}{2}} Q \Lambda^{-\frac{1}{2}} \right\|_{\text{Tr}} \\ &\quad \times \sup_{h, \Delta t \in (0, 1]} \left\| \hat{\phi}_2 \right\|_{L^2([0, 1] \times \Omega, \mathcal{L}(\dot{H}^0))} \left\| D X_1(T) \Lambda^{\frac{1}{2}} \right\|_{L^2([0, T] \times \Omega, \mathcal{L}(\dot{H}^0))} \\ &\lesssim \sup_{\tau \in [0, T]} \left\| P^1 \left(\tilde{E}(T-\tau) - E(T-\tau) \right) \Theta^{-\beta} B \Lambda^{\frac{1}{2}} \Lambda^{\beta-\frac{1}{2}} Q \Lambda^{-\frac{1}{2}} \right\|_{\text{Tr}}, \end{aligned}$$

where the final inequality follows from Lemma 3.8 and the fact that by (2.14), (3.13) and Assumption 3.6(i),

$$\sup_{h, \Delta t \in (0, 1]} \|\hat{\phi}_2\|_{L^2([0, 1] \times \Omega, \mathcal{L}(\dot{H}^0))} < \infty.$$

For any $s \in [0, T]$, by (2.1) and (2.7),

$$\begin{aligned} & \left\| P^1 \left(\tilde{E}(T - \tau) - E(T - \tau) \right) \Theta^{-\beta} B \Lambda^{\frac{1}{2}} \Lambda^{\beta - \frac{1}{2}} Q \Lambda^{-\frac{1}{2}} \right\|_{\text{Tr}} \\ & \leq \| P^1 \left(\tilde{E}(T - \tau) - E(T - \tau) \right) \Theta^{-\beta} \|_{\mathcal{L}(\mathcal{H}, \dot{H}^0)} \left\| \Lambda^{\beta - \frac{1}{2}} Q \Lambda^{-\frac{1}{2}} \right\|_{\text{Tr}}, \end{aligned}$$

and thus, by Lemma 3.2 and Assumption 3.6(iii),

$$\text{III} \lesssim h^{\min(2\beta \frac{\kappa}{\kappa+1}, \kappa)} + \Delta t^{\min(2\beta \frac{\rho}{\rho+1}, 1)}.$$

In the exact same way, one deduces that

$$\text{IV} \lesssim h^{\min(2\beta \frac{\kappa}{\kappa+1}, \kappa)} + \Delta t^{\min(2\beta \frac{\rho}{\rho+1}, 1)},$$

and since $2\beta \geq \max(2\nu, \beta)$ this finishes the proof in the case when Assumption 3.6(iii) is used.

If Assumption 3.6(iv) holds in place of Assumption 3.6(iii), terms I and II are analyzed in the same way. For III, we proceed similarly as before, except for that we use the commutativity condition on ϕ'' along with (2.3) and Lemma 3.8 to deduce that

$$\begin{aligned} \text{III} & \leq \int_0^1 \int_0^T \mathbb{E} \left[\left\| \left(\Lambda^{\frac{1}{2}} D_r X_1(T) \right)^* \hat{\phi}_2(s)^* \Lambda^{-\frac{1}{2}} P^1 \left(\tilde{E}(T - r) - E(T - r) \right) B Q \right\|_{\text{Tr}} \right] dr ds \\ & \lesssim \sup_{r \in [0, T]} \left\| \Lambda^{-\frac{1}{2}} P^1 \left(\tilde{E}(T - r) - E(T - r) \right) B Q \right\|_{\text{Tr}} \\ & \leq \text{Tr}(Q) \sup_{r \in [0, T]} \left\| \Lambda^{-\frac{1}{2}} P^1 \left(\tilde{E}(T - r) - E(T - r) \right) B \right\|_{\mathcal{L}(\dot{H}^0)}. \end{aligned}$$

Therefore, by Lemma 3.2,

$$\text{III} \lesssim h^{\min(2\beta \frac{\kappa}{\kappa+1}, \kappa)} + \Delta t.$$

Finally, term IV is treated the same way as above, which finishes the proof. \square

4. EXAMPLES AND NUMERICAL SIMULATION

In this section we outline a few examples for which our theory yields weak convergence rates that are greater than the available strong convergence rates. Continuing in the setting of the previous section, where $\dot{H}^0 = L^2(\mathcal{D})$ for a domain $\mathcal{D} \in \mathbb{R}^d$, $d = 1, 2, 3$, we recall that $G(t) = I$ for all $t \in [0, T]$.

