IMA Journal of Numerical Analysis (2018) **00**, 1–33
doi:10.1093/imanum/dry030

Error estimates of finite element method for semilinear stochastic strongly damped wave equation

Ruisheng Qi
School of Mathematics and Statistics, Northeastern University at Qinhuangdao, Qinhuangdao, China
qirsh@neuq.edu.cn

and

Xiaojie Wang*
School of Mathematics and Statistics, Central South University, Changsha, China
*Corresponding author: x.j.wang7@csu.edu.cn

[Received on 18 May 2017; revised on 15 January 2018]

In this paper we consider a semilinear stochastic strongly damped wave equation driven by additive Gaussian noise. Following a semigroup framework we establish existence, uniqueness and space-time regularity of a mild solution to such equation. Unlike the usual stochastic wave equation without damping the underlying problem with space-time white noise ($Q = I$) allows for a mild solution with a positive order of regularity in multiple spatial dimensions. Further, we analyze a spatio-temporal discretization of the problem, performed by a standard finite element method (FEM) in space and a well-known linear implicit Euler scheme in time. The analysis of the approximation error forces us to significantly enrich existing error estimates of semidiscrete and fully discrete FEMs for the corresponding linear deterministic equation. The main results show optimal convergence rates in the sense that the orders of convergence in space and in time coincide with the orders of the spatial and temporal regularity of the mild solution, respectively. Numerical examples are finally included to confirm our theoretical findings.

Keywords: strongly damped wave equation; Q-Wiener process; finite element method; linear implicit Euler scheme; strong approximation.

1. Introduction

The present work is concerned with the following semilinear stochastic evolution equation subject to additive noise, described by

$$\begin{cases} \mathrm{d}u_t = \alpha L u_t \,\mathrm{d}t + Lu \,\mathrm{d}t + F(u)\,\mathrm{d}t + \mathrm{d}W(t), & \text{in } \mathcal{D} \times (0,\ T], \\ u(\cdot,0) = \varphi,\ u_t(\cdot,0) = \psi, & \text{in } \mathcal{D}, \\ u = 0, & \text{on } \partial\mathcal{D} \times (0,\ T], \end{cases} \tag{1.1}$$

where $\mathcal{D} \subset \mathbb{R}^d$, $d = 1, 2, 3$ is a bounded, convex and polynomial domain with a boundary $\partial\mathcal{D}$ and $\alpha > 0$ is a fixed positive constant. Let $L = \sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\big(l_{ij}(x)\frac{\partial}{\partial x_j}\big)$, $x \in \mathcal{D}$ be a linear second-order elliptic operator with smooth coefficients and $\{l_{ij}\}$ being uniformly positive definite. Let $\{W(t)\}_{t\in[0,T]}$ be a (possibly cylindrical) Q-Wiener process on $\big(L_2(\mathcal{D}), \|\cdot\|, (\cdot,\cdot)\big)$, defined on a stochastic basis

$(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{t\in[0,T]})$ with respect to a normal filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ and let φ, ψ be $\mathcal{F}_0$-measurable random variables.

The considered problem (1.1) is referred to as stochastic strongly damped wave equation (SSDWE for short) thereafter. The deterministic counterpart of (1.1) finds many applications in viscoelastic theory (Fitzgibbon, 1981; Massatt, 1983; Pata & Squassina, 2005), and its linear version has been numerically studied by Larsson *et al.* (1991) and Thomée & Wahlbin (2004), where a finite element method (FEM) is used for spatial discretization and rational approximations for analytic semigroup. Particularly when $\alpha = 0$ the problem (1.1) reduces to a stochastic wave equation (SWE) without damping, numerical approximations of which have been recently investigated by many authors (Quer-Sardanyons & Sanz-Solé, 2006; Walsh, 2006; Cao & Yin, 2007; Hausenblas, 2010; Kovács *et al.*, 2012, 2013; Cohen *et al.*, 2013; Qi & Yang, 2013; Qi *et al.*, 2013; Wang *et al.*, 2014; Jacobe de Naurois *et al.*, 2015; Wang, 2015; Anton *et al.*, 2016; Cohen & Quer-Sardanyons, 2016). In contrast to the SWE case ($\alpha = 0$) the SSDWEs ($\alpha > 0$) are much less well understood, from both theoretical and numerical points of view. In the book by Da Prato & Zabczyk (2014, Example 6.25) a linear version of SSDWE with multiplicative noise was examined and its unique mild solution was verified. To the best of our knowledge regularity analysis and numerical treatment of such stochastic problem are both missing in the literature. This article aims to fill the gap and investigate the regularity properties and strong approximations (Kloeden & Platen, 1992; Jentzen & Kloeden, 2011; Kruse, 2013; Lord *et al.*, 2014) of SSDWE (1.1). In our later work (Qi & Wang, 2017) we analyzed a kind of accelerated exponential scheme for (1.1).

Reformulating (1.1) as a Cauchy problem of first order in a Hilbert space we follow the semigroup framework as in the book by Da Prato & Zabczyk (2014) to show existence, uniqueness and space-time regularity of a mild solution to (1.1). Under some standard assumptions (Assumptions 2.1–2.3) it is revealed that (see Theorem 2.4) the unique mild solution $\{u(t)\}_{t\in[0,T]}$ exhibits the Sobolev and Hölder regularity properties as follows:

$$u \in L^{\infty}\big([0,T]; L^2(\Omega; \dot{H}^{\gamma+1})\big), \qquad \sup_{t,s\in[0,T], t\neq s} \frac{\|u(t)-u(s)\|_{L^2(\Omega;\dot{H}^0)}}{|t-s|^{(\gamma+1)/2}} < \infty, \tag{1.2}$$

where, as specified later, $\dot{H}^{\delta} := D(A^{\frac{\delta}{2}})$, $\delta \subset \mathbb{R}$ and the parameter $\gamma \in [-1,1]$ satisfying $\|A^{\frac{\gamma-1}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$ quantifies the spatial correlation of the noise process (Assumption 2.2). If $\gamma \in [0,1]$ and $\psi \in L^2(\Omega; \dot{H}^{\gamma})$ then

$$u_t \in L^{\infty}\big([0,T]; L^2(\Omega; \dot{H}^{\gamma})\big), \qquad \sup_{t,s\in[0,T], t\neq s} \frac{\|u_t(t)-u_t(s)\|_{L^2(\Omega;\dot{H}^0)}}{|t-s|^{\gamma/2}} < \infty. \tag{1.3}$$

In order to achieve (1.2) and (1.3) we exploit further spatial and temporal regularity properties of the linear deterministic equation (Lemmas 2.8 and 2.9), based on an existing spatial regularity result (Lemma 2.7) in the study by Larsson *et al.* (1991). From (1.2) and (1.3) it is easy to realize that, the mild solution of (1.1) ($\alpha > 0$) enjoys higher spatial and temporal regularity than that of the usual SWE ($\alpha = 0$), which only admits a mild solution taking values in $L^{\infty}\big([0,T]; L^2(\Omega; \dot{H}^{\gamma} \times \dot{H}^{\gamma-1})\big)$ and satisfying $\sup_{t,s\in[0,T]} \frac{\|u(t)-u(s)\|_{L^2(\Omega;\dot{H}^0)}}{|t-s|^{\gamma}} < \infty$ under the same assumptions (Wang, 2015; Anton *et al.*, 2016). This benefits from smoothing effect of the analytic semigroup $\mathcal{S}(t)$ generated by the dominant linear operator $\mathcal{A}$. In particular, different from both the stochastic heat equation and the SWE, the strongly damped problem driven by space-time white noise ($Q = I$) allows for a mild solution with a positive order of regularity in multiple spatial dimensions ($d > 1$). For example, the space-time white noise case when $d = 2$ admits a mild solution $u \in L^\infty\big([0,T]; L^2(\Omega, \dot{H}^\alpha)\big)$ for any $\alpha < 1$ (consult Remark 2.5 for more details).

As the second contribution of this article we analyze the mean-square approximation errors caused by finite element spatial semidiscretization and space-time full discretization of (1.1). More precisely, we measure the discrepancy between the mild solution $(u(t), u_t(t))'$ and the finite element spatial approximation $(u_h(t), u_{h,t}(t))'$ as follows (Theorem 4.1):

$$\|u(t) - u_h(t)\|_{L^2(\Omega;\dot{H}^0)} = O\big(h^{1+\gamma}\big), \quad \|u_t(t) - u_{h,t}(t)\|_{L^2(\Omega;\dot{H}^0)} = O\big(h^\gamma\big), \tag{1.4}$$

where the parameter γ restricted to $\gamma \in [0,1]$, similarly as before, characterizes the spatial correlation of the Wiener process. By a combination of the FEM together with a linear implicit Euler–Maruyama time-stepping scheme, we also investigate a spatio-temporal discretization of (1.1). As stated in Theorem 4.3 the corresponding strong approximation error satisfies, for $\gamma \in [0,1]$,

$$\|u(t_n) - U^n\|_{L^2(\Omega;\dot{H}^0)} = O\left(h^{\gamma+1} + k^{\frac{\gamma+1}{2}}\right), \; \|u_t(t_n) - V^n\|_{L^2(\Omega;\dot{H}^0)} = O\big(h^\gamma + k^{\frac{\gamma}{2}}\big). \tag{1.5}$$

Here U^n and V^n are full-discrete approximations of $u(t_n)$ and $u_t(t_n)$, respectively. Comparing the convergence results (1.4) and (1.5) with the regularity results (1.2) and (1.3) one can readily observe that the convergence rate in space obtained here is optimal (Thomée, 2006; Kruse, 2014) in the sense that the rate of convergence in space coincides with the order of the spatial regularity of the mild solution. Moreover, the convergence rate in time coincides with the order of the temporal regularity of the exact solution. This essentially differs from the SWE setting ($\alpha = 0$), where the strong rates $O\big(h^{\frac{2}{3}\gamma} + k^{\frac{\gamma}{2}}\big)$ of the FEM coupled with the linear implicit Euler scheme are lower than orders of the spatial and temporal regularity of the mild solution (e.g., Kovács *et al.*, 2010; Kovács *et al.*, 2013).

Before proving (1.4) and (1.5) we formulate in Section 3 a rich variety of error estimates for the finite element semidiscretization and full discretization of the corresponding deterministic linear problem. Some of such error estimates can be straightforwardly derived from existing ones in the study by Larsson *et al.* (1991) by ingenious modifications or by interpolation arguments (Theorems 3.1 and 3.4). Nevertheless, we must stress that error estimates available in Larsson *et al.* (1991) are far from enough for the purpose of our error analysis. For instance, as one can see later, two completely new error estimates of integral form such as (3.16) and (3.63) are indispensable in the error analysis and their proofs turn out to be quite involved. To show the error estimate (3.16) for the semidiscretization we rely on energy arguments and interpolation theory (see the proof of Theorem 3.3). The proof of (3.63) for the full discretization is, however, more complicated and more technical. In addition to energy arguments and interpolation theory we need some further integral versions of regularity results of the linear deterministic problem as presented in Lemma 3.5. The newly derived error estimates, on the one hand, enrich existing ones for the deterministic problem. On the other hand, they enable us to establish (1.4) and (1.5) for the stochastic problem (see Section 4 for the details).

The outline of this paper is as follows. In the next section some preliminaries are collected and the well-posedness of the considered problem is elaborated. Section 3 is devoted to error estimates of semidiscrete and full-discrete FEMs for the corresponding deterministic linear problem. The main convergence results for the stochastic problem are presented in Section 4. Numerical experiments are finally performed in Section 5 to confirm the theoretical results.

2. The SSDWE

Let U and H be two separable $\mathbb{R}$-Hilbert spaces and by $\mathcal{L}(U,H)$ we denote the Banach space of all linear bounded operators from U into H and by $\mathcal{L}_2(U,H)$ the Hilbert space of all Hilbert–Schmidt operators from U into H. When $H=U$ we write $\mathcal{L}(U)=\mathcal{L}(U,U)$ and $\mathrm{HS}=\mathcal{L}_2(U,U)$ for ease of notation. Also, we denote the space of the Hilbert–Schmidt operators from $Q^{\frac{1}{2}}(U)$ to H by $\mathcal{L}_2^0:=\mathcal{L}_2(Q^{\frac{1}{2}}(U),H)$ and the corresponding norm is given by $\|\Gamma\|_{\mathcal{L}_2^0}=\|\Gamma Q^{\frac{1}{2}}\|_{\mathcal{L}_2(U,H)}$. It is well known that $\|ST\|_{\mathcal{L}_2(U,H)}\leq \|T\|_{\mathcal{L}_2(U,H)}\|S\|_{\mathcal{L}(U)}, T\in\mathcal{L}_2(U,H),\ S\in\mathcal{L}(U)$. Additionally, we denote $A=-L$ with the domain $D(A)=H^2(\mathcal{D})\cap H_0^1(\mathcal{D})$ and define

$$\dot{H}^s=D\left(A^{\frac{s}{2}}\right),\ \|v\|_{\dot{H}^s}=\left\|A^{\frac{s}{2}}v\right\|=|v|_s,\ s\in\mathbb{R}.$$

It is clear that $\dot{H}^0=L_2(\mathcal{D})$, $\dot{H}^1=H_0^1(\mathcal{D})$ and $\dot{H}^2=H^2(\mathcal{D})\cap H_0^1(\mathcal{D})$. Now we reformulate the stochastic equation (1.1) as the following abstract form:

$$\begin{cases} \mathrm{d}X(t)=-\mathcal{A}X(t)\,\mathrm{d}t+\mathbf{F}(X(t))\,\mathrm{d}t+\mathbf{B}\,\mathrm{d}W(t), & t\in(0,T],\\ X(0)=X_0, \end{cases} \tag{2.1}$$

where $X(t)=(u(t),u_t(t))'$, $X_0=(\varphi,\psi)'$ and

$$\mathcal{A}:=\begin{bmatrix} 0 & -I\\ A & \alpha A\end{bmatrix},\quad \mathbf{F}(X):=\begin{bmatrix} 0\\ F(u)\end{bmatrix}\ \text{ and }\ \mathbf{B}:=\begin{bmatrix} 0\\ I\end{bmatrix}.$$

Here $-\mathcal{A}$ generates an analytic semigroup $\mathcal{S}(t)=e^{-t\mathcal{A}}$ in $\dot{H}^s\times\dot{H}^{s-\sigma}$, $s\in\mathbb{R},\sigma\in[0,2]$ (see the study by Larsson *et al.*, 1991, Lemma 2.1). For the purpose of the existence, uniqueness and regularity of the mild solution to (2.1) we put standard assumptions on the nonlinear term F, the noise process $W(t)$ and the initial data $(\varphi,\psi)'$.

Assumption 2.1 (Nonlinearity). Let $F\colon\dot{H}^0\to\dot{H}^0$ be a deterministic mapping such that

$$\|F(x)-F(y)\|\leq K\|x-y\|,\qquad \forall\, x,y\in\dot{H}^0, \tag{2.2}$$

$$\|F(x)\|\leq K(1+\|x\|),\qquad \forall\, x\in\dot{H}^0, \tag{2.3}$$

where $K\in(0,\infty)$ is a positive constant.

Assumption 2.2 (Q-Wiener process). Let $W(t)$ be a (possibly cylindrical) Q-Wiener process on $\dot{H}^0$, with the covariance operator $Q\colon\dot{H}^0\to\dot{H}^0$ being a symmetric non-negative operator satisfying

$$\left\|A^{\frac{\gamma-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathrm{HS}}<\infty,\qquad \text{for some } \gamma\in[-1,1]. \tag{2.4}$$

Assumption 2.3 (Initial data). Let φ,ψ be $\mathcal{F}_0$-measurable and assume $(\varphi,\psi)'\in L^2(\Omega;\dot{H}^{\gamma+1})\times L^2(\Omega;\dot{H}^{\gamma-1})$.

Here we let $\mathbf{E}$ be the expectation in the probability space and let $L^2(\Omega;H)$ be the space of H-valued integrable random variables, equipped with the norm $\|v\|_{L^2(\Omega;H)} = \big(\mathbf{E}[\|v\|_H^2]\big)^{\frac{1}{2}}$. Owing to the above assumptions we have the following regularity results of the mild solution of (2.1).

