Mittag-Leffler Euler Integrator for a Stochastic Fractional Order Equation with Additive Noise

Citation for the original published paper (version of record):

Kovacs, M., Larsson, S., Saedpanah, F. (2020)
Mittag-Leffler Euler Integrator for a Stochastic Fractional Order Equation with Additive Noise
SIAM Journal on Numerical Analysis, 58(1): 66-85
http://dx.doi.org/10.1137/18M1177895

SIAM J. NUMER. ANAL.
Vol. 58, No. 1, pp. 66–85

MITTAG–LEFFLER EULER INTEGRATOR FOR A STOCHASTIC FRACTIONAL ORDER EQUATION WITH ADDITIVE NOISE*

MIHÁLY KOVÁCS†, STIG LARSSON‡, AND FARDIN SAEDPANAH§

Abstract. Motivated by fractional derivative models in viscoelasticity, a class of semilinear stochastic Volterra integro-differential equations, and their deterministic counterparts, are considered. A generalized exponential Euler method, named here the Mittag–Leffler Euler integrator, is used for the temporal discretization, while the spatial discretization is performed by the spectral Galerkin method. The temporal rate of strong convergence is found to be (almost) twice compared to when the backward Euler method is used together with a convolution quadrature for time discretization. Numerical experiments that validate the theory are presented.

Key words. Euler integrator, fractional equations, stochastic differential equations, strong convergence, integro-differential equations, Riesz kernel

AMS subject classifications. 34A08, 45D05, 45K05, 60H15, 60H35, 65M12, 65M60

DOI. 10.1137/18M1177895

1. Introduction. We study the numerical approximation of a class of semilinear Volterra integro-differential equations in a real, separable, infinite-dimensional Hilbert space H of the form

$$\mathrm{d}u(t) + \int_0^t b(t-s)Au(s)\,\mathrm{d}s\,\mathrm{d}t = F(u(t))\,\mathrm{d}t + \mathrm{d}W(t),\ t \in (0,T];\ u(0) = u_0, \tag{1.1}$$

where A is a self-adjoint, positive definite, not necessarily bounded, operator on the Hilbert space H, W is an H-valued Wiener process with covariance operator Q, $F\colon H \to H$ is a nonlinear operator, b is a locally integrable scalar kernel, and u_0 is an H-valued random variable. Our main example of b is the Riesz kernel

$$b(t) = \frac{t^{\alpha-1}}{\Gamma(\alpha)}, \quad 0 < \alpha < 1, \tag{1.2}$$

where $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}\mathrm{e}^{-t}\,\mathrm{d}t$ is the gamma function.

By introducing the fractional integral of order α denoted by J_0^α (see, for example, [18]) as

$$(J_0^\alpha g)(t) = \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}g(s)\,\mathrm{d}s, \quad 0 < \alpha < 1,$$

(1.1) becomes a fractional order equation of the form

$$\mathrm{d}u(t) + J_0^\alpha(Au)(t)\,\mathrm{d}t = F(u(t))\,\mathrm{d}t + \mathrm{d}W(t),\ t \in (0,T]; \quad u(0) = u_0. \tag{1.3}$$

*Received by the editors March 28, 2018; accepted for publication (in revised form) November 4, 2019; published electronically January 7, 2020.
https://doi.org/10.1137/18M1177895
Funding: The work of the authors was partially supported by Nordforsk grant 74756.
†Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE-41296 Gothenburg, Sweden, and Pázmány Péter Catholic University, Faculty of Information Technology and Bionics, H-1444 Budapest, Hungary (kovacs.mihaly@itk.ppke.hu).
‡Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE-41296 Gothenburg, Sweden (stig@chalmers.se).
§Department of Mathematics, University of Kurdistan, Sanandaj, Iran (f.saedpanah@uok.ac.ir, fardin.saedpanah@gmail.com).

We note that the present framework applies also to slightly more general kernels, which have similar smoothing effects, such as the tempered Riesz kernel

$$b(t) = \frac{1}{\Gamma(\alpha)} t^{\alpha-1} \mathrm{e}^{-\eta t}, \quad 0 < \alpha < 1,\ \eta \geq 0, \tag{1.4}$$

and even to certain kernels with finite smoothness; see Remark 1 for further discussion. Recalling that $0 < \alpha < 1$ in (1.2), we denote henceforth

$$\rho = \alpha + 1, \quad 1 < \rho < 2, \tag{1.5}$$

so that $b(t) = t^{\rho-2}/\Gamma(\rho-1)$.