Below, we also only consider time-independent F , so that η can be chosen arbitrarily large. We take F to be a Nemytskij operator, which for $u \in \dot{H}^0$ are given by $F(u)(x) = f(u(x))$ for *a.e.* $x \in \mathcal{D}$. Here $f: \mathbb{R} \rightarrow \mathbb{R}$ is a differentiable function such that, for a constant $C > 0$, $|f(x)| \leq C(1 + |x|)$, $|f'(x)| \leq C$ and $|f'(x) - f'(y)| \leq C|x - y|$ for all $x, y \in \mathbb{R}$. In Appendix A, we show that with these conditions on f , Assumption 2.2(ii) is fulfilled for all $\theta \in [0, 1/2)$. If it also holds that $f(0) = 0$, then the assumption is also fulfilled for $\theta \in (1/2, 1]$. Moreover, we show that the derivative of F , given by $(F'(u)v)(x) = f'(u(x))v(x)$ for $v \in \dot{H}^0$ and *a.e.* $x \in \mathcal{D}$, fulfills Assumption 3.3 for all $\nu \in [0, 1/2) \cup (1/2, 1)$ and μ such that $\mu \geq \max(\nu, d/2 + \epsilon)$ for an arbitrary small number $\epsilon > 0$.

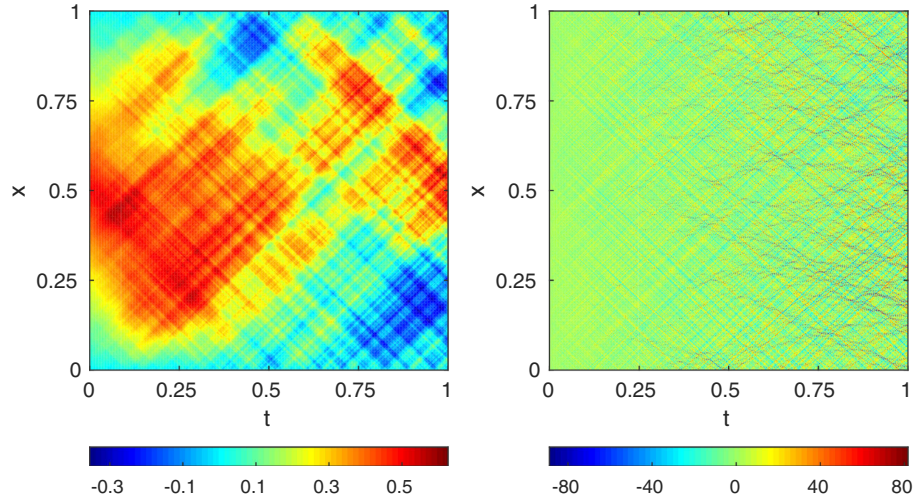


FIGURE 1. A sample of \tilde{u} (left) and \tilde{v} (right) for $\tilde{X} = [\tilde{u}, \tilde{v}]^\top$ computed with the parameters of Section 4.1 and $h = \Delta t = 2^{-9}$.

4.1. The white noise case

Suppose that $Q = I$, so that we are considering space-time white noise. For Assumption 2.2(i) to be fulfilled we then must have $d = 1$, and Assumption 3.6 is fulfilled for all $\beta < 1/2$, see Remark 4.6 of [19]. Suppose that $\delta = 1$ so that $r = \beta$, and choose $\nu = r$ maximal. For $0 < \epsilon \ll 1/2$ we set $\mu = 1/d + \epsilon = 1/2 + \epsilon > \nu$. Theorem 3.10 therefore yields the weak convergence result

$$\left| \mathbb{E} \left[\phi(X(T)) - \phi(\tilde{X}(T)) \right] \right| \lesssim h^{2\beta \frac{\kappa}{\kappa+1}} + \Delta t^{2\beta \frac{\rho}{\rho+1}}.$$

In contrast, Theorem 3.5 ensures that

$$\|\tilde{X}_1(T) - X_1(T)\|_{L^2(\Omega, \dot{H}^0)} \lesssim h^{\beta \frac{\kappa}{\kappa+1}} + \Delta t^{\beta \frac{\rho}{\rho+1}}.$$

We note that since $2\beta < 1 + \theta$, the value of θ has no influence on the convergence rate in this case.

Below we illustrate this case with $\mathcal{D} = (0, 1)$, $T = 1$. We choose $u_0(x) = x\chi_{[0, 1/2)}(x) + (1-x)\chi_{[1/2, 1)}(x)$ and $v_0(x) = \chi_{[0, 1/2)}(x)$, $x \in \mathcal{D}$. With these choices $X_0 = [u_0, v_0]^\top \in \mathcal{H}^1$. Moreover, we set $f = \cos(\cdot)$ and use piecewise linear finite elements (*i.e.*, $\kappa = 2$) and the Crank–Nicolson method (*i.e.*, $\rho = 2$) in our approximation. See Figure 1 for a sample of \tilde{X} with these parameters.