THEOREM 2.4 Under Assumptions 2.1–2.3, the problem (2.1) admits a unique mild solution given by

$$X(t) = \mathcal{S}(t)X_0 + \int_0^t \mathcal{S}(t-s)\mathbf{F}(X(s))\,\mathrm{d}s + \int_0^t \mathcal{S}(t-s)\mathbf{B}\,\mathrm{d}W(s),\ t \in [0,T],\ \text{a.s.} \tag{2.5}$$

Furthermore, the mild solution $\{u(t)\}_{t\in[0,T]}$ has the following space-time regularity properties:

$$\sup_{t\in[0,T]} \|u(t)\|_{L^2(\Omega;\dot H^{\gamma+1})} \le C\left(1 + \|\varphi\|_{L^2(\Omega;\dot H^{\gamma+1})} + \|\psi\|_{L^2(\Omega;\dot H^{\gamma-1})}\right), \tag{2.6}$$

$$\|u(t)-u(s)\|_{L^2(\Omega;\dot H^{0})} \le C|t-s|^{\frac{\gamma+1}{2}}\left(1 + \|\varphi\|_{L^2(\Omega;\dot H^{\gamma+1})} + \|\psi\|_{L^2(\Omega;\dot H^{\gamma-1})}\right). \tag{2.7}$$

Additionally, if Assumption 2.2 is satisfied with $\gamma \in [0,1]$ and $\psi \in L^2(\Omega;\dot H^{\gamma})$ then

$$\sup_{t\in[0,T]} \|u_t(t)\|_{L^2(\Omega;\dot H^{\gamma})} \le C\big(1 + \|\varphi\|_{L^2(\Omega;\dot H^{\gamma})} + \|\psi\|_{L^2(\Omega;\dot H^{\gamma})}\big), \tag{2.8}$$

$$\|u_t(t)-u_t(s)\|_{L^2(\Omega;\dot H^{0})} \le C|t-s|^{\frac{\gamma}{2}}\big(1 + \|\varphi\|_{L^2(\Omega;\dot H^{\gamma})} + \|\psi\|_{L^2(\Omega;\dot H^{\gamma})}\big). \tag{2.9}$$

REMARK 2.5 We highlight that the mild solution $\{u(t)\}_{t\in[0,T]}$ of (1.1) driven by space-time white noise ($Q = I$) can enjoy a positive order of regularity in multiple spatial dimensions ($d > 1$). To see this we first note that Assumption 2.2 holds in the sense that $\big\|A^{\frac{\gamma-1}{2}}\big\|_{\mathrm{HS}} < \infty$ for $-1 \le \gamma < \frac{2-d}{2}$, $d = 1, 2, 3$, by taking the asymptotics of the eigenvalues of A into account. If $d \in \{2,3\}$ then the estimate (2.6) ensures that $\{u(t)\}_{t\in[0,T]}$ can have a positive order of spatial regularity since $\gamma+1 > 0$. As a comparison we recall that the mild solutions of the stochastic heat equation (Yan, 2005, Corollary 2.5) and the SWE (Kovács *et al.*, 2010, Remark 3.2) subject to the space-time white noise only survive in one spatial dimension.

REMARK 2.6 Here and below C denotes a generic positive constant that may vary from line to line, depending on T, K and $\|A^{\frac{\gamma-1}{2}}Q^{\frac{1}{2}}\|_{\mathrm{HS}}$, but independent of step sizes h, k. In addition, we make further comments on the initial data. Since our main interest lies in the influence due to the presence of the noise we work with smooth initial data here and below (e.g., $\varphi \in L^2(\Omega;\dot H^{\gamma+1})$, $\psi \in L^2(\Omega;\dot H^{\gamma})$). However, as indicated in the study by Larsson *et al.* (1991), such conditions can be relaxed with nonsmooth initial data, but at the cost of nonuniform error constants C blowing up as $T \to 0$.

In order to prove Theorem 2.4 we need some properties of the semigroup $\mathcal{S}(t)$, which rely on properties of the corresponding linear deterministic strongly damped wave equation

$$\begin{cases} u_{tt} + \alpha A u_t + Au = 0, & t \in (0,\ T], \\ u(0) = u_0,\ u_t(0) = v_0. \end{cases} \tag{2.10}$$

As mentioned earlier the linear problem has been examined in Larsson *et al.* (1991) and some spatial regularity results of the solution are already available there. Nevertheless, they are far from enough for our analysis in this work and we have to develop some new further regularity results. To begin with we recall the following spatial regularity result from Larsson *et al.* (1991, Lemma 2.3).

LEMMA 2.7 Let $u(t)$ be the solution of the strongly damped wave equation (2.10). For any integer $j \geq 0$, real numbers $\rho \in \mathbb{R}$ and $\sigma \in [0, 2]$ we have

$$\left|D_t^j u(t)\right|_\rho + \left|D_t^{j+1} u(t)\right|_{\rho-\sigma} \leq c\, t^{-j}(|u_0|_\rho + |v_0|_{\rho-\sigma}), \quad \text{for } t > 0. \tag{2.11}$$

Furthermore, we need the following integral versions of spatial regularity results.

LEMMA 2.8 Let $u(t)$ be the solution of the strongly damped wave equation (2.10), then

$$\alpha \int_0^t |u_t(s)|_\beta^2 \,\mathrm{d}s \leq \tfrac{1}{2}\big(|u_0|_\beta^2 + |v_0|_{\beta-1}^2\big), \quad \forall\, \beta \in \mathbb{R}, \tag{2.12}$$

$$\alpha t|u_t(t)|_\beta^2 + \int_0^t s|u_{tt}(s)|_{\beta-1}^2 \,\mathrm{d}s \leq C\big(|u_0|_\beta^2 + |v_0|_{\beta-1}^2\big), \quad \forall\, \beta \in \mathbb{R}. \tag{2.13}$$

Proof of Lemma 2.8. To prove (2.12) we multiply both sides of (2.10) by $A^{\beta-1}u_t$ to obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}r}|u_t(r)|_{\beta-1}^2 + \alpha|u_t(r)|_\beta^2 + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}r}|u(r)|_\beta^2 = 0, \tag{2.14}$$

which, after integration over $[s, t]$, suggests that

$$|u(t)|_\beta^2 + |u_t(t)|_{\beta-1}^2 + 2\alpha \int_s^t |u_t(r)|_\beta^2 \,\mathrm{d}r = |u(s)|_\beta^2 + |u_t(s)|_{\beta-1}^2. \tag{2.15}$$

Taking $s = 0$ implies (2.12) straightforwardly. To validate (2.13) we multiply (2.10) by $sA^{\beta-1}u_{tt}$ and do some manipulations to arrive at

$$s\big(u_{tt}, A^{\beta-1}u_{tt}\big) + \frac{\mathrm{d}}{\mathrm{d}s}\big(\tfrac{\alpha}{2}sAu_t + sAu, A^{\beta-1}u_t\big) = \big(s + \tfrac{\alpha}{2}\big)\big(Au_t, A^{\beta-1}u_t\big) + \big(Au, A^{\beta-1}u_t\big). \tag{2.16}$$

Similarly as before, by integration over $[0, t]$ and using (2.11) and (2.12), one can derive that

$$\begin{aligned}
\int_0^t s|u_{tt}(s)|_{\beta-1}^2 \,\mathrm{d}s + \tfrac{\alpha t}{2}|u_t(t)|_\beta^2 &= -(tu(t), A^\beta u_t(t)) + \int_0^t \Big[\big(s + \tfrac{\alpha}{2}\big)|u_t(s)|_\beta^2 + \big(u, A^\beta u_t\big)\Big] \,\mathrm{d}s \\
&\leq |u(t)|_\beta\, t|u_t(t)|_\beta + \big(T + \tfrac{\alpha+1}{2}\big) \int_0^t |u_t(s)|_\beta^2 \,\mathrm{d}s + \tfrac{1}{2}\int_0^t |u(s)|_\beta^2 \,\mathrm{d}s \\
&\leq C\big(|u_0|_\beta^2 + |v_0|_{\beta-1}^2\big) + C\big(|u_0|_\beta^2 + |v_0|_{\beta-1}^2\big) + CT\big(|u_0|_\beta^2 + |v_0|_{\beta-1}^2\big) \\
&\leq C\big(|u_0|_\beta^2 + |v_0|_{\beta-1}^2\big).
\end{aligned} \tag{2.17}$$

This thus concludes the proof of this lemma. □

Based on the above spatial regularity results one can tackle the temporal regularity properties.

LEMMA 2.9 Let $u(t)$ be the solution of equation (2.10). For $0 \leq s < t \leq T$ we have

$$\|u(t) - u(s)\| \leq C(t-s)^{\frac{\mu}{2}}(|u_0|_\mu + |v_0|_{\mu-2}), \quad \forall \mu \in [0,2], \tag{2.18}$$

$$\|u_t(t) - u_t(s)\| \leq Cs^{-\nu}(t-s)^{\nu}(|u_0|_2 + \|v_0\|), \quad \forall \nu \in [0,1], \tag{2.19}$$

$$\|u_t(t) - u_t(s)\| \leq C(t-s)^{\frac{\mu}{2}}(|u_0|_\mu + |v_0|_\mu), \quad \forall \mu \in [0,2], \tag{2.20}$$

$$\int_s^t \|u_t(r)\|^2 \, \mathrm{d}r \leq C(t-s)^{\nu}\big(|u_0|_\nu^2 + |v_0|_{\nu-1}^2\big), \quad \forall \nu \in [0,1]. \tag{2.21}$$

Proof of Lemma 2.9. Thanks to interpolation theory we only need to verify (2.18) for the two cases $\mu = 0$ and $\mu = 2$. With the aid of (2.11) with $\rho = j = 0, \sigma = 2$ one can see that

$$\|u(t) - u(s)\| \leq \|u(t)\| + \|u(s)\| \leq 2c(\|u_0\| + |v_0|_{-2}). \tag{2.22}$$

Likewise, using (2.11) with $\rho = \sigma = 2, j = 0$ leads us to

$$\|u(t) - u(s)\| \leq \int_s^t \|u_t(r)\| \, \mathrm{d}r \leq c|t-s|(|u_0|_2 + \|v_0\|). \tag{2.23}$$

With regard to (2.19) in the same manner we use (2.11) with $\rho = j = \sigma = 0$ to infer that

$$\|u_t(t) - u_t(s)\| \leq \|u_t(t)\| + \|u_t(s)\| \leq 2c(\|u_0\| + \|v_0\|). \tag{2.24}$$

At the same time, due to (2.11) with $\rho = \sigma = 2, j = 1$, we get

$$\|u_t(t) - u_t(s)\| \leq \int_s^t \|u_{tt}(r)\| \, \mathrm{d}r \leq c\int_s^t r^{-1}(|u_0|_2 + |v_0|_0) \, \mathrm{d}r \leq cs^{-1}|t-s|(|u_0|_2 + \|v_0\|). \tag{2.25}$$

To show (2.20) we recall $u_{tt} + \alpha A u_t + Au = 0$ and apply (2.11) with $\rho = 2, j = \sigma = 0$ to obtain

$$\|u_t(t) - u_t(s)\| \leq \int_s^t \|u_{tt}(r)\| \, \mathrm{d}r \leq \int_s^t (\alpha|u_t(r)|_2 + |u(r)|_2) \, \mathrm{d}r \leq C(t-s)(|u_0|_2 + |v_0|_2), \tag{2.26}$$

which combined with (2.24) implies (2.20) by interpolation. Finally, the proof of (2.21) for the cases $\nu = 0$ and $\nu = 1$ is direct consequences of (2.15) with $\beta = 0$ and (2.11), with $j = 0, \rho = \sigma = 1$, respectively. □

At this stage we are ready to associate the above regularity results with properties of the semigroup. To this end we come back to the linear problem (2.10) and reformulate it as a system of first order

$$\begin{cases} w_t + \mathcal{A}w = 0, \ \ t \in (0,T], \\ w(0) = w_0, \end{cases} \tag{2.27}$$

where we denote $v := u_t, w = (u, v)'$ and $w_0 = (u_0, v_0)'$. In terms of the semigroup $\mathcal{S}(t)$ the solution of (2.27) is given by $w(t) = (u(t), v(t))' = \mathcal{S}(t)w_0$ for $w_0 \in \dot{H}^s \times \dot{H}^{s-\sigma}, s \in \mathbb{R}, \sigma \in [0, 2]$. For two Hilbert spaces $H_i, i = 1, 2$ we additionally introduce two operators P_1 and P_2 defined by

$$P_i x = x_i, \ \forall x = (x_1, x_2)' \in H_1 \times H_2.$$

Noting that $u(t) = P_1\mathcal{S}(t)w_0$ and $v(t) = P_2\mathcal{S}(t)w_0$ one can reformulate the above regularity results in a semigroup way. For example, (2.11) with $j = 0$ can be rewritten as

$$|P_1\mathcal{S}(t)w_0|_\rho + |P_2\mathcal{S}(t)w_0|_{\rho-\sigma} \le C(|P_1w_0|_\rho + |P_2w_0|_{\rho-\sigma}), \quad \rho \in \mathbb{R}, \ \sigma \in [0, 2], \quad t > 0. \tag{2.28}$$

Moreover, Lemmas 2.8 and 2.9 suggest that

$$\alpha t|P_2\mathcal{S}(t)w_0|^2_\beta + \alpha \int_0^t |P_2\mathcal{S}(s)w_0|^2_\beta \, \mathrm{d}s \le C\big(|P_1w_0|^2_\beta + |P_2w_0|^2_{\beta-1}\big), \quad \forall \beta \in \mathbb{R}, \tag{2.29}$$

$$\|P_1\mathcal{S}(t)w_0 - P_1\mathcal{S}(s)w_0\| \le C(t-s)^{\frac{\mu}{2}}(|P_1w_0|_\mu + |P_2w_0|_{\mu-2}), \quad \forall \mu \in [0, 2], \tag{2.30}$$

$$\|P_2\mathcal{S}(t)w_0 - P_2\mathcal{S}(s)w_0\| \le Cs^{-\nu}(t-s)^\nu(|P_1w_0|_2 + \|P_2w_0\|), \quad \forall \nu \in [0, 1], \tag{2.31}$$

$$\|P_2\mathcal{S}(t)w_0 - P_2\mathcal{S}(s)w_0\| \le C(t-s)^{\frac{\mu}{2}}(|P_1w_0|_\mu + |P_2w_0|_\mu), \quad \forall \mu \in [0, 2], \tag{2.32}$$

$$\int_s^t \|P_2\mathcal{S}(r)w_0\|^2 \, \mathrm{d}r \le C(t-s)^\nu\big(|P_1w_0|^2_\nu + |P_2w_0|^2_{\nu-1}\big), \quad \forall \nu \in [0, 1]. \tag{2.33}$$

Proof of Theorem 2.4. Let $\varrho \in [-1, \gamma]$ and let $H = \dot{H}^{\varrho+1} \times \dot{H}^{\varrho-1}$ be equipped with the norm $\|X\|^2_H = |u|^2_{\varrho+1} + |v|^2_{\varrho-1}, \forall X = (u, v)' \in H$. It is easy to check that H is a separable Hilbert space. Then we show existence of a unique mild solution in H. Since Assumption 2.1 holds with $\gamma \in [-1, 1]$ and, due to the definition of $\|\cdot\|_H$ we realize that, for any $X = (u, v)'$ and $X_i = (u_i, v_i)', i = 1, 2,$

$$\|\mathbf{F}(X_1) - \mathbf{F}(X_2)\|_H = |F(u_1) - F(u_2)|_{\varrho-1} \le C\|u_1 - u_2\| \le C\|X_1 - X_2\|_H, \tag{2.34}$$

$$\|\mathbf{F}(X)\|_H = |F(u)|_{\varrho-1} \le C(1 + \|u\|) \le C(1 + \|X\|_H). \tag{2.35}$$

Additionally, the definition of the Hilbert–Schmidt norm and Assumption 2.2 enable us to deduce that

$$\|\mathbf{B}\|^2_{\mathcal{L}^2_0} = \sum_{i\in\mathbb{N}} \Big\|A^{\frac{\varrho-1}{2}}Q^{\frac{1}{2}}e_i\Big\|^2 = \Big\|A^{\frac{\varrho-1}{2}}Q^{\frac{1}{2}}\Big\|^2_{\mathrm{HS}} \le \Big\|A^{\frac{\gamma-1}{2}}Q^{\frac{1}{2}}\Big\|^2_{\mathrm{HS}} < \infty. \tag{2.36}$$

In view of Theorem 7.4 in the book by Da Prato & Zabczyk (2014), (2.34)–(2.36) together with the fact that $\mathcal{A}$ generates an analytic semigroup $\mathcal{S}(t)$ in H guarantee a unique mild solution given by (2.5) that satisfies

$$\sup_{t\in[0,T]} \Big(\|u(t)\|^2_{L^2(\Omega;\dot{H}^{\varrho+1})} + \|u_t(t)\|^2_{L^2(\Omega;\dot{H}^{\varrho-1})}\Big) \le C\Big(1 + \|\varphi\|^2_{L^2(\Omega;\dot{H}^{\varrho+1})} + \|\psi\|^2_{L^2(\Omega;\dot{H}^{\varrho-1})}\Big), \tag{2.37}$$

for all $\varrho \in [-1, \gamma]$. Taking $\varrho = \gamma$ thus confirms (2.6). As another consequence of (2.37) we have

$$\sup_{s\in[0,T]} \|F(u(s))\|_{L^2(\Omega;\dot{H}^0)} \le C\big(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^0)} + \|\psi\|_{L^2(\Omega;\dot{H}^{-2})}\big). \tag{2.38}$$