We motivate our main example (1.3) by a model from linear viscoelasticity; for more examples see, e.g., [16, 20] and references therein. In one spatial dimension, considering the class of viscoelastic materials which exhibit a simple power-law creep, the evolution equation that describes the response variable w (chosen among the field variables: displacement, stress, the strain, or the particle velocity) is given by

$$w(x,t) = w(x,0+) + tw_t(x,0+) + \frac{c}{\Gamma(\rho)} \int_0^t (t-s)^{\rho-1} w_{xx}(x,s)\,\mathrm{d}s, \quad 1 < \rho < 2;$$

see, for example, [16]. Assuming that $w_t(x,0+) = 0$ and that w is continuous at $t = 0+$ with $w(x,0+) = w_0(x)$, one arrives at the Cauchy problem

$$w_t(x,t) = \frac{c}{\Gamma(\rho-1)} \int_0^t (t-s)^{\rho-2} w_{xx}(x,s)\,\mathrm{d}s; \quad w(x,0) = w_0(x).$$

With $\alpha = \rho - 1$ and $c = 1$ we get

$$w_t(x,t) = J_0^\alpha(w_{xx}(x,\cdot))(t); \quad w(x,0) = w_0(x).$$

Now, if w is chosen to be the particle velocity, and f represents a nonlinear, external viscous force, which is perturbed by Gaussian noise $\dot{\xi}$, then the equation for the particle velocity reads as

$$w_t(x,t) = J_0^\alpha(w_{xx}(x,\cdot))(t) + f(w(x,t)) + \dot{\xi}(x,t); \quad w(x,0) = w_0(x). \tag{1.6}$$

Considering the equation on an interval $[0, L]$ and supplementing the equation with nonslip boundary conditions, we arrive at a special instance of (1.3), with $H = L_2(0,L)$, A being the Dirichlet Laplacian in H and F the Nemytskij operator $F(v)(\cdot) = f(v(\cdot))$. We remark that, without the noise and f, (1.6) is often referred to as a fractional wave equation.

We note that when the kernel b in (1.1) is smooth, e.g., exponential kernels, these equations reveal a hyperbolic behavior, whereas for weakly singular kernels, e.g., the Riesz kernel (1.2), they exhibit certain parabolic features.

The literature on numerical methods for stochastic PDEs, such as stochastic parabolic and hyperbolic PDEs, is mature. In some works, by using exponential integrators [8], the strong rate of convergence has been improved for the stochastic heat equation (see, e.g., [4, 14, 23]) and for the stochastic wave equation (see, e.g., [2] and the references therein). The drawback of the exponential integrators for stochastic PDEs is that the eigenfunctions of the operator A and of the covariance operator Q of the noise must coincide and must be known explicitly, so that the scheme can be implemented.

However, the literature on numerical analysis of stochastic Volterra equations is more scarce, containing only [1, 11, 12] and a few recent papers specifically for the fractional stochastic heat equation (where there is a derivative in front of J_0^α in (1.3)), based on a convolution quadrature; see, for example, [6, 7].

Here, we study a full discretization of (1.1), as well as its deterministic counterpart, i.e., the special case when $W = 0$. We use a generalized exponential Euler method, named here the Mittag–Leffler Euler integrator, for the temporal discretization. Full discretization is then formulated by the spectral Galerkin method for spatial discretization.

As is the case for stochastic equations with no memory effects, the time integration is based on the mild formulation of the equation. However, there is a major difference, namely, the solution operator in our case does not have the semigroup property and hence the integrator uses the approximate solution from all previous time levels, not just from the current one, in order to advance to the next time level. This phenomenon is of course present in the convolution quadrature setting as well and makes the error analysis more difficult compared to the memoryless case.

The main novelty in this work is the introduction of a new temporal discretization method for (1.1) and its error analysis. The analysis of the spatial discretisation is more or less standard. In particular, we prove that the strong rate of temporal convergence is (almost) twice the rate of the Euler method combined with Lubich's convolution quadrature of order 1 [1]. As a consequence, for trace-class noise, we recover (almost) the optimal rate 1 in time.