Choosing $\phi = \|\cdot\|^2$ we approximate our weak error by the Monte Carlo estimate

$$\left| E_N \left[\|P^1 X_{h, \Delta t}^{N_{\Delta t}}\|_{\dot{H}^0}^2 - \|P^1 X_{h', \Delta t'}^{N_{\Delta t'}}\|_{\dot{H}^0}^2 \right] \right| = \left| \frac{1}{N} \sum_{i=1}^N \left(\|P^1 X_{h, \Delta t}^{N_{\Delta t}}\|_{\dot{H}^0}^2 - \|P^1 X_{h', \Delta t'}^{N_{\Delta t'}}\|_{\dot{H}^0}^2 \right)^{(i)} \right|,$$

where N is the number of iid samples $\left(\|P^1 X_{h, \Delta t}^{N_{\Delta t}}\|_{\dot{H}^0}^2 - \|P^1 X_{h', \Delta t'}^{N_{\Delta t'}}\|_{\dot{H}^0}^2 \right)^{(i)}$ of $\|P^1 X_{h, \Delta t}^{N_{\Delta t}}\|_{\dot{H}^0}^2 - \|P^1 X_{h', \Delta t'}^{N_{\Delta t'}}\|_{\dot{H}^0}^2$. The strong error is approximated by

$$\left(E_N \left[\|P^1 X_{h, \Delta t}^{N_{\Delta t}} - P^1 X_{h', \Delta t'}^{N_{\Delta t'}}\|_{\dot{H}^0}^2 \right] \right)^{\frac{1}{2}}.$$

A reference solution $X_{h', \Delta t'}^{N_{\Delta t'}}$, $h', \Delta t' \in (0, 1]$, replaces the analytical solution since this is not available. We set $\Delta t = h$ and compute errors for $h = 2^{-1}, 2^{-2}, \dots, 2^{-6}$. We use a reference solution with $\Delta t' = h' = 2^{-8}$ and use

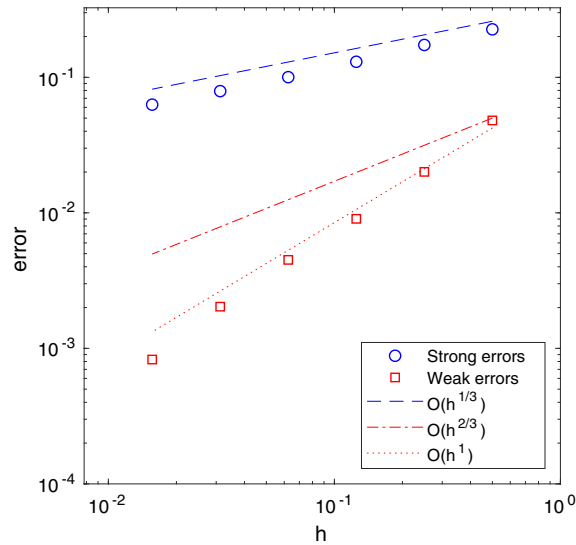


FIGURE 2. Monte Carlo estimates of strong and weak errors in the setting of Section 4.1: The white noise case.

$N = 2000$ samples in our Monte Carlo simulation. As one can see from Figure 2, the behaviour of the strong errors are consistent with our theoretical results while the weak errors appear to decay faster than expected. This is in line with [28] where numerical convergence rates of 1 were reported for a Crank–Nicolson discretization of the stochastic wave equation driven by white noise.

4.2. The trace-class noise case

If we assume that Q is of trace-class, Assumption 2.2(i) holds for all $\beta \leq 1$ but in general not for $\beta > 1$. If $\Lambda Q = Q\Lambda$, or if $\Lambda Q \neq Q\Lambda$ with ϕ as in Section 4.1, then Assumption 3.6(iii) or (iv) is fulfilled for $\beta \leq 1$, respectively. Let us take $\beta = 1$ and suppose that $\min(1 + \theta, \delta) \geq 2$ (letting $f(0) = 0$ so that $\theta = 1$), which ensures that $r = \beta = 1$. For arbitrary $0 < \epsilon \ll 1/2$, we choose $\mu = \max(d/2 + \epsilon, 1)$. In $d = 1$ we choose $\nu = \beta - \epsilon/2 = 1 - \epsilon/2 < \mu$ so that $r' = 2\nu = 2 - \epsilon$ in Theorem 3.10. Our weak convergence result in that theorem then states that

$$\left| \mathbb{E} \left[\phi(X_1(T)) - \phi(\tilde{X}_1(T)) \right] \right| \lesssim h^{(2-\epsilon)\frac{\kappa}{\kappa+1}} + \Delta t^{\min((2-\epsilon)\frac{\rho}{\rho+1}, 1)},$$

while Theorem 3.5 yields the (for sufficiently small ϵ) slower strong convergence rate

$$\|\tilde{X}_1(T) - X_1(T)\|_{L^2(\Omega, \dot{H}^0)} \lesssim h^{\frac{\kappa}{\kappa+1}} + \Delta t^{\frac{\rho}{\rho+1}}.$$