Concerning the temporal regularity we apply the Itô isometry to obtain for all $\varrho \in [-1, \gamma]$

$$\|u(t) - u(s)\|_{L^2(\Omega;\dot{H}^0)} \leq \|P_1(\mathcal{S}(t-s) - \mathcal{S}(0))X(s)\|_{L^2(\Omega;\dot{H}^0)} + \int_s^t \|P_1\mathcal{S}(t-r)\mathbf{F}(X(r))\|_{L^2(\Omega;\dot{H}^0)}\,\mathrm{d}r$$
$$+ \left(\int_s^t \left\|P_1\mathcal{S}(t-r)\mathbf{B}Q^{\frac{1}{2}}\right\|_{\mathrm{HS}}^2 \mathrm{d}r\right)^{\frac{1}{2}} := \mathbb{I}_1 + \mathbb{I}_2 + \mathbb{I}_3. \tag{2.39}$$

Combining (2.30) and (2.37) shows

$$\mathbb{I}_1 \leq C(t-s)^{\frac{\varrho+1}{2}} \left(\|P_1X(s)\|_{L^2(\Omega;\dot{H}^{\varrho+1})} + \|P_2X(s)\|_{L^2(\Omega;\dot{H}^{\varrho-1})}\right)$$
$$\leq C\left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^{\varrho+1})} + \|\psi\|_{L^2(\Omega;\dot{H}^{\varrho-1})}\right)(t-s)^{\frac{\varrho+1}{2}}. \tag{2.40}$$

For the estimate of $\mathbb{I}_2$ one can recall (2.28) with $\rho = \sigma = 0$ together with (2.38) to derive

$$\mathbb{I}_2 \leq C\int_s^t \|F(u(r))\|_{L^2(\Omega;\dot{H}^0)}\,\mathrm{d}r \leq C\left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^0)} + \|\psi\|_{L^2(\Omega;\dot{H}^{-2})}\right)(t-s). \tag{2.41}$$

To treat the remaining term $\mathbb{I}_3$ we first use the definition of the Hilbert–Schmidt norm (2.30) to get

$$\left\|P_1\mathcal{S}(t-r)\mathbf{B}Q^{\frac{1}{2}}\right\|_{\mathrm{HS}} = \left\|P_1(\mathcal{S}(0) - \mathcal{S}(t-r))\mathbf{B}Q^{\frac{1}{2}}\right\|_{\mathrm{HS}} \leq C(t-r)^{\frac{\varrho+1}{2}} \left\|A^{\frac{\varrho-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathrm{HS}}, \tag{2.42}$$

where we also used the fact $P_1\mathcal{S}(0)\mathbf{B}Q^{\frac{1}{2}}\phi = P_1\mathbf{B}Q^{\frac{1}{2}}\phi = 0$ for $\phi \in \dot{H}^0$. This yields the estimate of $\mathbb{I}_3$:

$$\mathbb{I}_3 \leq C(t-r)^{\frac{\varrho+2}{2}} \left\|A^{\frac{\varrho-1}{2}}Q^{\frac{1}{2}}\right\|_{\mathrm{HS}}. \tag{2.43}$$

Putting the above three estimates together implies

$$\|u(t) - u(s)\|_{L^2(\Omega;\dot{H}^0)} \leq C\left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^{\varrho+1})} + \|\psi\|_{L^2(\Omega;\dot{H}^{\varrho-1})}\right)(t-s)^{\frac{\varrho+1}{2}}, \tag{2.44}$$

which gives (2.7) by taking $\varrho = \gamma$. Next we shall look at the regularity of $\{u_t(t)\}_{t\in[0,T]}$ when Assumption 2.2 holds for $\gamma \in [0, 1]$. The Itô isometry ensures

$$\|u_t(t)\|_{L^2(\Omega;\dot{H}^\gamma)} \leq \|P_2\mathcal{S}(t)X_0\|_{L^2(\Omega;\dot{H}^\gamma)} + \int_0^t \left\|P_2\mathcal{S}(t-s)\mathbf{F}(X(s))\right\|_{L^2(\Omega;\dot{H}^\gamma)}\mathrm{d}s$$
$$+ \left(\int_0^t \left\|A^{\frac{\gamma}{2}}P_2\mathcal{S}(t-s)\mathbf{B}Q^{\frac{1}{2}}\right\|_{\mathrm{HS}}^2 \mathrm{d}s\right)^{\frac{1}{2}} := I_1 + I_2 + I_3. \tag{2.45}$$

In what follows, we estimate I_1, I_2, I_3 separately. The use of (2.28) with $\rho = \gamma, \sigma = 0$ guarantees that

$$I_1 \leq C\left(\|\varphi\|_{L^2(\Omega;\dot{H}^\gamma)} + \|\psi\|_{L^2(\Omega;\dot{H}^\gamma)}\right). \tag{2.46}$$

Considering (2.29) with $\beta = \gamma$ and (2.38) shows that

$$I_2 \leq C \int_0^t (t-s)^{-\frac{1}{2}} \left\| A^{\frac{\gamma-1}{2}} F(u(s)) \right\|_{L^2(\Omega;\dot{H}^0)} \, \mathrm{d}s \leq C \left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^0)} + \|\psi\|_{L^2(\Omega;\dot{H}^{-2})} \right). \tag{2.47}$$

Finally, using (2.29) with $\beta = \gamma$ and Assumption 2.2 yields

$$I_3 \leq C \left\| A^{\frac{\gamma-1}{2}} Q^{\frac{1}{2}} \right\|_{\mathrm{HS}} < \infty. \tag{2.48}$$

This together with (2.46) and (2.47) gives (2.8). To prove (2.9) it holds by (2.28), (2.32) and (2.33)

$$\begin{aligned} \|u_t(t) - u_t(s)\|_{L^2(\Omega;\dot{H}^0)} &\leq \|P_2(\mathcal{S}(t-s) - \mathcal{S}(0))X(s)\|_{L^2(\Omega;\dot{H}^0)} + \int_s^t \|P_2 \mathcal{S}(t-r)\mathbf{F}(X(r))\|_{L^2(\Omega;\dot{H}^0)} \, \mathrm{d}r \\ &\quad + \left(\int_s^t \left\| P_2 \mathcal{S}(t-r) \mathbf{B} Q^{\frac{1}{2}} \right\|_{\mathrm{HS}}^2 \mathrm{d}r \right)^{\frac{1}{2}} \\ &\leq C(t-s)^{\frac{\gamma}{2}} \left(\|P_1 X(s)\|_{L^2(\Omega;\dot{H}^\gamma)} + \|P_2 X(s)\|_{L^2(\Omega;\dot{H}^\gamma)} \right) \\ &\quad + C \int_s^t \|F(u(r))\|_{L^2(\Omega;\dot{H}^0)} \, \mathrm{d}r + C(t-s)^{\frac{\gamma}{2}} \left\| A^{\frac{\gamma-1}{2}} Q^{\frac{1}{2}} \right\|_{\mathrm{HS}} \\ &\leq C \left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^\gamma)} + \|\psi\|_{L^2(\Omega;\dot{H}^\gamma)} \right) (t-s)^{\frac{\gamma}{2}}, \end{aligned} \tag{2.49}$$

where we also used the fact $P_1 \mathbf{B} Q^{\frac{1}{2}} \phi = 0$, (2.8), (2.38) and (2.37) with $\varrho = \gamma - 1$, $\gamma \in [0,1]$. □

3. Error estimates for the finite element semidiscretization and full discretization of the deterministic linear problem

In this section we consider the semidiscrete and full-discrete finite element approximations of the deterministic linear strongly damped wave equation (2.10). A variety of error estimates will be derived, which play an important role in the mean-square convergence analysis of the FEM for the SSDWE.

For simplicity of presentation we assume that $L = \Delta$ in the following. Let $V_h \subset H_0^1(\mathcal{D})$, $h \in (0,1]$ be the space of continuous functions that are piecewise linear over the triangulation $\mathcal{T}_h$ of $\mathcal{D}$. Then we define the discrete Laplace operator $A_h \colon V_h \to V_h$ by

$$(A_h v_h, \chi_h) = a(v_h, \chi_h) := (\nabla v_h, \nabla \chi_h), \ \forall\, v_h, \chi_h \in V_h. \tag{3.1}$$

Additionally, we introduce a Riesz representation operator $\mathcal{R}_h : H_0^1(\mathcal{D}) \to V_h$ defined by

$$a(\mathcal{R}_h v, \chi_h) = a(v, \chi_h), \ \forall\, v \in H_0^1(\mathcal{D}), \ \forall\, \chi_h \in V_h \tag{3.2}$$

and a generalized projection operator $\mathcal{P}_h : \dot{H}^{-1} \to V_h$ given by

$$(\mathcal{P}_h v, \chi_h) = (v, \chi_h), \ \forall\, v \in \dot{H}^{-1}, \ \forall\, \chi_h \in V_h. \tag{3.3}$$

It is well known that (see e.g., (2.15) and (2.16) in the study by Andersson & Larsson, 2016) the operators $\mathcal{P}_h$ and $\mathcal{R}_h$ defined as above satisfy

$$\left\| A^{\frac{s}{2}}(I-\mathcal{R}_h)A^{-\frac{r}{2}} \right\|_{\mathcal{L}(\dot{H}^0)} \leq Ch^{r-s}, \quad 0 \leq s \leq 1 \leq r \leq 2, \tag{3.4}$$

$$\left\| A^{\frac{s}{2}}(I-\mathcal{P}_h)A^{-\frac{r}{2}} \right\|_{\mathcal{L}(\dot{H}^0)} \leq Ch^{r-s}, \quad 0 \leq s \leq 1, \quad 0 \leq s \leq r \leq 2. \tag{3.5}$$

Moreover, the operators A and A_h obey

$$C_1 \left\| A_h^{\frac{\gamma}{2}} \mathcal{P}_h v \right\| \leq \left\| A^{\frac{\gamma}{2}} v \right\| \leq C_2 \left\| A_h^{\frac{\gamma}{2}} \mathcal{P}_h v \right\|, \quad v \in \dot{H}^{-\gamma}, \ \gamma \in [-1, 1]. \tag{3.6}$$

Furthermore, we denote by $T\colon L_2(\mathcal{D}) \to H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$ the solution operator of equation $Au = f$ and $T_h : L_2(\mathcal{D}) \to V_h$ approximation of T, so that

$$a(T_h f, \chi_h) = (f, \chi_h), \quad \forall f \in L_2(\mathcal{D}), \chi_h \in V_h. \tag{3.7}$$

By the definition of the operator $\mathcal{R}_h$ we observe $T_h = A_h^{-1}\mathcal{P}_h = \mathcal{R}_h T$, and T_h is self-adjoint, positive semidefinite on $L_2(\mathcal{D})$ and positive definite on V_h. Furthermore, as a consequence of (3.4) we have

$$\|(T_h - T)f\| \leq Ch^s \|f\|_{s-2}, \quad f \in \dot{H}^{s-2}, \ s \in [1, 2]. \tag{3.8}$$

3.1 *Error estimates of semidiscrete scheme*

In this subsection we focus on the semidiscrete finite element approximation of the deterministic linear problem (2.10) and prove some useful estimates. We mention that such error estimates for the semidiscrete scheme will be derived based on energy arguments and some known results in the study by Larsson *et al.* (1991).

Note first that the weak variational form of (2.10) is to find $(u, v)' \in H_0^1(\mathcal{D}) \times H_0^1(\mathcal{D})$ such that

$$\begin{cases} (v_t, \vartheta_1) + \alpha(\nabla v, \nabla \vartheta_1) + (\nabla u, \nabla \vartheta_1) = 0, & \vartheta_1 \in H_0^1(\mathcal{D}), \\ (\nabla u_t, \nabla \vartheta_2) - (\nabla v, \nabla \vartheta_2) = 0, & \vartheta_2 \in H_0^1(\mathcal{D}). \end{cases} \tag{3.9}$$

The corresponding semidiscrete FEM is thus to find $(u_h, v_h)' \in V_h \times V_h$ such that

$$\begin{cases} (v_{h,t}, \vartheta_1) + \alpha(\nabla v_h, \nabla \vartheta_1) + (\nabla u_h, \nabla \vartheta_1) = 0, & \vartheta_1 \in V_h, \\ (\nabla u_{h,t}, \nabla \vartheta_2) - (\nabla v_h, \nabla \vartheta_2) = 0, & \vartheta_2 \in V_h. \end{cases} \tag{3.10}$$

In terms of the discrete Laplace operator defined by (3.1) we can equivalently write (3.10) as

$$\begin{bmatrix} u_{h,t}(t) \\ v_{h,t}(t) \end{bmatrix} + \begin{bmatrix} 0 & -I \\ A_h & \alpha A_h \end{bmatrix} \begin{bmatrix} u_h(t) \\ v_h(t) \end{bmatrix} = 0 \quad \text{with } u_h(0) = \mathcal{P}_h u_0, \ \ v_h(0) = \mathcal{P}_h v_0. \tag{3.11}$$

Similarly to (2.27) we can also reformulate it as

$$\begin{cases} w_{h,t} + \mathcal{A}_h w_h = 0, & t \in (0,T], \\ w_h(0) = w_{0h}, \end{cases} \tag{3.12}$$

where we denote $w_h(t) = (u_h(t), v_h(t))' \in V_h \times V_h, w_{0h} = (\mathcal{P}_h u_0, \mathcal{P}_h v_0)'$ and

$$\mathcal{A}_h := \begin{bmatrix} 0 & -I \\ A_h & \alpha A_h \end{bmatrix}.$$

Here $-\mathcal{A}_h$ generates analytic semigroups $\mathcal{S}_h(t)$ in $V_h \times V_h$ supplied with the norm of $L_2(\mathcal{D}) \times L_2(\mathcal{D})$ (Larsson *et al.*, 1991). Let $\mathbf{P}_h$ denote a projection operator from $\dot{H}^{-1} \times \dot{H}^{-1}$ to $V_h \times V_h$ defined by $\mathbf{P}_h x = (\mathcal{P}_h x_1, \mathcal{P}_h x_2)'$, $\forall x = (x_1, x_2) \in \dot{H}^{-1} \times \dot{H}^{-1}$. Then the solution of (3.12) can be written as $w_h(t) = \mathcal{S}_h(t)\mathbf{P}_h w_0$. The following results can be regarded as an extension of error estimates of integer order (i.e., $\beta = 2, q = 2$) in the study by Larsson *et al.* (1991) to cover intermediate cases.

THEOREM 3.1 Let $w_h(t) = (u_h(t), v_h(t))'$ and $w = (u(t), v(t))'$ be the solutions of (3.11) and (2.27), respectively. Let the setting in the beginning of Section 3 be fulfilled and define an error operator as

$$F_h(t)w_0 := w_h(t) - w(t) = (\mathcal{S}_h(t)\mathbf{P}_h - \mathcal{S}(t))w_0, \quad w_0 = (u_0, v_0)'. \tag{3.13}$$

Then it holds that

$$\|P_1 F_h(t) w_0\| \leq C h^{\beta} (|u_0|_{\beta} + |v_0|_{\beta-2}), \quad 1 \leq \beta \leq 2, \tag{3.14}$$

$$\|P_2 F_h(t) w_0\| \leq C h^{q} (|u_0|_{q} + |v_0|_{q}), \quad 0 \leq q \leq 2. \tag{3.15}$$

Proof of Theorem 3.1. The estimate (3.14) for the special case $\beta = 2$ can be found in the study by Larsson *et al.* (1991, Theorem 3.1). Following the basic lines there and taking (3.8) into account one can readily justify (3.14) for the intermediate cases $\beta \in [1,2]$. For (3.15) the case $q = 0$ and $q = 2$ can be immediately achieved by applying the stability of $\mathcal{S}(t)$ and $\mathcal{S}_h(t)\mathbf{P}_h$ in $L_2(\mathcal{D}) \times L_2(\mathcal{D})$, and Theorem 3.4 in Larsson *et al.* (1991), respectively. The interpolation theory thus results in (3.15) for the general case. □

REMARK 3.2 It is worthwhile to point out that (3.14) cannot hold for the range $\beta \in [0,1)$ because (3.8) is valid only for $s \in [1,2]$. Indeed, the projection $\mathbf{P}_h w_0$ is not well defined if $v_0 \in \dot{H}^{\beta-2}$, $\beta < 1$. Also, this comment applies to the error estimate (3.40) below for the full discretization.

Subsequently, we will present a completely new error estimate of integral form, which requires weaker regularity assumption on v_0 and cannot be derived directly from existing results in Larsson *et al.* (1991). Moreover, a nonsmooth data error estimate like (3.17) is obtained.