When $H = L_2(\mathcal{D})$, where $\mathcal{D} \subset \mathbb{R}^d$ is a bounded domain, with appropriately smooth boundary, the framework presented here allows for a general class of Nemytskij operators F when $d = 1, 2, 3$, with some restriction on ρ when $d = 3$. For space-time white noise we must have $d = 1$, while for colored noise $d > 1$ is allowed.

The outline of the paper is as follows. In section 2, we introduce notation and the abstract framework, state our main assumptions, and present some preliminary results on the solution of (1.1). In section 3, we introduce the numerical scheme (3.7) and, in Theorem 1, we state and prove our main result on the order of strong convergence. In section 4, we discuss the implementation of the scheme and present some numerical experiments to illustrate the theory. Throughout the paper C denotes a generic constant that may have different values at different occurrences, but its value is independent of the discretization parameters.

2. Preliminaries.

2.1. The abstract setting. Let H be a real, separable, infinite-dimensional Hilbert space with inner product $(\cdot,\cdot)$ and norm $\|\cdot\|$ and let A be a self-adjoint, positive definite, not necessarily bounded operator in H with compact inverse. An important example is $H = L_2(\mathcal{D})$ and $A = -\Delta$ with homogeneous Dirichlet boundary conditions. Let $\{(\lambda_k, \varphi_k)\}_{k=1}^\infty$ be the eigenpairs of A, i.e.,

$$A\varphi_k = \lambda_k \varphi_k, \quad k \in \mathbb{N}. \tag{2.1}$$

It is known that $0 < \lambda_1 \le \lambda_2 \le \cdots \le \lambda_k \le \cdots$ with $\lim_{k\to\infty} \lambda_k = \infty$ and the eigenvectors $\{\varphi_k\}_{k=1}^\infty$ form an orthonormal basis for H. We introduce the fractional order spaces

$$\dot H^\nu := \operatorname{dom}\left(A^{\frac{\nu}{2}}\right), \quad \|v\|_\nu^2 := \left\|A^{\frac{\nu}{2}} v\right\|^2 = \sum_{k=1}^\infty \lambda_k^\nu (v, \varphi_k)^2, \quad \nu \in \mathbb{R},\ v \in \dot H^\nu.$$

Let $\mathcal{L} = \mathcal{L}(H)$ denote the space of all bounded linear operators on H. We also consider the space of Hilbert–Schmidt operators, that is, the space of all operators $T \in \mathcal{L}$ for which

$$\|T\|_{\mathrm{HS}}^2 = \sum_{k=1}^{\infty} \|T\varphi_k\|^2 < \infty.$$

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space, $(\mathcal{F}_t)_{t\in[0,T]}$ being a normal filtration, with Bochner spaces $L_p(\Omega; H) = L_p\big((\Omega, \mathcal{F}, \mathbb{P}); H\big)$, $p \geq 2$. We let $Q \in \mathcal{L}$ be a self-adjoint, positive semidefinite operator and $H_0 = Q^{\frac{1}{2}}(H)$ be the Hilbert space with the inner product $\langle u, v\rangle_{H_0} = \langle Q^{-\frac{1}{2}}u, Q^{-\frac{1}{2}}v\rangle$, where $Q^{-\frac{1}{2}}$ denotes the pseudoinverse of $Q^{\frac{1}{2}}$, when it is not injective, and $Q^{\frac{1}{2}}$ is the unique positive semidefinite square root of Q. By $\mathcal{L}_2^0 = \mathcal{L}_2^0(H)$ we denote the space of Hilbert–Schmidt operators $H_0 \to H$. Thus, $\|T\|_{\mathcal{L}_2^0} = \|TQ^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$, for $T \in \mathcal{L}_2^0$. Then we let W be Q-Wiener process in H with respect to $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$. We recall the Itô isometry,

$$\left\| \int_0^t \phi(s)\,\mathrm{d}W(s) \right\|_{L_2(\Omega;H)} = \left\| \left(\int_0^t \|\phi(s)\|_{\mathcal{L}_2^0}^2 \,\mathrm{d}s \right)^{\frac{1}{2}} \right\|_{L_2(\Omega;\mathbb{R})}, \tag{2.2}$$

and the Burkholder–Davis–Gundy inequality, for $p \geq 2$,

$$\left\| \int_0^t \phi(s)\,\mathrm{d}W(s) \right\|_{L_p(\Omega;H)} \leq C_p \left\| \left(\int_0^t \|\phi(s)\|_{\mathcal{L}_2^0}^2 \,\mathrm{d}s \right)^{\frac{1}{2}} \right\|_{L_p(\Omega;\mathbb{R})}, \tag{2.3}$$

for strongly measurable functions $\phi\colon [0,T] \to \mathcal{L}_2^0$ [5].