In $d = 2$, we choose β and ν as before and $\mu = 1 + \epsilon > \nu$. Our strong convergence result remains the same as in $d = 1$ while the weak rate becomes

$$\left| \mathbb{E} \left[\phi(X_1(T)) - \phi(\tilde{X}_1(T)) \right] \right| \lesssim h^{(2-\epsilon)\frac{\kappa}{\kappa+1} - \epsilon} + h^{-\epsilon} \Delta t.$$

Note that in both $d = 1$ and $d = 2$, the Crank–Nicolson scheme provides no essential benefit, in terms of the weak convergence rate, over the backward Euler scheme in this setting. In either case we have a weak rate that is essentially twice as big as the strong rate. In the case $d = 3$, however, we need to have $\mu > 3/2$, which means that we get a factor of $h^{-\frac{1}{2}}$ in Theorem 3.10. Therefore, while Theorem 3.10 still yields greater spatial

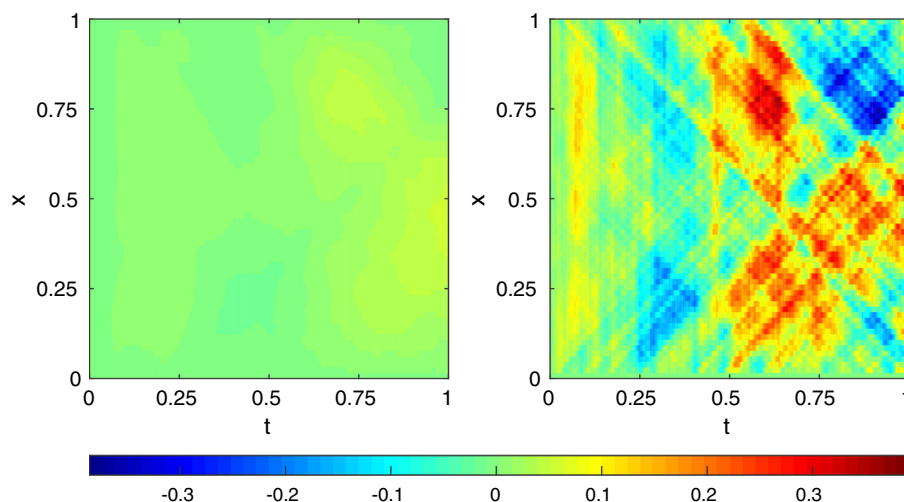


FIGURE 3. A sample of \tilde{u} (left) and \tilde{v} (right) for $\tilde{X} = [\tilde{u}, \tilde{v}]'$ computed with the parameters of Section 4.2 and $h^2 = \Delta t = 2^{-12}$.

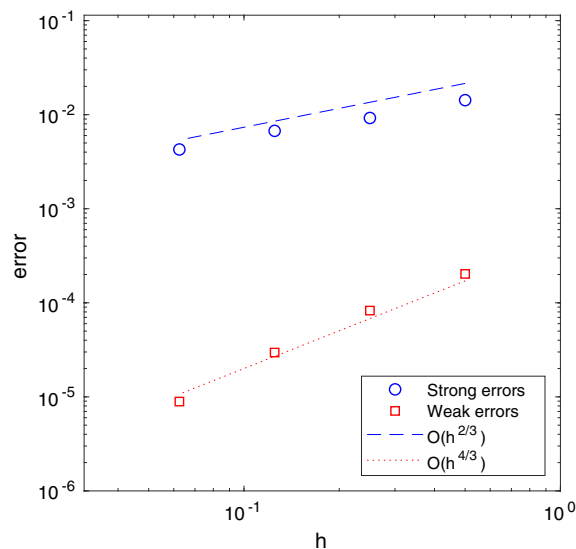


FIGURE 4. Monte Carlo estimates of strong and weak errors in the setting of Section 4.2: The trace-class noise case.

convergence rates compared to Theorem 3.5 for appropriate parameter configurations, the temporal convergence rate will be significantly lower.

In $d = 1$, we now compute weak and strong errors numerically in the setting outlined above with the same choices of \mathcal{D} , T , ρ and κ as in Section 4.1. Let Q be the integral operator defined by

$$\langle Qu, v \rangle_{\dot{H}^0} = \int_{\mathcal{D} \times \mathcal{D}} q(x, y) u(x) v(y) \, dx \, dy$$

for all $u, v \in \dot{H}^0$. We choose, for $x, y \in \mathcal{D}$, the exponential covariance kernel $q(x, y) = q(x - y) = \exp(-25|x - y|)/16$, $f(x) = \sin(x)$, and $u_0 = v_0 = 0$. See Figure 3 for a sample of \tilde{X} with these parameters.