THEOREM 3.3 Under the assumptions stated in Theorem 3.1 it holds that

$$\|P_1 F_h(t) w_0\| + \left(\int_0^t \|P_2 F_h(s) w_0\|^2 \, \mathrm{d}s \right)^{\frac{1}{2}} \leq C h^{\beta} (|u_0|_{\beta} + |v_0|_{\beta-1}), \quad 0 \leq \beta \leq 2, \tag{3.16}$$

$$\|P_2 F_h(t) w_0\| \leq C h^{q} t^{-\frac{q-s}{2}} (|u_0|_{s+2} + |v_0|_{s}), \quad 0 \leq s \leq q \leq 2. \tag{3.17}$$

Proof of Theorem 3.3. By interpolation, we only need to show (3.16) for $\beta = 0$ and $\beta = 2$. For the case $\beta = 0$ we set $\vartheta_1 = T_h v_h$ and $\vartheta_2 = T_h u_h$ in (3.10), and add the resulting two equations to get

$$(v_{h,t}, T_h v_h) + \alpha(v_h, v_h) + (u_{h,t}, u_h) = 0, \tag{3.18}$$

where the definitions of T_h and A_h were also used. Equivalently, we can recast it as

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}s}\|u_h(s)\|^2 + \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}s}\left\|T_h^{\frac{1}{2}} v_h(s)\right\|^2 + \alpha\|v_h(s)\|^2 = 0, \tag{3.19}$$

which, after integration over $[0, t]$ and employing (3.6), lead to

$$\begin{aligned}\|u_h(t)\|^2 + \left\|T_h^{\frac{1}{2}} v_h(t)\right\|^2 + 2\alpha\int_0^t \|v_h(s)\|^2\,\mathrm{d}s &= \|u_h(0)\|^2 + \left\|T_h^{\frac{1}{2}} v_h(0)\right\|^2 \\ &= \|\mathcal{P}_h u_0\|^2 + \left\|T_h^{\frac{1}{2}}\mathcal{P}_h v_0\right\|^2 \le C\big(\|u_0\|^2 + |v_0|_{-1}^2\big).\end{aligned} \tag{3.20}$$

In the same spirit as (3.20) one can derive that

$$\|u(t)\|^2 + \left\|T^{\frac{1}{2}} v(t)\right\|^2 + 2\alpha\int_0^t \|v(s)\|^2\,\mathrm{d}s = \|u_0\|^2 + |v_0|_{-1}^2. \tag{3.21}$$

Therefore, combining (3.6), (3.20) and (3.21) enables us to get

$$\|u(t) - u_h(t)\|^2 + \|v(t) - v_h(t)\|_{-1}^2 + \int_0^t \|v - v_h\|^2\,\mathrm{d}s \le C\big(\|u_0\|^2 + |v_0|_{-1}^2\big). \tag{3.22}$$

This verifies (3.16) in the case $\beta = 0$. To prove (3.16) for $\beta = 2$ we introduce notations as follows:

$$\begin{aligned}\theta_1 &:= u_h - \mathcal{R}_h u, \quad \rho_1 = (\mathcal{R}_h - I)u, \\ \theta_2 &:= v_h - \mathcal{P}_h v, \quad \rho_2 = (\mathcal{P}_h - I)v, \\ e_1 &:= u_h - u = \theta_1 + \rho_1, \quad e_2 := v_h - v = \theta_2 + \rho_2.\end{aligned} \tag{3.23}$$

Subtracting (3.9) from (3.10) shows

$$(e_{2,t}, \vartheta_1) + \alpha(\nabla e_2, \nabla\vartheta_1) + (\nabla e_1, \nabla\vartheta_1) = 0, \quad \forall \vartheta_1 \in V_h, \tag{3.24}$$

$$(\nabla e_{1,t}, \nabla\vartheta_2) - (\nabla e_2, \nabla\vartheta_2) = 0, \quad \forall \vartheta_2 \in V_h. \tag{3.25}$$

Inserting $e_i = \theta_i + \rho_i, i = 1, 2$ thus gives

$$(\theta_{2,t}, \vartheta_1) + \alpha(\nabla\theta_2, \nabla\vartheta_1) + (\nabla\theta_1, \nabla\vartheta_1) = -(\rho_{2,t}, \vartheta_1) - \alpha(\nabla\rho_2, \nabla\vartheta_1) - (\nabla\rho_1, \nabla\vartheta_1),\ \forall \vartheta_1 \in V_h, \tag{3.26}$$

$$(\nabla\theta_{1,t}, \nabla\vartheta_2) - (\nabla\theta_2, \nabla\vartheta_2) = -(\nabla\rho_{1,t}, \nabla\vartheta_2) + (\nabla\rho_2, \nabla\vartheta_2), \quad \forall \vartheta_2 \in V_h. \tag{3.27}$$

Further, the orthonormal properties of the operators $\mathcal{R}_h$ and $\mathcal{P}_h$ help us to arrive at

$$(\theta_{2,t}, \vartheta_1) + \alpha(\nabla\theta_2, \nabla\vartheta_1) + (\nabla\theta_1, \nabla\vartheta_1) = -\alpha(\nabla\mathcal{R}_h\rho_2, \nabla\vartheta_1), \quad \forall\vartheta_1 \in V_h, \tag{3.28}$$

$$(\nabla\theta_{1,t}, \nabla\vartheta_2) - (\nabla\theta_2, \nabla\vartheta_2) = (\nabla\mathcal{R}_h\rho_2, \nabla\vartheta_2), \qquad \forall\vartheta_2 \in V_h. \tag{3.29}$$

Setting $\vartheta_1 = T_h\theta_2$, $\vartheta_2 = T_h\theta_1$, adding the resulting two equations and taking the definitions of T_h, A_h into account yield that

$$(\theta_{1,t}, \theta_1) + (\theta_{2,t}, T_h\theta_2) + \alpha(\theta_2, \theta_2) = -\alpha(\mathcal{R}_h\rho_2, \theta_2) + (\mathcal{R}_h\rho_2, \theta_1). \tag{3.30}$$

Exploiting similar arguments as before in conjunction with Cauchy–Schwarz inequality shows

$$\begin{aligned}\|\theta_1(t)\|^2 &+ \left\|T_h^{\frac{1}{2}}\theta_2(t)\right\|^2 + \alpha\int_0^t \|\theta_2(s)\|^2\,\mathrm{d}s \\ &\leq (\alpha+1)\int_0^t \|\mathcal{R}_h\rho_2(s)\|^2\,\mathrm{d}s + \int_0^t \|\theta_1(s)\|^2\,\mathrm{d}s + \|\theta_1(0)\|^2,\end{aligned} \tag{3.31}$$

where we also used the fact that $\theta_2(0) = v_h(0) - \mathcal{P}_h v(0) = 0$. Applying Gronwall's inequality yields

$$\begin{aligned}\|\theta_1(t)\|^2 + \left\|T_h^{\frac{1}{2}}\theta_2(t)\right\|^2 + \alpha\int_0^t \|\theta_2(s)\|^2\,\mathrm{d}s &\leq C\left(\int_0^t \|\mathcal{R}_h\rho_2(s)\|^2\,\mathrm{d}s + \|\theta_1(0)\|^2\right) \\ &\leq Ch^4\left(\int_0^t |u_t(s)|_2^2\,\mathrm{d}s + |u_0|_2^2\right) \leq Ch^4\big(|u_0|_2^2 + |v_0|_1^2\big),\end{aligned} \tag{3.32}$$

where we also used (2.12) and the facts that $\|\theta_1(0)\| = \|\mathcal{P}_h(\mathcal{R}_h - I)u_0\| \leq Ch^2|u_0|_2$ and that

$$\|\mathcal{R}_h\rho_2\| = \|\mathcal{R}_h(\mathcal{P}_h - I)v\| = \|\mathcal{P}_h(\mathcal{R}_h - I)v\| \leq Ch^2|v|_2. \tag{3.33}$$

Furthermore, (2.11), (2.12), (3.4) and (3.5) promise that

$$\|\rho_1(t)\|^2 + \int_0^t \|\rho_2(s)\|^2\,\mathrm{d}s \leq Ch^4|u(t)|_2^2 + Ch^4\int_0^t |v(s)|_2^2\,\mathrm{d}s \leq Ch^4\big(|u_0|_2^2 + |v_0|_1^2\big). \tag{3.34}$$

The triangle inequality shows (3.16) for $\beta = 2$ and the interpolation argument finally concludes the proof of (3.16). The assertion (3.17) can be also deduced by interpolation between $s = 0$ and $s = q$ (see, e.g., Thomée, 2006, Theorem 3.5). The case $s = q$ is an immediate consequence of (3.15). To prove (3.17) for $s = 0$ we again use interpolation arguments. The case $s = 0$, $q = 0$ is a special case of (3.15) and the case $s = 0$, $q = 2$ is covered in the study by Larsson *et al.* (1991, Theorem 3.1). □

3.2 *Error estimates of the full-discrete scheme*

In this part we turn our attention to the full-discrete finite element approximation of (2.10) with the backward Euler scheme for the time discretization. With the help of previous findings for the spatially discrete scheme we will derive some error estimates for the full-discrete scheme. We mention that the error analysis in the full-discrete setting is more involved, and one needs to explore further regularity results of the linear deterministic problem and the analysis relies heavily on energy arguments.

Let k be a time step size such that $k = \frac{T}{N}$, $N \in \mathbb{N}$ and denote $t_n = nk$, for $0 \leq n \leq N$. Applying the backward Euler scheme to the semidiscretization problem (3.11) gives the full-discrete finite element approximation. More accurately, we are to find $W^n = (U^n, V^n)' \in V_h \times V_h$ such that

$$\begin{bmatrix} U^n \\ V^n \end{bmatrix} - \begin{bmatrix} U^{n-1} \\ V^{n-1} \end{bmatrix} + k \begin{bmatrix} 0 & -I \\ A_h & \alpha A_h \end{bmatrix} \begin{bmatrix} U^n \\ V^n \end{bmatrix} = 0 \quad \text{with} \quad \begin{bmatrix} U^0 \\ V^0 \end{bmatrix} = \begin{bmatrix} \mathcal{P}_h u_0 \\ \mathcal{P}_h v_0 \end{bmatrix}, \tag{3.35}$$

or in a compact way,

$$W^n - W^{n-1} + k\mathcal{A}_h W^n = 0, \quad W^0 = \mathbf{P}_h w_0. \tag{3.36}$$

Denoting $r(\lambda) = (1+\lambda)^{-1}, \lambda > 0$ we can rewrite (3.36) in the form

$$W^n = r(k\mathcal{A}_h)^n \mathbf{P}_h w_0, \tag{3.37}$$

where the operator $r(k\mathcal{A}_h)^n$ is stable in the following sense (Larsson *et al.*, 1991, Lemma 5.2):

$$\|r(k\mathcal{A}_h)^n\|_{\mathcal{L}(\dot{H}^0 \times \dot{H}^0)} \leq C, \quad \text{for} \quad n \geq 0. \tag{3.38}$$

It is worthwhile to mention that the constant C in (3.38) depends on α and blows up as α tends to zero, that is, $r(k\mathcal{A}_h)^n$ is stable in the norm $\dot{H}^0 \times \dot{H}^0$, independent of k, h, but dependent of α. In contrast, the estimate (3.68) below implies that the operator $r(k\mathcal{A}_h)^n$ is stable unconditionally in the norm $\dot{H}^0 \times \dot{H}^{-1}$, independent of k, h and α. Next our aim is to analyze various error estimates of $W^n - w(t_n)$. As the first part we can derive the time-discrete analog of Theorem 3.1.

THEOREM 3.4 Let all the conditions in Theorem 3.1 be fulfilled and define

$$F^n_{kh} w_0 := W^n - w(t_n) = (r(k\mathcal{A}_h)^n \mathbf{P}_h - \mathcal{S}(t_n)) w_0, \quad w_0 = (u_0, v_0)', \tag{3.39}$$

where W^n and $w(t_n)$ obey (3.37) and (2.27), respectively. Then it holds that

$$\left\| P_1 F^n_{kh} w_0 \right\| \leq C \left(h^\beta + k^{\frac{\beta}{2}} \right) (|u_0|_\beta + |v_0|_{\beta-2}), \quad 1 \leq \beta \leq 2, \tag{3.40}$$

$$\left\| P_2 F^n_{kh} w_0 \right\| \leq C \left(h^q + k^{\frac{q}{2}} \right) (|u_0|_q + |v_0|_q), \quad 0 \leq q \leq 2. \tag{3.41}$$

Proof of Theorem 3.4. Owing to interpolation arguments we only need to show (3.40) for $\beta = 1$ and $\beta = 2$. The latter case $\beta = 2$ can be found in the study by Larsson *et al.* (1991, Theorem 5.3). For the case $\beta = 1$, with the spatial approximation error (3.14) at disposal, it remains to prove

$$\|P_1 F_n \mathbf{P}_h w_0\| = \|P_1 (r(k\mathcal{A}_h)^n - \mathcal{S}_h(t_n)) \mathbf{P}_h w_0\| \leq C k^{\frac{1}{2}} (|u_0|_1 + |v_0|_{-1}), \tag{3.42}$$

where we introduced an error operator $F_n \colon V_h \to V_h$ defined by

$$F_n := r(k\mathcal{A}_h)^n - \mathcal{S}_h(t_n). \tag{3.43}$$

As shown in the study by Larsson *et al.* (1991, Theorem 4.2) we have

$$\|F_n \mathcal{A}_h^{-1}\|_{\mathcal{L}(\dot{H}^0\times\dot{H}^0)} \leq Ck, \quad \text{for} \quad n \geq 1. \tag{3.44}$$

Before proceeding further we also observe that

$$P_1 F_n \mathbf{P}_h w_0 = P_1 M_h F_n \mathbf{P}_h w_0, \qquad M_h = \begin{bmatrix} I & 0 \\ 0 & T_h \end{bmatrix}, \tag{3.45}$$

and thus $\|P_1 F_n \mathbf{P}_h w_0\| = \|P_1 M_h F_n \mathbf{P}_h w_0\| \leq \|M_h F_n \mathbf{P}_h w_0\|_{L_2(\mathcal{D})\times L_2(\mathcal{D})}$. Accordingly, it suffices to focus on the estimate of $\|M_h F_n \mathbf{P}_h w_0\|_{L_2(\mathcal{D})\times L_2(\mathcal{D})}$. By denoting $\widetilde{F}_n := M_h F_n M_h^{-1}$ one can write

$$M_h F_n \mathbf{P}_h w_0 = M_h F_n M_h^{-1} M_h \mathbf{P}_h w_0 = \widetilde{F}_n M_h \mathbf{P}_h w_0. \tag{3.46}$$

Further, using eigenfunction expansions one can easily check that

$$\widetilde{F}_n = r(k\widetilde{\mathcal{A}}_h)^n - e^{-t_n \widetilde{\mathcal{A}}_h}, \quad \text{where} \quad \widetilde{\mathcal{A}}_h = M_h \mathcal{A}_h M_h^{-1} = \begin{bmatrix} 0 & -A_h \\ I & \alpha A_h \end{bmatrix}. \tag{3.47}$$

Since $\widetilde{\mathcal{A}}_h$ shares the same eigenvalues as $\mathcal{A}_h$, (3.38), (3.44) and the stability property $\|e^{-t\widetilde{\mathcal{A}}_h}\|_{\mathcal{L}(\dot{H}^0\times\dot{H}^0)} \leq C$ also hold for $\widetilde{\mathcal{A}}_h$. These facts enable us to derive from (3.46) that

$$\begin{aligned} \|M_h F_n \mathbf{P}_h w_0\|_{L_2(\mathcal{D})\times L_2(\mathcal{D})} &= \|\widetilde{F}_n \widetilde{\mathcal{A}}_h^{-1} \widetilde{\mathcal{A}}_h M_h \mathbf{P}_h w_0\|_{L_2(\mathcal{D})\times L_2(\mathcal{D})} \leq Ck \|\widetilde{\mathcal{A}}_h M_h \mathbf{P}_h w_0\|_{L_2(\mathcal{D})\times L_2(\mathcal{D})} \\ &= Ck \left\| \begin{bmatrix} 0 & -I \\ I & \alpha I \end{bmatrix} \mathbf{P}_h w_0 \right\|_{L_2(\mathcal{D})\times L_2(\mathcal{D})} \leq Ck(\|\mathcal{P}_h u_0\| + \|\mathcal{P}_h v_0\|), \end{aligned} \tag{3.48}$$

and that

$$\begin{aligned} \|M_h F_n \mathbf{P}_h w_0\|_{L_2(\mathcal{D})\times L_2(\mathcal{D})} &= \|\widetilde{F}_n M_h \mathbf{P}_h w_0\|_{L_2(\mathcal{D})\times L_2(\mathcal{D})} \leq C \|M_h \mathbf{P}_h w_0\|_{L_2(\mathcal{D})\times L_2(\mathcal{D})} \\ &\leq C\big(\|\mathcal{P}_h u_0\| + \|A_h^{-1} \mathcal{P}_h v_0\|\big). \end{aligned} \tag{3.49}$$

Then by interpolation we obtain

$$\|P_1 F_n \mathbf{P}_h w_0\| \leq \|M_h F_n \mathbf{P}_h w_0\|_{L_2(\mathcal{D})\times L_2(\mathcal{D})} \leq Ck^{\frac{s}{2}} \left(\left\| A_h^{\frac{s}{2}} \mathcal{P}_h u_0 \right\| + \left\| A_h^{\frac{s-2}{2}} \mathcal{P}_h v_0 \right\| \right), \text{ for } 0 \leq s \leq 2, \tag{3.50}$$

which, after assigning $s = 1$ and using (3.6), imply (3.42), that is,

$$\|P_1(r(k\mathcal{A}_h)^n - \mathcal{S}_h(t_n))\mathbf{P}_h w_0\| \leq Ck^{\frac{1}{2}} \left(\left\| A_h^{\frac{1}{2}} \mathcal{P}_h u_0 \right\| + \left\| A_h^{-\frac{1}{2}} \mathcal{P}_h v_0 \right\| \right) \leq Ck^{\frac{1}{2}} (|u_0|_1 + |v_0|_{-1}). \tag{3.51}$$

With regard to (3.41) the case $q = 0$ can be directly obtained by the stability of $e^{-t\mathcal{A}}$ and $r(k\mathcal{A}_h)^n$, and the case $q = 2$ is available in the study by Larsson *et al.* (1991, Theorem 5.1). Again, the interpolation gives (3.41). □

Similarly to the semidiscrete problem as before we expect a time-discrete analog of Theorem 3.3, which requires weaker regularity assumption on v_0 than Theorem 3.4 does. However, this is not an easy job and, as one can see below, the proof becomes much more involved. First, we need further regularity results of the linear strongly damped wave equation (2.10).