We recall from (1.5) that $\rho = \alpha + 1$, $\rho \in (1,2)$, $\alpha \in (0,1)$.

Assumption 1. We quantify the regularity of the noise by $\beta \in (0, \frac{1}{\rho}]$ through the assumption that there is a constant B such that

$$\left\| A^{\frac{\beta-\frac{1}{\rho}}{2}} \right\|_{\mathcal{L}_2^0} = \left\| A^{\frac{\beta-\frac{1}{\rho}}{2}} Q^{\frac{1}{2}} \right\|_{\mathrm{HS}} \leq B. \tag{2.4}$$

Trace class noise, that is, when $\mathrm{Tr}(Q) = \|Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 < \infty$, corresponds to $\beta = \frac{1}{\rho}$. When $A = -\Delta$ is the Dirichlet Laplacian, we may take $Q = A^{-s}$ with $s \geq 0$. Then (2.4) is satisfied with $\beta < s + \frac{1}{\rho} - \frac{d}{2}$, because $\lambda_j \approx j^{\frac{2}{d}}$ as $j \to \infty$. We note that $s = 0$ corresponds to space-time white noise $Q = I$, and in this case, $d = 1$ and $\beta < \frac{1}{\rho} - \frac{1}{2}$.

2.2. The linear deterministic problem. We assume that there exists a strongly continuous family $\{S(t)\}_{t\geq 0}$ of bounded linear operators on H such that the function $u(t) = S(t)u_0$, $u_0 \in H$, is the unique solution of

$$u(t) + A\int_0^t B(t-s)u(s)\,\mathrm{d}s = u_0, \quad t \geq 0,$$

with $B(t) = \int_0^t b(s)\,\mathrm{d}s$. When $t \to u(t) = S(t)u_0$ is differentiable for $t > 0$, then u is the unique solution of

$$\dot{u}(t) + A\int_0^t b(t-s)u(s)\,\mathrm{d}s = 0,\ t > 0; \quad u(0) = u_0.$$

We refer to the monograph [20] for a comprehensive theory of resolvent families for Volterra equations. An important feature of the resolvent family $\{S(t)\}_{t\geq 0}$ is that it does not have the semigroup property; that is, $S(t+s) \neq S(t)S(s)$. This is the mathematical reflection of the fact that the solution possesses a nontrivial memory. In our special setting, using the spectral decomposition of A, an explicit representation of $S(t)$ is given by the Fourier series

$$S(t)v = \sum_{k=1}^{\infty} s_k(t)(v, \varphi_k)\varphi_k, \tag{2.5}$$

where the functions $s_k(t)$ are the solutions of

$$\dot{s}_k(t) + \lambda_k \int_0^t b(t-s)s_k(s)\,\mathrm{d}s = 0,\ t>0; \quad s_k(0) = 1. \tag{2.6}$$

Next, we collect our precise assumptions on the resolvent family $\{S(t)\}_{t\geq 0}$.

Assumption 2. We assume that the resolvent family $\{S(t)\}_{t\geq 0}$ is strongly continuously differentiable for $t>0$ and enjoys the following smoothing properties: There is M such that for $t>0$, we have

$$\|A^s S(t)\|_{\mathcal{L}} \leq M t^{-s\rho}, \qquad s \in \left[0, \frac{1}{\rho}\right], \tag{2.7}$$

$$\|A^s \dot{S}(t)\|_{\mathcal{L}} \leq M t^{-s\rho-1}, \qquad s \in \left[0, \frac{1}{\rho}\right], \tag{2.8}$$

$$\|A^{-s} \dot{S}(t)\|_{\mathcal{L}} \leq M t^{s\rho-1}, \qquad s \in [0,1]. \tag{2.9}$$