The temporal step size is set to $\Delta t = h^2$. With this choice, we expect to see a weak and strong convergence rate of approximate order $O(h^{4/3})$ and $O(h^{2/3})$, respectively. We compute errors for $h = 2^{-1}, 2^{-2}, \dots, 2^{-5}$ and use a reference solution with $\Delta t' = h'^2 = 2^{-12}$, employing $N = 500$ samples in our Monte Carlo simulations. As one can see from Figure 4, the decay of the errors is consistent with our theoretical results.

APPENDIX A. NEMYTSKIJ OPERATORS

Let $\dot{H}^0 = L^2(\mathcal{D})$ for a convex bounded domain \mathcal{D} in \mathbb{R}^d , $d = 1, 2, 3$. For $m \in \mathbb{N}_0$, we denote by $H^m = H^m(\mathcal{D}) = W^{m,2}(\mathcal{D})$ the classical Sobolev space of order m . For $s = m + \sigma$, $m \in \mathbb{N}_0$, $\sigma \in (0, 1)$, we employ the same notation for the fractional Sobolev space H^s (see [14]) equipped with the Sobolev–Slobodeckij norm

$$\|u\|_{H^s} = \left(\|u\|_{H^m}^2 + \sum_{|\alpha|=m} \int_{\mathcal{D} \times \mathcal{D}} \frac{|D^\alpha u(x) - D^\alpha u(y)|^2}{|x - y|^{d+2\sigma}} dx dy \right)^{1/2},$$

where $u \in H^s$. The spaces $(\dot{H}^s)_{s \in [0, 2]}$ are related to $(H^s)_{s \in [0, 2]}$ by (see, e.g., [30], Thm. 4.5)

$$\dot{H}^s = \begin{cases} H^s & \text{if } s \in [0, 1/2), \\ \{u \in H^s : u = 0 \text{ a.e. on } \partial\mathcal{D}\} & \text{if } s \in (1/2, 3/2) \cup (3/2, 2], \end{cases} \quad (\text{A.1})$$

with norm equivalence.

The aim of this appendix is to show some results on *Nemytskij operators*, i.e., operators F that are for $u \in \dot{H}^0$ given by $F(u)(x) = f(u(x))$ for almost every $x \in \mathcal{D}$, where $f : \mathbb{R} \rightarrow \mathbb{R}$ is a measurable function. We assume f to be Lipschitz continuous, i.e., that there exists a constant $C > 0$ such that

$$|f(x) - f(y)| \leq C|x - y| \quad (\text{A.2})$$

for all $x, y \in \mathbb{R}$, implying the existence of a constant $C > 0$ such that

$$|f(x)| \leq C(1 + |x|) \quad (\text{A.3})$$

for all $x \in \mathbb{R}$.

If f is also a once continuously differentiable function with a bounded first derivate, i.e., if $f' : \mathbb{R} \rightarrow \mathbb{R}$ is continuous and there exists a constant $C > 0$ such that $|f'(x)| \leq C$ for all $x \in \mathbb{R}$, then $F \in \mathcal{G}_b^1(\dot{H}^0)$ (see e.g., [1], Thm. 2.7, Chap. 1) and the derivative of F at u is given by

$$(F'(u)v)(x) = f'(u(x))v(x)$$

for all $v \in H$ and almost every $x \in \mathcal{D}$.

We first show that F is Lipschitz continuous on \dot{H}^0 and that it fulfills a linear growth condition.

Proposition A.1. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a Lipschitz continuous function and let F be the corresponding Nemytskij operator. Then, there exists a constant $C > 0$ such that*

$$\|F(u) - F(v)\|_{\dot{H}^0} \leq C\|u - v\|_{\dot{H}^0} \quad (\text{A.4})$$

for all $u, v \in \dot{H}^0$ and

$$\|F(u)\|_{\dot{H}^\theta} \leq C(1 + \|u\|_{\dot{H}^\theta}) \quad (\text{A.5})$$

for all $u \in \dot{H}^\theta$, $\theta \in [0, 1/2)$. If also $f(0) = 0$, then (A.5) holds for $\theta \in (1/2, 1)$. If, in addition, f is continuously differentiable with a bounded derivative f' , then (A.5) holds for $\theta = 1$.

Proof. The inequality (A.4) is a direct consequence of (A.2) via

$$\|F(u) - F(v)\|_{\dot{H}^0}^2 = \int_{\mathcal{D}} |f(u(x)) - f(v(x))|^2 dx \leq C^2 \int_{\mathcal{D}} |u(x) - v(x)|^2 dx = C^2 \|u - v\|_{\dot{H}^0}^2.$$