LEMMA 3.5 Let $u(t)$ be the solution of the strongly damped wave equation (2.10), then it holds that

$$\int_0^t |u_{tt}(s)|_\beta^2 \,\mathrm{d}s \le C\big(|u_0|_{\beta+2}^2 + |v_0|_{\beta+1}^2\big), \quad \beta \in \mathbb{R}, \tag{3.52}$$

$$\int_0^t s^2 |u_{ttt}(s)|_\beta^2 \,\mathrm{d}s \le C\big(|u_0|_{\beta+2}^2 + |v_0|_{\beta+1}^2\big), \quad \beta \in \mathbb{R}, \tag{3.53}$$

$$\int_0^t s^2 |u_{tt}(s)|_\beta^2 \,\mathrm{d}s \le C\big(|u_0|_{\beta}^2 + |v_0|_{\beta-1}^2\big), \quad \beta \in \mathbb{R}. \tag{3.54}$$

Proof of Lemma 3.5. In order to prove (3.52) we multiply both sides of (2.10) by $A^\beta u_{tt}$ to obtain

$$|u_{tt}(s)|_\beta^2 + \frac{\alpha}{2}\frac{\mathrm{d}}{\mathrm{d}s}|u_t(s)|_{\beta+1}^2 = -\big(Au(s), A^\beta u_{tt}(s)\big). \tag{3.55}$$

Integration over $[0,t]$ and the Cauchy–Schwarz inequality yield

$$\int_0^t |u_{tt}(s)|_\beta^2 \,\mathrm{d}s + \frac{\alpha}{2}|u_t(t)|_{\beta+1}^2 \le \frac{\alpha}{2}|v_0|_{\beta+1}^2 + \frac{1}{2}\int_0^t |u(s)|_{\beta+2}^2 \,\mathrm{d}s + \frac{1}{2}\int_0^t |u_{tt}(s)|_\beta^2 \,\mathrm{d}s. \tag{3.56}$$

Further, using (2.11) with $\rho = \beta+2, \sigma = 1$ gives

$$\int_0^t |u_{tt}(s)|_\beta^2 \,\mathrm{d}s + \alpha|u_t(t)|_{\beta+1}^2 \le \alpha|v_0|_{\beta+1}^2 + \int_0^t |u(s)|_{\beta+2}^2 \,\mathrm{d}s \le C\big(|u_0|_{\beta+2}^2 + |v_0|_{\beta+1}^2\big). \tag{3.57}$$

To confirm (3.53) we differentiate (2.10) with respect to t and multiply both sides by $s^2 A^\beta u_{ttt}$ to get

$$s^2\big(u_{ttt}, A^\beta u_{ttt}\big) + \alpha s^2\big(Au_{tt}, A^\beta u_{ttt}\big) + s^2\big(Au_t, A^\beta u_{ttt}\big) = 0, \tag{3.58}$$

which can be equivalently written as

$$s^2\big(u_{ttt}, A^\beta u_{ttt}\big) + \frac{\alpha}{2}\frac{\mathrm{d}}{\mathrm{d}s}s^2\big(Au_{tt}, A^\beta u_{tt}\big) = -\frac{\mathrm{d}}{\mathrm{d}s}s^2\big(Au_t, A^\beta u_{tt}\big) + \big(s(\alpha+s)Au_{tt} + 2sAu_t, A^\beta u_{tt}\big). \tag{3.59}$$

Integration over $[0,t]$ and using the Cauchy–Schwarz inequality suggest that

$$\int_0^t s^2|u_{ttt}(s)|_\beta^2 \,\mathrm{d}s + \frac{\alpha}{2}t^2|u_{tt}(t)|_{\beta+1}^2 \le t^2|u_t(t)|_{\beta+1}|u_{tt}(t)|_{\beta+1} + \int_0^t \big[(\alpha+2s)s|u_{tt}(s)|_{\beta+1}^2 + |u_t(s)|_{\beta+1}^2\big]\,\mathrm{d}s.$$

Applying Gronwall's inequality and taking Lemmas 2.7 and 2.8 into consideration show that

$$\begin{aligned}\int_0^t s^2|u_{ttt}(s)|_\beta^2\,\mathrm{d}s + t^2|u_{tt}(t)|_{\beta+1}^2 &\le C\Big(t^2|u_t(t)|_{\beta+1}|u_{tt}(t)|_{\beta+1} + \int_0^t \big[s|u_{tt}(s)|_{\beta+1}^2 + |u_t(s)|_{\beta+1}^2\big]\,\mathrm{d}s\Big)\\ &\le C\big(|u_0|_{\beta+2}^2 + |v_0|_{\beta+1}^2\big). \end{aligned} \tag{3.60}$$

This validates (3.53). For the estimate of (3.54), we, similarly as before, differentiate (2.10) with respect to t and multiply both sides by $sA^{\frac{\beta}{2}-1}$ to acquire

$$s\alpha A^{\frac{\beta}{2}}u_{tt} = -sA^{\frac{\beta}{2}-1}u_{ttt} - sA^{\frac{\beta}{2}}u_t. \tag{3.61}$$

Squaring both sides before integration over $[0,t]$ and combining (2.11) and (3.53) lead us to

$$\alpha^2\int_0^t s^2|u_{tt}(s)|_\beta^2\,\mathrm{d}s \le 2\int_0^t s^2|u_{ttt}(s)|_{\beta-2}^2\,\mathrm{d}s + 2\int_0^t s^2|u_t(s)|_\beta^2\,\mathrm{d}s \le C\big(|u_0|_\beta^2 + |v_0|_{\beta-1}^2\big). \tag{3.62}$$

This completes the proof of this lemma. □

Now we are ready to formulate the time-discrete analog of Theorem 3.3.

THEOREM 3.6 Under the assumptions of Theorem 3.4 it holds that

$$\left\|P_1F_{kh}^n w_0\right\| + \left(k\sum_{m=1}^n \left\|P_2F_{kh}^m w_0\right\|^2\right)^{\frac{1}{2}} \le C\left(h^\beta + k^{\frac{\beta}{2}}\right)(|u_0|_\beta + |v_0|_{\beta-1}),\ 0\le\beta\le 2, \tag{3.63}$$

$$\left\|P_2F_{kh}^n w_0\right\| \le C\left(h^q + k^{\frac{q}{2}}\right)t_n^{-\frac{q-s}{2}}(|u_0|_{s+2} + |v_0|_s),\ 0\le s\le q\le 2. \tag{3.64}$$

Proof of Theorem 3.6. Since the proof of (3.64) is easy we do this first. The case $s=q$ with $s\in[0,2]$ is a direct consequence of (3.41). In addition, the case $s=0, q=2$ can be found in the study by Larsson *et al.* (1991, Theorem 5.4). Similarly to the proof of (3.17) the desired intermediate case is obvious by interpolation. In what follows we focus on the proof of (3.63). Note first that the full-discrete weak variational form of (3.35) is to find $(U^n,V^n)'\in V_h\times V_h$ such that

$$\begin{aligned}\big(\overline{\partial}_n V^n,\chi_1\big) + \alpha(\nabla V^n,\nabla\chi_1) + (\nabla U^n,\nabla\chi_1) = 0,\ \forall\chi_1\in V_h,\\ \big(\nabla\overline{\partial}_n U^n,\nabla\chi_2\big) - (\nabla V^n,\nabla\chi_2) = 0,\ \forall\chi_2\in V_h,\end{aligned} \tag{3.65}$$

where $\overline{\partial}_n V^n := \frac{V^n - V^{n-1}}{k}$. Once again, we use interpolation arguments to obtain (3.63). For the case $\beta=0$ setting $\chi_1 = T_hV^n$, $\chi_2 = T_hU^n$ in (3.65) and adding the resulting two equations give

$$\big(\overline{\partial}_n U^n, U^n\big) + \big(\overline{\partial}_n V^n, T_hV^n\big) + \alpha\left(V^n,V^n\right) = 0. \tag{3.66}$$

Observing that $(\bar{\partial}_n U^n, U^n) \ge \frac{1}{2k}(\|U^n\|^2 - \|U^{n-1}\|^2)$ and $(\bar{\partial}_n V^n, T_h V^n) \ge \frac{1}{2k}\left(\|T_h^{\frac{1}{2}} V^n\|^2 - \|T_h^{\frac{1}{2}} V^{n-1}\|^2\right)$ by the Cauchy–Schwarz inequality we derive from (3.66) that

$$\|U^n\|^2 - \left\|U^{n-1}\right\|^2 + \left\|T_h^{\frac{1}{2}} V^n\right\|^2 - \left\|T_h^{\frac{1}{2}} V^{n-1}\right\|^2 + 2\alpha k\|V^n\|^2 \le 0. \tag{3.67}$$

By summation on n and noting $U^0 = \mathcal{P}_h u_0, V^0 = \mathcal{P}_h v_0$ we deduce that

$$\|U^n\|^2 + \left\|T_h^{\frac{1}{2}} V^n\right\|^2 + 2\alpha \sum_{m=1}^{n} k\|V^m\|^2 \le \|\mathcal{P}_h u_0\|^2 + \left\|T_h^{\frac{1}{2}} \mathcal{P}_h v_0\right\|^2 \le C\big(\|u_0\|^2 + |u_0|_{-1}^2\big). \tag{3.68}$$

Now we only need to bound $\|u(t_n)\|^2 + \sum_{m=1}^n k\|v(t_m)\|^2$ before we can show (3.63) for $\beta = 0$. Observing

$$v(t_m) = \frac{1}{k}\int_{t_{m-1}}^{t_m} (s - t_{m-1}) v_t(s)\,\mathrm{d}s + \frac{1}{k}\int_{t_{m-1}}^{t_m} v(s)\,\mathrm{d}s, \tag{3.69}$$

due to integration by parts we additionally use (2.12) and (3.54) to derive

$$\begin{aligned}\sum_{m=1}^{n} k\|v(t_m)\|^2 &\le 2\sum_{m=1}^{n} k\left(\left\|\frac{1}{k}\int_{t_{m-1}}^{t_m} (s - t_{m-1}) v_t(s)\,\mathrm{d}s\right\|^2 + \left\|\frac{1}{k}\int_{t_{m-1}}^{t_m} v(s)\,\mathrm{d}s\right\|^2\right)\\ &\le 2\int_0^{t_n} s^2\|v_t(s)\|^2\,\mathrm{d}s + 2\int_0^{t_n} \|v(s)\|^2\,\mathrm{d}s \le C\big(\|u_0\|^2 + |v_0|_{-1}^2\big),\end{aligned} \tag{3.70}$$

where at the second step we used $0 \le s - t_{m-1} \le s$. This together with (2.11) and (3.68) verifies (3.63) for $\beta = 0$. Next we validate (3.63) for $\beta = 2$. Similarly to (3.23) we introduce some notations as

$$\begin{aligned}&\theta_1^n = U^n - \mathcal{R}_h u(t_n),\ \rho_1^n = (\mathcal{R}_h - I)u(t_n),\\ &\theta_2^n = V^n - \mathcal{P}_h v(t_n),\ \rho_2^n = (\mathcal{P}_h - I)v(t_n).\end{aligned} \tag{3.71}$$

Combining (3.9) and (3.65) yields

$$\begin{aligned}\big(\bar{\partial}_n V^n - v_t(t_n), \chi_1\big) + \alpha\big(\nabla\big(V^n - v(t_n)\big), \nabla\chi_1\big) + \big(\nabla(U^n - u(t_n)), \nabla\chi_1\big) = 0,\ \forall\chi_1 \in V_h,\\ \big(\nabla(\bar{\partial}_n U^n - u_t(t_n)), \nabla\chi_2\big) - \big(\nabla(V^n - v(t_n)), \nabla\chi_2\big) = 0,\ \forall\chi_2 \in V_h.\end{aligned} \tag{3.72}$$

Taking the definitions of $\mathcal{P}_h$ and $\mathcal{R}_h$ into account and plugging the notations proposed in (3.71) show

$$\big(\bar{\partial}_n\theta_2^n, \chi_1\big) + \alpha\big(\nabla\theta_2^n, \nabla\chi_1\big) + \big(\nabla\theta_1^n, \nabla\chi_1\big) = \big(\rho_4^n, \chi_1\big) - \alpha\big(\nabla\rho_2^n, \nabla\chi_1\big),\ \forall\chi_1 \in V_h, \tag{3.73}$$

$$\big(\nabla\bar{\partial}_n\theta_1^n, \nabla\chi_2\big) - \big(\nabla\theta_2^n, \nabla\chi_2\big) = \big(\nabla\rho_3^n, \nabla\chi_2\big) + \big(\nabla\rho_2^n, \nabla\chi_2\big),\ \forall\chi_2 \in V_h, \tag{3.74}$$

where further notations were also introduced:

$$\rho_3^n := u_t(t_n) - \overline{\partial}_n u(t_n) \quad \text{and} \quad \rho_4^n := v_t(t_n) - \overline{\partial}_n v(t_n). \tag{3.75}$$

As in the proof of (3.30) setting $\chi_1 = T_h\theta_2^n$, $\chi_2 = T_h\theta_1^n$ in (3.73) and (3.74) and adding together give

$$\begin{aligned}\big(\overline{\partial}_n\theta_2^n, T_h\theta_2^n\big) + \big(\overline{\partial}_n\theta_1^n, \theta_1^n\big) + \alpha\big\|\theta_2^n\big\|^2 &= \big(T_h\mathcal{P}_h\rho_4^n, \theta_2^n\big) - \alpha\big(\mathcal{R}_h\rho_2^n, \theta_2^n\big) + \big(\mathcal{R}_h\rho_3^n, \theta_1^n\big) + \big(\mathcal{R}_h\rho_2^n, \theta_1^n\big)\\ &= \Big(\tfrac{1}{\sqrt{\alpha}}T_h\mathcal{P}_h\rho_4^n - \sqrt{\alpha}\mathcal{R}_h\rho_2^n, \sqrt{\alpha}\theta_2^n\Big) + \big(\mathcal{R}_h\rho_3^n + \mathcal{R}_h\rho_2^n, \theta_1^n\big).\end{aligned} \tag{3.76}$$

Using the facts $\big(\overline{\partial}_n\theta_1^n, \theta_1^n\big) \geq \frac{1}{2k}(\|\theta_1^n\|^2 - \|\theta_1^{n-1}\|^2)$ and $\big(\overline{\partial}_n\theta_2^n, T_h\theta_2^n\big) \geq \frac{1}{2k}\Big(\|T_h^{\frac{1}{2}}\theta_2^n\|^2 - \|T_h^{\frac{1}{2}}\theta_2^{n-1}\|^2\Big)$ shows

$$\begin{aligned}&\tfrac{1}{2k}\Big(\big\|\theta_1^n\big\|^2 - \big\|\theta_1^{n-1}\big\|^2 + \Big\|T_h^{\frac{1}{2}}\theta_2^n\Big\|^2 - \Big\|T_h^{\frac{1}{2}}\theta_2^{n-1}\Big\|^2\Big) + \alpha\big\|\theta_2^n\big\|^2\\ &\quad\leq \tfrac{1}{\alpha}\big\|T_h\mathcal{P}_h\rho_4^n\big\|^2 + (\alpha+1)\big\|\mathcal{R}_h\rho_2^n\big\|^2 + \tfrac{\alpha}{2}\big\|\theta_2^n\big\|^2 + \big\|\mathcal{R}_h\rho_3^n\big\|^2 + \tfrac{1}{2}\big\|\theta_1^n\big\|^2.\end{aligned} \tag{3.77}$$