Remark 1. These are verified in [17, Theorem 5.5] for the Riesz kernel and in [3, Lemma A.4] for more general kernels. We note that for the Riesz kernel (1.2), which is our main example, estimates (2.7) and (2.8) hold also for $s\in[0,1]$ (see [17, Theorem 5.5]), but we do not need this extended range of s for the present analysis. A more general class of kernels b for which (2.7)–(2.9) are satisfied is the class of 4-monotone kernels such that $0 \neq b \in L_{1,\mathrm{loc}}(\mathbb{R}_+)$, $\lim_{t\to\infty} b(t) = 0$, with

$$\rho := 1 + \frac{2}{\pi}\sup\left\{|\arg \widehat{b}(z)| : \operatorname{Re} z > 0\right\} \in (1,2),$$

and $\widehat{b}(z) \leq Cz^{1-\rho}$ for $z>1$, where this latter condition may be substituted by the condition $\|b\|_{L_1(0,t)} \leq Ct^{\rho-1}$, $t\in(0,1)$; see [3, Remark 3.8 and Lemma A.4]. In particular, b does not have to be analytic. (Here $\widehat{b}$ denotes the Laplace transform of b.)

2.3. Well-posedness of the semilinear stochastic problem. The mild solution of the semilinear stochastic equation (1.1) is an adapted H-valued stochastic process, $u(t)$, such that, for $t\in[0,T]$, $\mathbb{P}$-almost surely,

$$u(t) = S(t)u_0 + \int_0^t S(t-s)F(u(s))\,\mathrm{d}s + \int_0^t S(t-s)\,\mathrm{d}W(s). \tag{2.10}$$

Assumption 3. In addition to the singularity exponent $\rho = \alpha+1 \in (1,2)$ from (1.2) and the regularity parameter $\beta \in (0, \frac{1}{\rho}]$ in (2.4), we assume that there are $\delta \in [1, \frac{2}{\rho})$, $\gamma \in [0,\beta)$, $\eta \in [1, \frac{2}{\rho})$, and a constant $L>0$, such that

$$\|F(u)\| \le L(1+\|u\|), \quad \|F'(u)v\| \le L\|v\|, \qquad u, v \in H, \tag{2.11}$$

$$\|F'(u)v\|_{-\delta} \le L(1+\|u\|_\gamma)\|v\|_{-\gamma}, \qquad u \in \dot H^\gamma,\ v \in \dot H^{-\gamma}, \tag{2.12}$$

$$\|F''(u)(v_1, v_2)\|_{-\eta} \le L\|v_1\|\|v_2\|, \qquad v_1, v_2 \in H. \tag{2.13}$$

Our main example is $H = L_2(\mathcal{D})$ with $\mathcal{D} \subset \mathbb{R}^d$ a bounded domain with appropriately smooth boundary, and $A = -\Delta$, the negative of the Dirichlet Laplacian. Here F can be taken to be a Nemytskij operator defined by $F(u)(x) = f(u(x))$, where $f \colon \mathbb{R} \to \mathbb{R}$ is a smooth function with bounded derivatives of orders 1 and 2. Then (2.11) clearly holds and (2.13) is satisfied with $\eta > d/2$ because of Sobolev's inequality. The additional assumption $\eta < \frac{2}{\rho}$ puts a restriction on ρ, namely, $1 < \rho < 4/d$. For (2.12) we refer to Lemma 4.4 in [22], which can be extended from $d = 1$ to $d \le 3$, again in case $\delta > d/2$ and thus $1 < \rho < 4/d$.

Lemma 1. *Suppose that Assumption* 1, (2.7) *from Assumption* 2*, and* (2.11) *from Assumption* 3 *hold. Let $p \ge 2$, and assume $\|u_0\|_{L_p(\Omega;\dot H^\gamma)} \le K$. Then, there is a unique mild solution $u \in \mathcal{C}([0,T]; L_p(\Omega;H))$ of* (2.10)*. Furthermore, for a constant $C = C(B,K,L,M,T,\beta,\gamma,\rho,p)$,*

$$\sup_{t\in[0,T]} \|u(t)\|_{L_p(\Omega;\dot H^\gamma)} \le C. \tag{2.14}$$

Proof. The existence and uniqueness of a mild solution $u \in \mathcal{C}([0,T]; L_p(\Omega;H))$ of (2.10) can be proved, even only under assumption (2.11), via a standard Banach fixed point argument using (2.4) and (2.7); see, for example, the proof of [3, Theorem 3.3]. Therefore,

$$\|u(t)\|_{L_p(\Omega;H)} \le C, \quad t \in [0,T], \tag{2.15}$$

which is (2.14) with $\gamma = 0$. For $\gamma \in (0,\beta)$, using (2.10), we have