For (A.5) with $\theta \in [0, 1/2)$ we also make use of (A.1) and (A.3) to see that

$$\begin{aligned} \|F(u)\|_{\dot{H}^\theta}^2 &\lesssim \|F(u)\|_{\dot{H}^\theta}^2 \\ &= \|F(u)\|_{\dot{H}^0}^2 + \int_{\mathcal{D} \times \mathcal{D}} \frac{|f(u(x)) - f(u(y))|^2}{|x - y|^{d+2\theta}} dx dy \\ &\lesssim 1 + \|u\|_{\dot{H}^0}^2 + \int_{\mathcal{D} \times \mathcal{D}} \frac{|u(x) - u(y)|^2}{|x - y|^{d+2\theta}} dx dy \\ &= 1 + \|u\|_{\dot{H}^\theta}^2 \lesssim 1 + \|u\|_{\dot{H}^\theta}^2. \end{aligned}$$

The same argument is used for (A.5) when $\theta \in (1/2, 1)$, noting that the condition $f(0) = 0$ causes $F(u)$ to inherit the boundary condition of $u \in \dot{H}^\theta$. For $\theta = 1$ we simply note that, due to the definition of $\|\cdot\|_{\dot{H}^1}$, the chain rule for weak derivatives and the assumption that $|f'(x)| \leq C$ for all $x \in \mathbb{R}$,

$$\|F(u)\|_{\dot{H}^1}^2 = \int_{\mathcal{D}} |\nabla f(u(x))|^2 dx = \int_{\mathcal{D}} |f'(u(x))|^2 |\nabla u(x)|^2 dx \leq C^2 \|u\|_{\dot{H}^1}^2,$$

which completes the proof of the proposition. \square

Next, we show that F' fulfills a negative norm bound if f' is Lipschitz continuous.

Proposition A.2. *Let f be a continuously differentiable function of at most linear growth with bounded and Lipschitz continuous derivative f' and let F be the corresponding Nemytskij operator. Then, there exists a constant $C > 0$ such that*

$$\|F'(u)v\|_{\dot{H}^0} \leq C \|v\|_{\dot{H}^0}, \quad (\text{A.6})$$

for all $u, v \in \dot{H}^0$, and

$$\|F'(u)v\|_{\dot{H}^{-\mu}} \leq C (1 + \|u\|_{\dot{H}^\nu}) \|v\|_{\dot{H}^{-\nu}}, \quad (\text{A.7})$$

for all $u \in \dot{H}^\nu$ and $v \in \dot{H}^{-\nu}$ where $\epsilon > 0$, $\nu \in [0, 1/2) \cup (1/2, 1)$ and $\mu \geq \max(d/2 + \epsilon, \nu)$. If, in addition, f' is differentiable with a bounded derivative f'' , then (A.7) holds for $\nu = 1$.

Proof. The first result is a direct consequence of the assumption that there exists a constant $C > 0$ such that $|f'(x)| \leq C$ for all $x \in \mathbb{R}$ via the estimate

$$\|F'(u)v\|_{\dot{H}^0}^2 = \int_{\mathcal{D}} |f'(u(x))v(x)|^2 dx \leq C^2 \int_{\mathcal{D}} |v(x)|^2 dx = C^2 \|v\|_{\dot{H}^0}^2$$

for all $u, v \in \dot{H}^0$. This also shows (A.7) for $\nu = 0$. For $\nu > 0$, we mimic the approach of Lemma 4.4 from [29]. Let $u \in \dot{H}^\nu$ and $v \in \dot{H}^\mu$. Then $(F'(u)v)(x) = f'(u(x))v(x) = 0$ for a.e. $x \in \partial\mathcal{D}$ since $v(x) = 0$ for a.e. $x \in \partial\mathcal{D}$, as a consequence of $\mu > 1/2$ and (A.1). We may therefore use (A.1) to obtain, if $\nu \in (0, 1/2) \cup (1/2, 1)$, that

$$\|F'(u)v\|_{\dot{H}^\nu}^2 \lesssim \|F'(u)v\|_{\dot{H}^\nu}^2 = \|F'(u)v\|_{\dot{H}^0}^2 + \int_{\mathcal{D} \times \mathcal{D}} \frac{|f'(u(x))v(x) - f'(u(y))v(y)|^2}{|x - y|^{d+2\nu}} dx dy.$$

By the Sobolev embedding theorem $\dot{H}^\mu \subset L^\infty(\mathcal{D})$ continuously. Employing this fact along with (A.6), the inequality $(a+b)^2 \leq 2(a^2+b^2)$, $a, b \in \mathbb{R}$ and the Lipschitz assumption on f' yields