Hence, by summation and detecting that $\theta_2^0 = V^0 - \mathcal{P}_h v_0 = 0$ we infer

$$\Big\|T_h^{\frac{1}{2}}\theta_2^n\Big\|^2 + \alpha k\sum_{m=1}^n\big\|\theta_2^m\big\|^2 + \big\|\theta_1^n\big\|^2 \leq Ck\sum_{m=1}^n\Big(\big\|\mathcal{R}_h\rho_3^m\big\|^2 + \big\|T_h\mathcal{P}_h\rho_4^m\big\|^2 + \big\|\mathcal{R}_h\rho_2^m\big\|^2 + \big\|\theta_1^m\big\|^2\Big) + \big\|\theta_1^0\big\|^2.$$

Applying the discrete Gronwall inequality helps us to get

$$\Big\|T_h^{\frac{1}{2}}\theta_2^n\Big\|^2 + \alpha k\sum_{m=1}^n\big\|\theta_2^m\big\|^2 + \big\|\theta_1^n\big\|^2 \leq Ck\sum_{m=1}^n\Big(\big\|\mathcal{R}_h\rho_3^m\big\|^2 + \big\|T_h\mathcal{P}_h\rho_4^m\big\|^2 + \big\|\mathcal{R}_h\rho_2^m\big\|^2\Big) + Ch^4|u_0|_2^2, \tag{3.78}$$

where $\|\theta_1^0\| = \|U_0 - \mathcal{R}_h u_0\| = \|\mathcal{P}_h(I - \mathcal{R}_h)u_0\| \leq Ch^2|u_0|_2$. In the sequel we will estimate the remaining three terms separately. Note first that ρ_3^m admits the following expression:

$$\rho_3^m = u_t(t_m) - \overline{\partial}_m u(t_m) = \frac{1}{k}\int_{t_{m-1}}^{t_m}(s - t_{m-1})u_{tt}(s)\,\mathrm{d}s. \tag{3.79}$$

This together with (3.52) and (3.54) guarantees

$$
\begin{aligned}
k\sum_{m=1}^{n}\left\|\mathcal{R}_h\rho_3^m\right\|^2 &\le 2k\sum_{m=1}^{n}\left\|(I-\mathcal{R}_h)\rho_3^m\right\|^2+2k\sum_{m=1}^{n}\left\|\rho_3^m\right\|^2\\
&\le Ck\sum_{m=1}^{n}\left(h^4\left|\frac{1}{k}\int_{t_{m-1}}^{t_m}(s-t_{m-1})u_{tt}(s)\,\mathrm{d}s\right|_2^2+\left\|\frac{1}{k}\int_{t_{m-1}}^{t_m}(s-t_{m-1})u_{tt}(s)\,\mathrm{d}s\right\|^2\right)\\
&\le C\big(h^4+k^2\big)\int_0^{t_n}\left(s^2|u_{tt}(s)|_2^2+\|u_{tt}(s)\|^2\right)\mathrm{d}s\le C\big(h^4+k^2\big)\big(|u_0|_2^2+|v_0|_1^2\big),
\end{aligned}
\tag{3.80}
$$

where the fact $s-t_{m-1}\le s$ was used. Likewise, noting that $T_h\mathcal{P}_h=T_h$ and that ρ_4^m has the same expression as (3.79) with u replaced by v yields

$$
\begin{aligned}
k\sum_{m=1}^{n}\left\|T_h\mathcal{P}_h\rho_4^m\right\|^2 &\le 2k\sum_{m=1}^{n}\left\|(T-T_h)\rho_4^m\right\|^2+2k\sum_{m=1}^{n}\left\|T\rho_4^m\right\|^2\\
&\le Ck\sum_{m=1}^{n}\left(h^4\left\|\frac{1}{k}\int_{t_{n-1}}^{t_n}(s-t_{n-1})v_{tt}(s)\,\mathrm{d}s\right\|^2+\left\|\frac{1}{k}\int_{t_{n-1}}^{t_n}(s-t_{n-1})Tv_{tt}(s)\,\mathrm{d}s\right\|^2\right)\\
&\le Ch^4\int_0^{t_n}s^2\|u_{ttt}(s)\|^2\,\mathrm{d}s+Ck^2\int_0^{t_n}\left(\|u_{tt}(s)\|^2+\|u_t(s)\|^2\right)\mathrm{d}s\\
&\le C\big(h^4+k^2\big)\big(|u_0|_2^2+|v_0|_1^2\big),
\end{aligned}
\tag{3.81}
$$

where we also used (2.12), (3.8), (3.52), (3.53) and $Tu_{ttt}=-\alpha u_{tt}-u_t$. Using similar arguments as before and taking (3.4), (2.12), (3.54) and (3.69) into account, one can show that

$$
\begin{aligned}
k\sum_{m=1}^{n}\left\|\mathcal{R}_h\rho_2^m\right\|^2 &= k\sum_{m=1}^{n}\|\mathcal{P}_h(\mathcal{R}_h-I)v(t_m)\|^2\le Ch^4k\sum_{m=1}^{n}|v(t_m)|_2^2\\
&\le Ch^4k\sum_{m=1}^{n}\left|\frac{1}{k}\int_{t_{m-1}}^{t_m}(s-t_{m-1})v_t(s)\,\mathrm{d}s+\frac{1}{k}\int_{t_{m-1}}^{t_m}v(s)\,\mathrm{d}s\right|_2^2\\
&\le Ch^4\left(\int_0^{t_n}s^2|u_{tt}(s)|_2^2\,\mathrm{d}s+\int_0^{t_n}|u_t(s)|_2^2\,\mathrm{d}s\right)\le Ch^4\big(|u_0|_2^2+|v_0|_1^2\big).
\end{aligned}
\tag{3.82}
$$

Analogously, one can achieve

$$
k\sum_{m=1}^{n}\left\|\rho_2^m\right\|^2+\left\|\rho_1^n\right\|^2\le C\big(h^4+k^2\big)\big(|u_0|_2^2+|v_0|_1^2\big).
\tag{3.83}
$$

Finally, plugging (3.80)–(3.82) into (3.78) and considering (3.83) help us to get

$$k\sum_{m=1}^{n}\|V^m - v(t_m)\|^2 + \|U^n - u(t_n)\|^2 \le C\big(h^4 + k^2\big)\big(|u_0|_2^2 + |v_0|_1^2\big). \tag{3.84}$$

The intermediate cases follow by interpolation. □

4. FEM for the stochastic problem

This section is devoted to the finite element approximation of the stochastic problem (2.1). The convergence analysis relies on regularity properties of the mild solution of (2.1) as well as error estimates obtained in Section 3.

4.1 *Spatial semidiscretization*

In this subsection we shall follow notations introduced in Section 3 and analyze the semidiscrete finite element approximation of (2.1). Let V_h be the finite element space defined in the previous section. The semidiscrete approximation of (2.1) is to find $X_h(t) = (u_h(t), u_{h,t}(t))' \in V_h \times V_h$ such that

$$\mathrm{d}X_h(t) + \mathcal{A}_h X_h(t)\,\mathrm{d}t = \mathbf{P}_h\mathbf{F}(X_h(t))\,\mathrm{d}t + \mathbf{P}_h\mathbf{B}\,\mathrm{d}W(t), \quad \text{in } \ t \in (0, T], \ X_h(0) = \mathbf{P}_h X_0, \tag{4.1}$$

or in the mild form

$$X_h(t) = \mathcal{S}(t)\mathbf{P}_h X_0 + \int_0^t \mathcal{S}_h(t-s)\mathbf{P}_h\mathbf{F}(X_h(s))\,\mathrm{d}s + \int_0^t \mathcal{S}_h(t-s)\mathbf{P}_h\mathbf{B}\,\mathrm{d}W(s), \ t \in [0, T]. \tag{4.2}$$

The first main convergence result is as follows.

THEOREM 4.1 Let Assumptions 2.1–2.3 hold with $\gamma \in [0, 1]$ and let the setting in the beginning of Section 3 be fulfilled. Let $(u(t), u_t(t))'$ and $(u_h(t), u_{h,t}(t))'$ be the mild solutions of the problems (2.1) and (4.1), respectively. Then for all $t \in [0, T]$ it holds that

$$\|u(t) - u_h(t)\|_{L^2(\Omega;\dot{H}^0)} \le Ch^{1+\gamma}\left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^{\gamma+1})} + \|\psi\|_{L^2(\Omega;\dot{H}^{\gamma-1})}\right). \tag{4.3}$$

If additionally $\psi \in L^2(\Omega; \dot{H}^\gamma)$ then for all $t \in [0, T]$ it holds that

$$\|u_t(t) - u_{h,t}(t)\|_{L^2(\Omega;\dot{H}^0)} \le Ch^{\gamma}\left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^{\gamma})} + \|\psi\|_{L^2(\Omega;\dot{H}^{\gamma})}\right). \tag{4.4}$$

REMARK 4.2 When Assumption 2.3 is fulfilled with $\gamma \in [-1, 0)$ the problem (1.1) can still admit a mild solution $\{u(t)\}_{t\in[0,T]}$ that exhibits a positive order of regularity. However, in this situation the Wiener process takes values in $\dot{H}^\delta$ for $\delta = \gamma - 1 < -1$, which destroys the well-posedness of $\mathbf{P}_h$, as also explained in Remark 3.2. Accordingly, we restrict ourselves to $\gamma \in [0, 1]$ throughout this section. Also, we remark that the error constants C in (4.3) and (4.4) essentially depend on $\frac{1}{\alpha}$ and thus $C \to \infty$ when $\alpha \to 0$. To simply see this fact, let us, for example, just recall the spatial error estimate $\left(\int_0^t \|P_2 F_h(s) w_0\|^2\,\mathrm{d}s\right)^{\frac{1}{2}} = \left(\int_0^t \|v_h - v\|^2\,\mathrm{d}s\right)^{\frac{1}{2}} \le Ch^\beta(|u_0|_\beta + |v_0|_{\beta-1})$ in (3.16) of Theorem 3.3. To prove it we resort to interpolation argument, where, e.g., the estimate (3.32) clearly indicates that the above constant C must depend on $\frac{1}{\alpha}$. Likewise, the dependence of $\frac{1}{\alpha}$ can be easily detected in estimates (3.77) and (3.78), which are used for error estimates (4.17) and (4.18) of space-time full discretization. Actually, apart from numerical error estimates, many estimates for the continuous problem get involved with $\frac{1}{\alpha}$, say, (2.12), (2.13) and (3.62), which is used to arrive at $\int_0^t s^2 |u_{tt}(s)|_\beta^2 \, \mathrm{d}s \leq C(|u_0|_\beta^2 + |v_0|_{\beta-1}^2)$ in (3.54).

Proof of Theorem 4.1. Subtracting (2.5) from (4.2) gives

$$\begin{aligned} X_h(t) - X(t) &= \big(\mathcal{S}_h(t)\mathbf{P}_h - \mathcal{S}(t)\big)X_0 + \int_0^t \big(\mathcal{S}_h(t-s)\mathbf{P}_h - \mathcal{S}(t-s)\big)\mathbf{F}(X(s))\,\mathrm{d}s \\ &\quad + \int_0^t \mathcal{S}_h(t-s)\mathbf{P}_h\Big(\mathbf{F}\big(X_h(s)\big) - \mathbf{F}\big(X(s)\big)\Big)\,\mathrm{d}s + \int_0^t \big(\mathcal{S}_h(t-s)\mathbf{P}_h - \mathcal{S}(t-s)\big)\mathbf{B}\,\mathrm{d}W(s) \\ &:= J_1 + J_2 + J_3 + J_4. \end{aligned} \tag{4.5}$$

Recalling $u(t) - u_h(t) = P_1(X(t) - X_h(t))$ we require to bound $P_1 J_i, i = 1,2,3,4$. For the term $P_1 J_1$ a combination with (3.14) and (3.16) enables us to claim that, for $\beta \in [0,\gamma]$, $i \in \{0,1\}$,

$$\|P_1 F_h(t) w_0\| \leq C h^{\beta+i}(|u_0|_{\beta+i} + |v_0|_{\beta-1}), \tag{4.6}$$

which together with Assumption 2.3 leads to

$$\|P_1 J_1\|_{L^2(\Omega;\dot{H}^0)} = \|P_1 F_h(t) X_0\|_{L^2(\Omega;\dot{H}^0)} \leq C h^{\beta+i}\left(\|\varphi\|_{L^2(\Omega;\dot{H}^{\beta+i})} + \|\psi\|_{L^2(\Omega;\dot{H}^{\beta-1})}\right). \tag{4.7}$$

Similarly, using (3.14) and (2.38) shows

$$\begin{aligned} \|P_1 J_2\|_{L^2(\Omega;\dot{H}^0)} &\leq \int_0^t \|P_1 F_h(t-s)\mathbf{F}(X(s))\|_{L^2(\Omega;\dot{H}^0)}\,\mathrm{d}s \leq C h^2 \int_0^t \|F(u(s))\|_{L^2(\Omega;\dot{H}^0)}\,\mathrm{d}s \\ &\leq C h^2\left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^0)} + \|\psi\|_{L^2(\Omega;\dot{H}^{-2})}\right). \end{aligned} \tag{4.8}$$

To bound $P_1 J_3$ we combine the stability of $\mathcal{S}_h(t)\mathbf{P}_h$ in $\dot{H}^0 \times \dot{H}^0$ with Assumption 2.2 to derive

$$\begin{aligned} \|P_1 J_3\|_{L^2(\Omega;\dot{H}^0)} &\leq \int_0^t \Big\| P_1 \mathcal{S}_h(t-s)\mathbf{P}_h\Big(\mathbf{F}\big(X_h(s)\big) - \mathbf{F}\big(X(s)\big)\Big)\Big\|_{L^2(\Omega;\dot{H}^0)}\,\mathrm{d}s \\ &\leq C\int_0^t \|F(u_h(s)) - F(u(s))\|_{L^2(\Omega;\dot{H}^0)}\,\mathrm{d}s \leq C\int_0^t \|u_h(s) - u(s)\|_{L^2(\Omega;\dot{H}^0)}\,\mathrm{d}s. \end{aligned} \tag{4.9}$$

Again, using (3.14) and the Itô isometry yields

$$\|P_1 J_4\|_{L^2(\Omega;\dot{H}^0)} = \left\| \int_0^t P_1 F_h(t-s) \mathbf{B}\, \mathrm{d}W(s) \right\|_{L^2(\Omega;\dot{H}^0)} = \left(\int_0^t \left\| P_1 F_h(t-s) \mathbf{B} Q^{\frac{1}{2}} \right\|_{\mathrm{HS}}^2 \mathrm{d}s \right)^{\frac{1}{2}}$$
$$\leq C\sqrt{T} h^{\gamma+1} \left\| A^{\frac{\gamma-1}{2}} Q^{\frac{1}{2}} \right\|_{\mathrm{HS}}. \tag{4.10}$$

Finally, putting the above estimates together and employing Gronwall's inequality give

$$\|u(t) - u_h(t)\|_{L^2(\Omega;\dot{H}^0)} \leq C h^{\beta+i} \left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^{\beta+i})} + \|\psi\|_{L^2(\Omega;\dot{H}^{\beta-1})} \right), \quad \beta \in [0,\gamma],\ i \in \{0,1\}. \tag{4.11}$$

Letting $\beta = \gamma$, $i = 1$ in (4.11) hence yields (4.3). Next we are to verify (4.4). Following the same notations as before we need to estimate $P_2 J_i, i = 1,2,3,4$. Using (3.15) with $q = \gamma$ gives

$$\|P_2 J_1\|_{L^2(\Omega;\dot{H}^0)} = \|P_2 F_h(t) X_0\|_{L^2(\Omega;\dot{H}^0)} \leq C h^{\gamma} \left(\|\varphi\|_{L^2(\Omega;\dot{H}^{\gamma})} + \|\psi\|_{L^2(\Omega;\dot{H}^{\gamma})} \right). \tag{4.12}$$