$$\begin{aligned} \|F'(u)v\|_{\dot{H}^\nu}^2 &\lesssim \|v\|_{\dot{H}^0}^2 + \int_{\mathcal{D} \times \mathcal{D}} \frac{|f'(u(x))(v(x) - v(y))|^2}{|x - y|^{d+2\nu}} \, dx \, dy \\ &\quad + \int_{\mathcal{D} \times \mathcal{D}} \frac{|(f'(u(x)) - f'(u(y)))v(y)|^2}{|x - y|^{d+2\nu}} \, dx \, dy \\ &\lesssim \|v\|_{\dot{H}^0}^2 + \int_{\mathcal{D} \times \mathcal{D}} \frac{|v(x) - v(y)|^2}{|x - y|^{d+2\nu}} \, dx \, dy \\ &\quad + \|v\|_{L^\infty(\mathcal{D})}^2 \int_{\mathcal{D} \times \mathcal{D}} \frac{|u(x) - u(y)|^2}{|x - y|^{d+2\nu}} \, dx \, dy \\ &\lesssim \|v\|_{\dot{H}^\nu}^2 + \|v\|_{L^\infty(\mathcal{D})}^2 \|u\|_{\dot{H}^\nu}^2 \lesssim \|v\|_{\dot{H}^\mu}^2 (1 + \|u\|_{\dot{H}^\nu}^2). \end{aligned}$$

If $\nu = 1$ and f' is differentiable with f'' bounded, we directly use the definition of \dot{H}^1 to see that, by the same arguments as above,

$$\begin{aligned} \|F'(u)v\|_{\dot{H}^\nu}^2 &= \int_{\mathcal{D}} |\nabla(f'(u(x))v(x))|^2 \, dx \lesssim \int_{\mathcal{D}} |f'(u(x))\nabla v(x)|^2 \, dx \\ &\quad + \int_{\mathcal{D}} |f''(u(x))\nabla u(x)v(x)|^2 \, dx \\ &\lesssim \|v\|_{\dot{H}^0}^2 + \|v\|_{L^\infty(\mathcal{D})}^2 \|u\|_{\dot{H}^1}^2 \lesssim \|v\|_{\dot{H}^\mu}^2 (1 + \|u\|_{\dot{H}^\nu}^2). \end{aligned}$$

In summary,

$$\|\Lambda^{\frac{\nu}{2}} F'(u) \Lambda^{-\frac{\mu}{2}}\|_{\mathcal{L}(\dot{H}^0)}^2 = \|F'(u)\|_{\mathcal{L}(\dot{H}^\mu, \dot{H}^\nu)}^2 \lesssim 1 + \|u\|_{\dot{H}^\nu}^2$$

and thus, using that $F'(u)$ is symmetric on \dot{H}^0 , we obtain for $v \in \dot{H}^0$ that

$$\begin{aligned} \|\Lambda^{-\frac{\mu}{2}} F'(u)v\|_{\dot{H}^0} &= \sup_{\substack{w \in \dot{H}^0 \\ \|w\|_{\dot{H}^0}=1}} \left| \left\langle \Lambda^{-\frac{\mu}{2}} F'(u)v, w \right\rangle_{\dot{H}^0} \right| \\ &= \sup_{\substack{w \in \dot{H}^0 \\ \|w\|_{\dot{H}^0}=1}} \left| \left\langle \Lambda^{-\frac{\nu}{2}} v, \Lambda^{\frac{\nu}{2}} F'(u) \Lambda^{-\frac{\mu}{2}} w \right\rangle_{\dot{H}^0} \right| \\ &\leq \|v\|_{\dot{H}^{-\nu}} \|\Lambda^{\frac{\nu}{2}} F'(u) \Lambda^{-\frac{\mu}{2}}\|_{\mathcal{L}(\dot{H}^0)} \end{aligned}$$

and since \dot{H}^0 is dense in $\dot{H}^{-\nu}$, this implies that

$$\|F'(u)\|_{\mathcal{L}(\dot{H}^{-\nu}, \dot{H}^{-\mu})}^2 \lesssim 1 + \|u\|_{\dot{H}^\nu}^2,$$

which is equivalent to (A.7). □

Acknowledgements. The authors wish to express their thanks to Stig Larsson for helpful comments. MK thanks the support of the Swedish Research Council (VR) through Grant No. 2017-04274 and the Marsden Fund of the Royal Society of New Zealand through Grant No. 18-UOO-143. AL and AP thank the support of the Knut and Alice Wallenberg foundation, the Swedish Research Council under Reg. No. 621-2014-3995 and the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation.