To deal with the term $P_2 J_2$ we employ (3.17) with $q = \gamma, s = 0$ and (2.38) to arrive at

$$\|P_2 J_2\|_{L^2(\Omega;\dot{H}^0)} \leq \int_0^t \left\| P_2 F_h(t-s) \mathbf{F}(X(s)) \right\|_{L^2(\Omega;\dot{H}^0)} \mathrm{d}s \leq C h^{\gamma} \int_0^t (t-s)^{-\frac{\gamma}{2}} \|F(u(s))\|_{L^2(\Omega;\dot{H}^0)} \mathrm{d}s$$
$$\leq C h^{\gamma} \left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^0)} + \|\psi\|_{L^2(\Omega;\dot{H}^{-2})} \right). \tag{4.13}$$

The stability of $\mathcal{S}_h(t)\mathbf{P}_h$ in $\dot{H}^0 \times \dot{H}^0$, (4.11) with $\beta = \gamma$, $i = 0$ and Assumption 2.1 ensure

$$\|P_2 J_3\|_{L^2(\Omega;\dot{H}^0)} \leq \int_0^t \left\| P_2 \mathcal{S}_h(t-s) \mathbf{P}_h \Big(\mathbf{F}(X(s)) - \mathbf{F}(X_h(s)) \Big) \right\|_{L^2(\Omega;\dot{H}^0)} \mathrm{d}s$$
$$\leq C \int_0^t \|u(s) - u_h(s)\|_{L^2(\Omega;\dot{H}^0)} \mathrm{d}s \leq C h^{\gamma} \left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^{\gamma})} + \|\psi\|_{L^2(\Omega;\dot{H}^{\gamma-1})} \right). \tag{4.14}$$

At last, Itô's isometry and (3.16) with $\beta = \gamma$ help us to estimate $P_2 J_4$ as follows:

$$\|P_2 J_4\|_{L^2(\Omega;\dot{H}^0)} = \left(\int_0^t \left\| P_2 F_h(t-s) \mathbf{B} Q^{\frac{1}{2}} \right\|_{\mathrm{HS}}^2 \mathrm{d}s \right)^{\frac{1}{2}} \leq C h^{\gamma} \left\| A^{\frac{\gamma-1}{2}} Q^{\frac{1}{2}} \right\|_{\mathrm{HS}}. \tag{4.15}$$

Now gathering the estimates of $P_2 J_i, i = 1,2,3,4$ together gives the estimate of $u_t(t) - u_{h,t}(t)$. □

4.2 *Full discretization*

Below, we proceed to treat the full-discrete scheme for (2.1). Let k be the time step size and write $t_n = nk$, for $n \geq 1$. We discretize (4.1) in time with a linear implicit Euler scheme and the resulting full discretization is thus to find $\mathcal{F}_{t_n}$-adapted random variables $X^n = (U^n, V^n)' \in V_h \times V_h$ such that

$$X^n - X^{n-1} + k\mathcal{A}_h X^n = k\mathbf{P}_h\mathbf{F}\big(X^{n-1}\big) + \mathbf{P}_h\mathbf{B}\Delta W_n,\ X^0 = \mathbf{P}_h X_0, \tag{4.16}$$

where $\Delta W_n := W(t_n) - W(t_{n-1})$ is the Wiener increment. Now we state our last convergence result.

Theorem 4.3 Let $(u(t), u_t(t))'$ and $(U^n, V^n)'$ be the solutions of (2.1) and (4.16), respectively. If Assumptions 2.1–2.3 hold with $\gamma \in [0,1]$ and the setting in the beginning of Section 3 holds, then

$$\left\|u(t_n) - U^n\right\|_{L^2(\Omega;\dot{H}^0)} \le C\left(h^{\gamma+1} + k^{\frac{\gamma+1}{2}}\right)\left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^{\gamma+1})} + \|\psi\|_{L^2(\Omega;\dot{H}^{\gamma-1})}\right). \tag{4.17}$$

If additionally $\psi \in L^2(\Omega;\dot{H}^\gamma)$ then

$$\|u_t(t_n) - V^n\|_{L^2(\Omega;\dot{H}^0)} \le C\left(h^\gamma + k^{\frac{\gamma}{2}}\right)\left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^\gamma)} + \|\psi\|_{L^2(\Omega;\dot{H}^\gamma)}\right). \tag{4.18}$$

We begin by introducing a crucial ingredient in the following convergence analysis.

Lemma 4.4 Suppose that $w_0 = (u_0, v_0)' \in \dot{H}^\mu \times \dot{H}^{\mu-1}$ for some $\mu \in [0,2]$. Then

$$\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \|P_2(\mathcal{S}(t_n - s) - \mathcal{S}(t_n - t_j))w_0\|^2\,\mathrm{d}s \le Ck^\mu\big(|u_0|_\mu^2 + |v_0|_{\mu-1}^2\big). \tag{4.19}$$

Proof of Lemma 4.4. Keep in mind that

$$P_2(\mathcal{S}(t_n - s) - \mathcal{S}(t_n - t_j))w_0 = u_t(t_n - s) - u_t(t_n - t_j), \tag{4.20}$$

where by abuse of notation we view $(u(t), u_t(t))'$ as the solution of equation (2.10). By interpolation we only need to verify (4.19) for the cases $\mu = 0$ and $\mu = 2$. Using (3.52) shows

$$\begin{aligned}\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \|u_t(t_n - s) - u_t(t_n - t_j)\|^2\,\mathrm{d}s &= \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \left\|\int_{t_j}^{s} u_{tt}(t_n - r)\,\mathrm{d}r\right\|^2\mathrm{d}s\\ &\le Ck^2\int_0^{t_n}\|u_{tt}(t_n - s)\|^2\,\mathrm{d}s \le Ck^2\big(|u_0|_2^2 + |v_0|_1^2\big).\end{aligned} \tag{4.21}$$

Also, employing (2.12) with $\beta = 0$ and (3.70) shows

$$\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \|u_t(t_n - s) - u_t(t_n - t_j)\|^2\,\mathrm{d}s \le 2k\sum_{j=1}^{n}\|u_t(t_j)\|^2 + 2\int_0^{t_n}\|u_t(t_n - s)\|^2\,\mathrm{d}s \le C\big(\|u_0\|^2 + |v_0|_{-1}^2\big).$$

This and (4.21) together concludes the proof of this lemma. □

Proof of Theorem 4.3. Equivalently, (4.16) can be reformulated as

$$X^n = r(k\mathcal{A}_h)^n \mathbf{P}_h X_0 + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} r(k\mathcal{A}_h)^{n-j} \mathbf{P}_h \mathbf{F}\big(X^j\big)\,\mathrm{d}s + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} r(k\mathcal{A}_h)^{n-j} \mathbf{P}_h \mathbf{B}\,\mathrm{d}W(s). \tag{4.22}$$

Therefore, the difference between X^n and $X(t_n)$ can be decomposed as follows:

$$\begin{aligned} X^n - X(t_n) &= (r(k\mathcal{A}_h)^n \mathbf{P}_h - \mathcal{S}(t_n))X_0 + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} r(k\mathcal{A}_h)^{n-j} \mathbf{P}_h \Big(\mathbf{F}(X^j) - \mathbf{F}(X(s))\Big)\,\mathrm{d}s \\ &\quad + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \Big(r(k\mathcal{A}_h)^{n-j} \mathbf{P}_h - \mathcal{S}(t_{n-j})\Big)\,\mathbf{F}(X(s))\,\mathrm{d}s \\ &\quad + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \big(\mathcal{S}(t_{n-j}) - \mathcal{S}(t_n - s)\big)\,\mathbf{F}(X(s))\,\mathrm{d}s \\ &\quad + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} \Big(r(k\mathcal{A}_h)^{n-j} \mathbf{P}_h - \mathcal{S}(t_n - s)\Big)\,\mathbf{B}\,\mathrm{d}W(s) := \mathbb{J}_1 + \mathbb{J}_2 + \mathbb{J}_3 + \mathbb{J}_4 + \mathbb{J}_5. \end{aligned} \tag{4.23}$$

Since $U^n - u(t_n) = P_1(X^n - X(t_n))$ the estimate $\|u(t_n) - U^n\|_{L^2(\Omega;\dot{H}^0)}$ can be achieved via estimates of $\|P_1\mathbb{J}_j\|_{L^2(\Omega;\dot{H}^0)}$, $j = 1, 2, \ldots, 5$. As in (4.6) combining (3.40) and (3.63) implies, for all $\beta \in [0, \gamma]$, $i \in \{0, 1\}$,

$$\big\|P_1 F_{kh}^n w_0\big\| \leq C\left(h^{\beta+i} + k^{\frac{\beta+i}{2}}\right)(|u_0|_{\beta+i} + |v_0|_{\beta-1}). \tag{4.24}$$

This immediately leads us to the estimate of $P_1\mathbb{J}_1$ as follows:

$$\|P_1\mathbb{J}_1\|_{L^2(\Omega;\dot{H}^0)} \leq C\left(h^{\beta+i} + k^{\frac{\beta+i}{2}}\right)\Big(\|\varphi\|_{L^2(\Omega;\dot{H}^{\beta+i})} + \|\psi\|_{L^2(\Omega;\dot{H}^{\beta-1})}\Big), \quad \beta \in [0, \gamma],\ i \in \{0, 1\}. \tag{4.25}$$

For $P_1\mathbb{J}_2$ we use the stability property (3.38), the regularity (2.44) with $\varrho = \beta + i - 1$, $i \in \{0,1\}$ and Assumption 2.1 to get

$$\begin{aligned}
\|P_1\mathbb{J}_2\|_{L^2(\Omega;\dot{H}^0)} &\leq \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \left\|P_1 r(k\mathcal{A}_h)^{n-j}\mathbf{P}_h\big(\mathbf{F}\big(X^j\big) - \mathbf{F}(X(s))\big)\right\|_{L^2(\Omega;\dot{H}^0)} \mathrm{d}s \\
&\leq C\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \left\|F\big(U^j\big) - F(u(s))\right\|_{L^2(\Omega;\dot{H}^0)} \mathrm{d}s \\
&\leq C\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \Big(\left\|U^j - u(t_j)\right\|_{L^2(\Omega;\dot{H}^0)} + \|u(t_j) - u(s)\|_{L^2(\Omega;\dot{H}^0)}\Big) \mathrm{d}s \\
&\leq Ck\sum_{j=0}^{n-1} \left\|U^j - u(t_j)\right\|_{L^2(\Omega;\dot{H}^0)} + Ck^{\frac{\beta+i}{2}}\Big(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^{\beta+i})} + \|\psi\|_{L^2(\Omega;\dot{H}^{\beta+i-2})}\Big).
\end{aligned} \tag{4.26}$$

To bound the term $P_1\mathbb{J}_3$ we recall (3.40) with $\beta = 2$ and derive that

$$\begin{aligned}
\|P_1\mathbb{J}_3\|_{L^2(\Omega;\dot{H}^0)} &\leq \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \left\|P_1 F_{kh}^{n-j}\mathbf{F}(X(s))\right\|_{L^2(\Omega;\dot{H}^0)} \mathrm{d}s \leq C\big(h^2 + k\big)T \sup_{s\in[0,T]} \|F(u(s))\|_{L^2(\Omega;\dot{H}^0)} \\
&\leq C\big(h^2 + k\big)\Big(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^0)} + \|\psi\|_{L^2(\Omega;\dot{H}^{-2})}\Big).
\end{aligned} \tag{4.27}$$

In the same spirit as before, but employing (2.30) with $\mu = 2$ instead, we obtain

$$\begin{aligned}
\|P_1\mathbb{J}_4\|_{L^2(\Omega;\dot{H}^0)} &\leq \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \|P_1(\mathcal{S}(t_{n-j}) - \mathcal{S}(t_n - s))\mathbf{F}(X(s))\|_{L^2(\Omega;\dot{H}^0)} \mathrm{d}s \\
&\leq Ck\Big(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^0)} + \|\psi\|_{L^2(\Omega;\dot{H}^{-2})}\Big).
\end{aligned} \tag{4.28}$$

Now it remains to estimate $P_1\mathbb{J}_5$. Employing (2.30), (3.40) and Itô's isometry together promises

$$\begin{aligned}
\|P_1\mathbb{J}_5\|^2_{L^2(\Omega;\dot{H}^0)} &= \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \left\|P_1\Big(r(k\mathcal{A}_h)^{n-j}\mathbf{P}_h - \mathcal{S}(t_n - s)\Big)\mathbf{B}Q^{\frac{1}{2}}\right\|^2_{\mathrm{HS}} \mathrm{d}s \\
&\leq 2\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \left(\left\|P_1 F_{kh}^{n-j}\mathbf{B}Q^{\frac{1}{2}}\right\|^2_{\mathrm{HS}} + \left\|P_1(\mathcal{S}(t_{n-j}) - \mathcal{S}(t_n - s))\mathbf{B}Q^{\frac{1}{2}}\right\|^2_{\mathrm{HS}}\right) \mathrm{d}s \\
&\leq C\big(h^{2(\gamma+1)} + k^{\gamma+1}\big)\left\|A^{\frac{\gamma-1}{2}}Q^{\frac{1}{2}}\right\|^2_{\mathrm{HS}}.
\end{aligned} \tag{4.29}$$

Putting the above five estimates together and applying Gronwall's inequality imply that,

$$\|u(t_n) - U^n\| \leq C\left(h^{\beta+i} + k^{\frac{\beta+i}{2}}\right)\left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^{\beta+i})} + \|\psi\|_{L^2(\Omega;\dot{H}^{\beta-1})}\right) \quad \text{for } \beta \in [0,\gamma],\ i \in \{0,1\}, \tag{4.30}$$

which validates (4.17) by taking $\beta = \gamma, i = 1$. In the sequel we turn our attention to the estimate of $V^n - v(t_n)$. Using (3.41) with $q = \gamma$ suggests

$$\|P_2\mathbb{J}_1\|_{L^2(\Omega;\dot{H}^0)} \leq C\big\|P_2F^n_{kh}X_0\big\|_{L^2(\Omega;\dot{H}^0)} \leq C\left(h^{\gamma} + k^{\frac{\gamma}{2}}\right)\left(\|\varphi\|_{L^2(\Omega;\dot{H}^{\gamma})} + \|\psi\|_{L^2(\Omega;\dot{H}^{\gamma})}\right). \tag{4.31}$$

Before treating $P_2\mathbb{J}_2$ we again recall the stability property (3.38). Following the same arguments as used in (4.26) and using Assumption 2.1, (2.44) with $\varrho = \gamma - 1$ and (4.30) with $\beta = \gamma, i = 0$ give

$$\begin{aligned}
\|P_2\mathbb{J}_2\|_{L^2(\Omega;\dot{H}^0)} &\leq \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \big\|P_2 r(k\mathcal{A}_h)^{n-j}\mathbf{P}_h\big(\mathbf{F}\big(X^j\big) - \mathbf{F}(X(s))\big)\big\|_{L^2(\Omega;\dot{H}^0)}\,\mathrm{d}s \\
&\leq C\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \big\|F\big(U^j\big) - F(u(s))\big\|_{L^2(\Omega;\dot{H}^0)}\,\mathrm{d}s \\
&\leq C\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \left(\big\|U^j - u(t_j)\big\|_{L^2(\Omega;\dot{H}^0)} + \|u(t_j) - u(s)\|_{L^2(\Omega;\dot{H}^0)}\right)\mathrm{d}s \\
&\leq C\left(h^{\gamma} + k^{\frac{\gamma}{2}}\right)\left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^{\gamma})} + \|\psi\|_{L^2(\Omega;\dot{H}^{\gamma-1})}\right).
\end{aligned} \tag{4.32}$$

Similarly to (4.27) we utilize (3.64) with $s = 0$, $q = 1$, (2.38) and Assumption 2.1 to achieve

$$\begin{aligned}
\|P_2\mathbb{J}_3\|_{L^2(\Omega;\dot{H}^0)} &\leq \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \big\|P_2F^{n-j}_{kh}\mathbf{F}(X(s))\big\|_{L^2(\Omega;\dot{H}^0)}\,\mathrm{d}s \\
&\leq C\left(h + k^{\frac{1}{2}}\right)\sum_{j=1}^{n-1} k t_{n-j}^{-\frac{1}{2}} \sup_{s\in[0,T]} \|F(u(s))\|_{L^2(\Omega;\dot{H}^0)} \\
&\leq C\left(h + k^{\frac{1}{2}}\right)\left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^0)} + \|\psi\|_{L^2(\Omega;\dot{H}^{-2})}\right).
\end{aligned} \tag{4.33}$$