REFERENCES

- [1] A. Ambrosetti and G. Prodi, A primer of nonlinear analysis. In: Vol. 34 of *Cambridge Studies in Advanced Mathematics*. Corrected reprint of the 1993 original. Cambridge University Press, Cambridge (1995).
- [2] A. Andersson, R. Kruse and S. Larsson, Duality in refined Sobolev–Malliavin spaces and weak approximation of SPDE. *Stoch. Partial Differ. Equ. Anal. Comput.* **4** (2016) 113–149.
- [3] R. Anton, D. Cohen, S. Larsson and X. Wang, Full discretization of semilinear stochastic wave equations driven by multiplicative noise. *SIAM J. Numer. Anal.* **54** (2016) 1093–1119.
- [4] G.A. Baker and J.H. Bramble, Semidiscrete and single step fully discrete approximations for second order hyperbolic equations. *RAIRO Anal. Numér.* **13** (1979) 75–100.
- [5] D. Blömker, Nonhomogeneous noise and Q-Wiener processes on bounded domains. *Stoch. Anal. Appl.* **23** (2005) 255–273.
- [6] D. Bolin, K. Kirchner and M. Kovács, Numerical solution of fractional elliptic stochastic PDEs with spatial white noise. *IMA J. Numer. Anal.* **12** (2018).
- [7] Y. Cao and L. Yin, Spectral Galerkin method for stochastic wave equations driven by space-time white noise. *Commun. Pure Appl. Math.* **6** (2007) 607–617.
- [8] D. Cohen and L. Quer-Sardanyons, A fully discrete approximation of the one-dimensional stochastic wave equation. *IMA J. Numer. Anal.* **36** (2015) 400–420.
- [9] D. Cohen, S. Larsson and M. Sigg, A trigonometric method for the linear stochastic wave equation. *SIAM J. Numer. Anal.* **51** (2013) 204–222.
- [10] S. Cox, A. Jentzen and F. Lindner, Weak convergence rates for temporal numerical approximations of stochastic wave equations with multiplicative noise. Preprint [arXiv:1901.05535](https://arxiv.org/abs/1901.05535) (2019).
- [11] G. Da Prato and J. Zabczyk, Stochastic equations in infinite dimensions. In: Vol. 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge (1992).
- [12] G. Da Prato and J. Zabczyk, Stochastic equations in infinite Dimensions, 2nd edition. In: Vol. 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge (2014).
- [13] R.C. Dalang, The stochastic wave equation. In: A Minicourse on Stochastic Partial Differential Equations, edited by D. Khoshnevisan and F. Rassoul-Agha. Springer, Berlin Heidelberg (2009) 39–71.
- [14] E. Di Nezza, G. Palatucci and E. Valdinoci, Hitchhikers guide to the fractional Sobolev spaces. *Bull. Sci. Math.* **136** (2012) 521–573.
- [15] R.D. Grigorieff, Diskrete approximation von eigenwertproblemen. *Numer. Math.* **25** (1975) 79–97.
- [16] E. Hausenblas, Weak approximation of the stochastic wave equation. *J. Comput. Appl. Math.* **235** (2010) 33–58.
- [17] L.J. de Naurois, A. Jentzen and T. Welti, Weak convergence rates for spatial spectral Galerkin approximations of semilinear stochastic wave equations with multiplicative noise. Preprint [arXiv:1508.05168](https://arxiv.org/abs/1508.05168) (2015).
- [18] M. Kovács, S. Larsson and F. Lindgren, Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise. *BIT Numer. Math.* **52** (2012) 85–108.
- [19] M. Kovács, S. Larsson and F. Lindgren, Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise II. Fully discrete schemes. *BIT Numer. Math.* **413** (2013) 497.
- [20] M. Kovács, F. Lindner and R.L. Schilling, Weak convergence of finite element approximations of linear stochastic evolution equations with additive Lévy noise. *SIAM-ASA J. Uncertainty Quantif.* **3** (2015) 1159–1199.
- [21] R. Kruse, Strong and weak approximation of semilinear stochastic evolution equations. In: Vol. 2093 of *Lecture Notes in Mathematics*. Springer (2014).
- [22] F. Lindgren, *On weak and strong convergence of numerical approximations of stochastic partial differential equations*. Ph.D. thesis, Chalmers University of Technology (2012).
- [23] D. Nualart, The Malliavin calculus and related topics, 2nd edition. In: *Probability and its Applications (New York)*. Springer-Verlag, Berlin (2006).
- [24] C. Prévôt and M. Röckner, A Concise course on stochastic partial differential equations. In: Vol. 1905 of *Lecture Notes in Mathematics*. Springer, Berlin (2007).
- [25] L. Quer-Sardanyons and M. Sanz-Solé, Space semi-discretisations for a stochastic wave equation. *Potential Anal.* **24** (2006) 303–332.
- [26] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems, second edition. In: Vol. 25 of *Springer Series in Computational Mathematics*. Springer (2006).
- [27] J.B. Walsh, On numerical solutions of the stochastic wave equation. *Illinois J. Math.* **50** (2006) 991–1018.
- [28] X. Wang, An exponential integrator scheme for time discretization of nonlinear stochastic wave equation. *J. Sci. Comput.* **64** (2015) 234–263.
- [29] X. Wang, S. Gan and J. Tang, Higher order strong approximations of semilinear stochastic wave equation with additive space-time white noise. *SIAM J. Sci. Comput.* **36** (2014) A2611–A2632.
- [30] A. Yagi, H^∞ -functional calculus and characterization of domains of fractional powers. In: Recent Advances in Operator Theory and Applications, edited by T. Ando, R.E. Curto, I.B. Jung and W.Y. Lee. Birkhäuser, Basel (2008) 217–235.