To handle the term $P_2\mathbb{J}_4$ by (2.31) with $\nu = \frac{1}{2}$ and (2.38) one can deduce

$$\begin{aligned}\|P_2\mathbb{J}_4\|_{L^2(\Omega;\dot{H}^0)} &\le \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \|P_2(\mathcal{S}(t_{n-j}) - \mathcal{S}(t_n - s))\mathbf{F}(X(s))\|_{L^2(\Omega;\dot{H}^0)}\,\mathrm{d}s \\ &\le C\int_0^{t_n} k^{\frac{1}{2}}(t_n - s)^{-\frac{1}{2}}\,\mathrm{d}s \sup_{s\in[0,T]} \|F(u(s))\|_{L^2(\Omega;\dot{H}^0)} \\ &\le Ck^{\frac{1}{2}}\left(1 + \|\varphi\|_{L^2(\Omega;\dot{H}^0)} + \|\psi\|_{L^2(\Omega;\dot{H}^{-2})}\right). \end{aligned} \tag{4.34}$$

Finally, we use Itô's isometry, Lemma 4.4 with $\mu = \gamma$ and (3.63) with $\beta = \gamma$ to show

$$\begin{aligned}\|P_2\mathbb{J}_5\|^2_{L^2(\Omega;\dot{H}^0)} &= \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \left\|P_2\left(r(k\mathcal{A}_h)^{n-j}\mathbf{P}_h - \mathcal{S}(t_n - s)\right)\mathbf{B}Q^{\frac{1}{2}}\right\|^2_{\mathrm{HS}}\mathrm{d}s \\ &\le 2\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} \left(\left\|P_2F_{kh}^{n-j}\mathbf{B}Q^{\frac{1}{2}}\right\|^2_{\mathrm{HS}} + \left\|P_2(\mathcal{S}(t_{n-j}) - \mathcal{S}(t_n - s))\mathbf{B}Q^{\frac{1}{2}}\right\|^2_{\mathrm{HS}}\right)\mathrm{d}s \\ &\le C\left(h^{2\gamma} + k^{\gamma}\right)\left\|A^{\frac{\gamma-1}{2}}Q^{\frac{1}{2}}\right\|^2_{\mathrm{HS}}. \end{aligned} \tag{4.35}$$

Gathering (4.31)–(4.35) together implies (4.18) and the proof of this theorem is thus complete. □

5. Numerical examples

In this section we report some numerical experiments to illustrate our previous findings. Let us consider the following strongly damped wave equation, subject to a perturbation of additive noise:

$$\begin{cases} u_{tt} = \Delta u + \Delta u_t - \sin(u) + \dot{W}(t), \quad t \in (0,1], \ \ x \in (0,1), \\ u(0,x) = \frac{\partial u}{\partial t}(0,x) = 0, \ x \in (0,1), \\ u(t,0) = u(t,1) = 0, \ t > 0. \end{cases} \tag{5.1}$$

In the following experiments we aim to test mean-square approximation errors as theoretically measured in (4.3), (4.4), (4.17) and (4.18). The expectation is approximated by the Monte Carlo approximation, using $M = 100$ path simulations. As the first task we examine the spatial approximation errors $\|u(T) - u_h(T)\|_{L^2(\Omega;\dot{H}^0)}$ and $\|u_t(T) - u_{h,t}(T)\|_{L^2(\Omega;\dot{H}^0)}$, with the endpoint $T = 1$ fixed. The 'true' solutions $u(T)$, $u_t(T)$ are identified with numerical ones using small step sizes $h_{\text{exact}} = 2^{-8}, k_{\text{exact}} = 2^{-14}$. The numerical approximations under various spatial mesh sizes $h = 2^{-i}, i = 1, 2, \ldots, 5$ are achieved via time-stepping with $k_{\text{exact}} = 2^{-14}$. The resulting computational errors are listed in Table 1, where two kinds of noises are considered, including the space-time white noise case ($Q = I$) and the trace-class noise case ($Q = A^{-0.5005}$). To clearly see the convergence rates we depict in Fig. 1 the errors versus mesh sizes in logarithmic scale. As expected the slopes of the errors (solid lines) and those of the reference dashed lines match well. More formally, the finite element spatial approximation errors in the space-time white noise case ($Q = I$) exhibit convergence rates of order $\frac{3}{2}$ for the displacement and

TABLE 1 *Mean-square spatial errors for the displacement and the velocity*

	h	$\left(\mathbf{E}\|u(T)-u_h(T)\|^2\right)^{\frac{1}{2}}$	$\left(\mathbf{E}\|u_t(T)-u_{h,t}(T)\|^2\right)^{\frac{1}{2}}$
$Q=I$	1/2	0.017262	0.144681
	1/4	0.006098	0.097764
	1/8	0.002158	0.063434
	1/16	7.694527e-004	0.038866
	1/32	2.723050e-004	0.022045
$Q=A^{-0.5005}$	1/2	0.007918	0.048106
	1/4	0.002289	0.023401
	1/8	6.467743e-004	0.011036
	1/16	1.800250e-004	0.004982
	1/32	4.888214e-005	0.002160

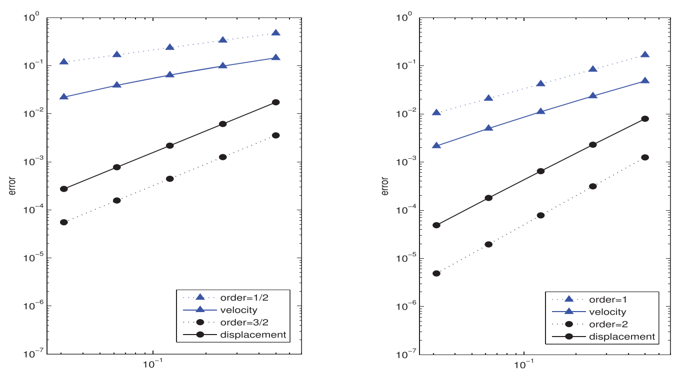

FIG. 1. Mean-square convergence rates for the spatial discretizations (left: $Q=I$, right: $Q=A^{-0.5005}$).

order $\frac{1}{2}$ for the velocity, which coincides with our previous theoretical findings in Theorem 4.1 when $\gamma<\frac{1}{2}$. For the other case ($Q=A^{-0.5005}$ and $\gamma=1$) the errors show the predicted rates of order 2 for the displacement and order 1 for the velocity (see the right plot in Fig. 1).

Next we proceed to tests on the convergence rates of temporal approximations. To this end we fix $h=2^{-7}, T=1$ and measure $\|u_h(T)-U^N\|_{L^2(\Omega;\dot{H}^0)}$ and $\|u_{h,t}(T)-V^N\|_{L^2(\Omega;\dot{H}^0)}$ for five different time step sizes $k=\frac{1}{N}, N=2^3,2^4,\ldots,2^7$. In order to obtain $u_h(T), u_{h,t}(T)$ we perform time-stepping using small time step size $k_{\text{exact}}=2^{-12}$. In Table 2 we present the computational errors for the two noise cases $Q=I$ and $Q=A^{-0.5005}$. Similarly as before these approximation errors are plotted versus time step sizes in Fig. 2, where one can easily observe the expected convergence rates. For example, in the trace-class noise case when $Q=A^{-0.5005}$, the approximation errors for the displacement and the velocity decrease at slopes of 1 and $\frac{1}{2}$, respectively, as the time step sizes decrease. Also, convergence rates of order $\frac{3}{4}$ and $\frac{1}{4}$ are detected for the displacement and the velocity in the space-time white noise case (see the left plot in Fig. 2). All in all, the above observations are all consistent with the previous theoretical results.

TABLE 2 *Mean-square temporal errors for the displacement and the velocity*

	k	$\left(\mathbf{E}\|u_h(T) - U^N\|^2\right)^{\frac{1}{2}}$	$\left(\mathbf{E}\|u_{h,t}(T) - V^N\|^2\right)^{\frac{1}{2}}$
$Q = I$	1/8	0.006226	0.166427
	1/16	0.004302	0.141315
	1/32	0.002560	0.116514
	1/64	0.001332	0.091829
	1/128	6.853130e-004	0.071157
$Q = A^{-0.5005}$	1/8	0.003446	0.068094
	1/16	0.002356	0.052651
	1/32	0.001377	0.038405
	1/64	6.993377e-004	0.025994
	1/128	3.512776e-004	0.017476

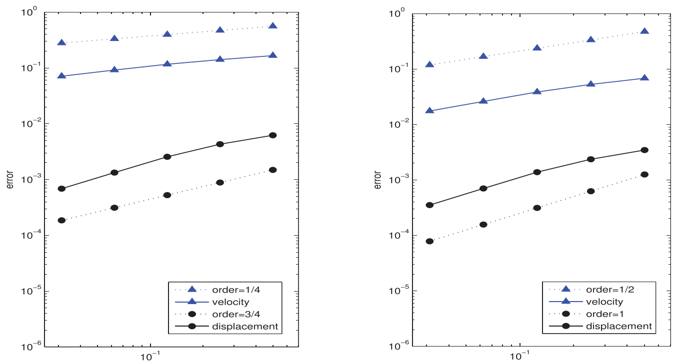

FIG. 2. Mean-square convergence rates for the temporal discretizations (left: $Q = I$, right: $Q = A^{-0.5005}$).

6. Concluding remarks

In this article we analyze strong error estimates of the finite element semidiscretization and full discretization for semilinear SWEs with strongly damping. To do so we significantly enrich existing error estimates of semidiscrete and fully discrete FEMs for the corresponding linear deterministic equation. Based on the obtained error estimates we derive optimal convergence rates for strong approximations of the stochastic problem.

An interesting idea for future works is to consider a more general model with possibly different diffusion operators, e.g., a fractional diffusion operator $L^{\gamma}, \gamma \in (0, 1)$, applied for the damping term. However, this would not make the current analysis persist anymore, and new techniques and arguments are required to handle it. This could be our future work if possible. Another interesting research topic is on ergodicity properties of the underlying model. Similarly to the weakly damping case (Barbu & Da Prato, 2002; Barbu *et al.*, 2007) the strong damping should make the wave equation ergodic with a unique invariant measure. For instance, both weakly and strongly dampings can, under reasonable assumptions, make moments of the mild solution uniformly bounded over long time, which helps to show the existence of invariant measure (see the book by Da Prato & Zabczyk, 1996, Hypothesis 5.1, Theorem 6.1.2 and comments on p. 94 for more details). With such ergodic systems one can further examine ergodicity of their numerical approximations and analyze the resulting approximation errors, as already done in the studies by Bréhier (2014), Bréhier & Kopec (2017) and Chen *et al.* (2017) for other types of stochastic partial differential equations. This also becomes one of our future works.

Acknowledgements

Part of this work was done when the authors attended a seminar on stochastic computations at AMSS, Beijing. The authors thank Jialin Hong and Tao Zhou for their help during the stay.

Funding

National Science Foundation of China (11701073 to R.Q., 11671405, 91630312, 11571373 to X.W.); National Science Foundation of Hebei Province (A2015501102 to R.Q.); National Science Foundation of Hunan Province (2016JJ3137 to X.W.); and Innovation Program of Central South University (2017CX017 to X.W.).

References

Andersson, A. & Larsson, S. (2016) Weak convergence for a spatial approximation of the nonlinear stochastic heat equation. *Math. Comput.*, **85**, 1335–1358.

Anton, R., Cohen, D., Larsson, S. & Wang, X. (2016) Full discretisation of semi-linear stochastic wave equations driven by multiplicative noise. *SIAM J. Numer. Anal.*, **54**, 1093–1119.

Barbu, V. & Da Prato, G. (2002) The stochastic nonlinear damped wave equation. *Appl. Math. Optim.*, **46**, 125–141.

Barbu, V., Da Prato, G. & Tubaro, L. (2007) Stochastic wave equations with dissipative damping. *Stochastic Process. Appl.*, **117**, 1001–1013.

Bréhier, C.-E. (2014) Approximation of the invariant measure with an Euler scheme for stochastic PDEs driven by space-time white noise. *Potential Anal.*, **40**, 1–40.

Bréhier, C.-E. & Kopec, M. (2017) Approximation of the invariant law of SPDEs: error analysis using a Poisson equation for a full-discretization scheme. *IMA J. Numer. Anal.*, **37**, 1375–1410.

Cao, Y. & Yin, L. (2007) Spectral Galerkin method for stochastic wave equations driven by space-time white noise. *Commun. Pure Appl. Anal.*, **6**, 607.

Chen, C., Hong, J. & Wang, X. (2017) Approximation of invariant measure for damped stochastic nonlinear Schrodinger equation via an ergodic numerical scheme. *Potential Anal.*, **46**, 323–367.

Cohen, D., Larsson, S. & Sigg, M. (2013) A trigonometric method for the linear stochastic wave equation. *SIAM J. Numer. Anal.*, **51**, 204–222.

Cohen, D. & Quer-Sardanyons, L. (2016) A fully discrete approximation of the one-dimensional stochastic wave equation. *IMA J. Numer. Anal.*, **36**, 400–420.

Da Prato, G. & Zabczyk, J. (1996) *Ergodicity for Infinite Dimensional Systems*, vol. 229. Cambridge: Cambridge University Press.

Da Prato, G. & Zabczyk, J. (2014) *Stochastic Equations in Infinite Dimensions*, vol. 152. Cambridge: Cambridge University Press.

Fitzgibbon, W. (1981) Strongly damped quasilinear evolution equations. *J. Math. Anal. Appl.*, **79**, 536–550.

Hausenblas, E. (2010) Weak approximation of the stochastic wave equation. *J. Comput. Appl. Math.*, **235**, 33–58.

Jacobe de Naurois, L., Jentzen, A. & Welti, T. (2015) Weak convergence rates for spatial spectral Galerkin approximations of semilinear stochastic wave equations with multiplicative noise. *Applied Mathematics & Optimization.* arXiv:1508.05168.

Jentzen, A. & Kloeden, P. E. (2011) *Taylor Approximations for Stochastic Partial Differential Equations*. Philadelphia: SIAM.

Kloeden, P. E. & Platen, E. (1992) *Numerical Solution of Stochastic Differential Equations*. Berlin: Springer.

Kovács, M., Larsson, S. & Lindgren, F. (2012) Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise. *BIT Numer. Math.*, **52**, 85–108.

Kovács, M., Larsson, S. & Lindgren, F. (2013) Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise II. Fully discrete schemes. *BIT Numer. Math.*, **53**, 497–525.

Kovács, M., Larsson, S. & Saedpanah, F. (2010) Finite element approximation of the linear stochastic wave equation with additive noise. *SIAM J. Numer. Anal.*, **48**, 408–427.

Kruse, R. (2013) *Strong and Weak Approximation of Semilinear Stochastic Evolution Equations*, vol. 2093. Berlin: Springer.

Kruse, R. (2014) Optimal error estimates of Galerkin finite element methods for stochastic partial differential equations with multiplicative noise. *IMA J. Numer. Anal.*, **34**, 217–251.

Larsson, S., Thomée, V. & Wahlbin, L. B. (1991) Finite-element methods for a strongly damped wave equation. *IMA J. Numer. Anal.*, **11**, 115–142.

Lord, G. J., Powell, C. E. & Shardlow, T. (2014) *An Introduction to Computational Stochastic PDEs*. Number 50. New York: Cambridge University Press.

Massatt, P. (1983) Limiting behavior for strongly damped nonlinear wave equations. *J. Differential Equations*, **48**, 334–349.

Pata, V. & Squassina, M. (2005) On the strongly damped wave equation. *Commun. Math. Phys.*, **253**, 511–533.

Qi, R. & Wang, X. (2017) An accelerated exponential time integrator for semi-linear stochastic strongly damped wave equation with additive noise. *J. Math. Anal. Appl.*, **447**, 988–1008.

Qi, R. & Yang, X. (2013) Weak convergence of finite element method for stochastic elastic equation driven by additive noise. *J. Sci. Comput.*, **56**, 450–470.

Qi, R., Yang, X. & Zhang, Y. (2013) Full-discrete finite element method for the stochastic elastic equation driven by additive noise. *Numer. Methods Partial Differential Equations*, **29**, 1946–1962.

Quer-Sardanyons, L. & Sanz-Solé, M. (2006) Space semi-discretisations for a stochastic wave equation. *Potential Anal.*, **24**, 303–332.

Thomée, V. (2006) *Galerkin Finite Element Methods for Parabolic Problems*. Berlin: Springer.

Thomée, V. & Wahlbin, L. (2004) Maximum-norm estimates for finite-element methods for a strongly damped wave equation. *BIT Numer. Math.*, **44**, 165–179.

Walsh, J. B. (2006) On numerical solutions of the stochastic wave equation. *Illinois J. Math.*, **50**, 991–1018.

Wang, X. (2015) An exponential integrator scheme for time discretization of nonlinear stochastic wave equation. *J. Sci. Comput.*, **64**, 234–263.

Wang, X., Gan, S. & Tang, J. (2014) Higher order strong approximations of semilinear stochastic wave equation with additive space-time white noise. *SIAM J. Sci. Comput.*, **36**, A2611–A2632.

Yan, Y. (2005) Galerkin finite element methods for stochastic parabolic partial differential equations. *SIAM J. Numer. Anal.*, **43**, 1363–1384.