$$\begin{aligned}\|u(t)\|_{L_p(\Omega;\dot H^\gamma)} \le{}& \|S(t)\|_{\mathcal{L}}\|u_0\|_{L_p(\Omega;\dot H^\gamma)} \\ &+ \int_0^t \|A^{\frac{\gamma}{2}}S(t-s)\|_{\mathcal{L}}\|F(u(s))\|_{L_p(\Omega;H)}\,\mathrm{d}s \\ &+ \Big\|\int_0^t A^{\frac{\gamma}{2}}S(t-s)\,\mathrm{d}W(s)\Big\|_{L_p(\Omega;H)}.\end{aligned}$$

By using (2.7) with $s = 0$, (2.7), (2.11), (2.3), and (2.15), we obtain

$$\begin{aligned}\|u(t)\|_{L_p(\Omega;\dot H^\gamma)} &\le \|u_0\|_{L_p(\Omega;\dot H^\gamma)} + L\int_0^t \|A^{\frac{\gamma}{2}}S(t-s)\|_{\mathcal{L}}\big(1+\|u(s)\|_{L_p(\Omega;H)}\big)\,\mathrm{d}s \\ &\quad + C_p\left\|\left(\int_0^t \|A^{\frac{\gamma}{2}}S(t-s)Q^{\frac12}\|_{\mathrm{HS}}^2\,\mathrm{d}s\right)^{\frac12}\right\|_{L_p(\Omega;\mathbb{R})} \\ &\le C + C\int_0^t (t-s)^{-\frac{\gamma\rho}{2}}\,\mathrm{d}s \\ &\quad + C_p\left\|A^{\frac{\beta-\frac{1}{\rho}}{2}}Q^{\frac12}\right\|_{\mathrm{HS}}\left(\int_0^t \left\|A^{\frac{(\gamma-\beta)+\frac{1}{\rho}}{2}}S(t-s)\right\|^2\mathrm{d}s\right)^{\frac12}.\end{aligned}$$

By using (2.4) and (2.7) again, we have

$$\|u(t)\|_{L_p(\Omega;\dot H^\gamma)} \le C + BC\left(\int_0^t (t-s)^{-1+(\beta-\gamma)\rho}\,\mathrm{d}s\right)^{\frac12},$$

where the integral is finite, since $(\beta-\gamma)\rho \in (0,1)$. This completes the proof. ∎

Remark 2. In the deterministic case, i.e., when $W = 0$, by following the proof of Lemma 1, it is straightforward to prove that, assuming $u_0 \in \dot{H}^{2\gamma}$ for some $\gamma \in [0, \frac{1}{\rho})$, we have the regularity estimate

$$\sup_{t\in[0,T]} \|u(t)\|_{\dot{H}^{2\gamma}} \leq C. \tag{2.16}$$

3. Full discretization. In this section we formulate a fully discrete method for approximation of (1.1). We use the spectral Galerkin method for spatial discretization in combination with a time discretization based on an exponential Euler type method. We refer to the proposed time discretization method as the Mittag–Leffler Euler integrator (MLEI) since the solution operator can be represented using the Mittag–Leffler function, in case the convolution kernel b is the Riesz kernel as in our main example (1.3). We give more details in section 4, where numerical examples are presented.

Let $0 = t_0 < t_1 < \cdots < t_M = T$ be a uniform partition of the time interval $[0, T]$, with time step $\Delta t = t_{m+1} - t_m,\ m = 0, 1, \ldots, M-1$. Then, for $m = 0, 1, \ldots, M$, by using the variation of constants formula (2.10), we have

$$\begin{aligned} u(t_m) &= S(t_m)u_0 + \int_0^{t_m} S(t_m-\sigma)F(u(\sigma))\,\mathrm{d}\sigma + \int_0^{t_m} S(t_m-\sigma)\,\mathrm{d}W(\sigma) \\ &= S(t_m)u_0 + \sum_{j=0}^{m-1}\int_{t_j}^{t_{j+1}} S(t_m-\sigma)F(u(\sigma))\,\mathrm{d}\sigma + \int_0^{t_m} S(t_m-\sigma)\,\mathrm{d}W(\sigma), \end{aligned} \tag{3.1}$$

Following the idea of exponential integrators, we formulate the MLEI as

$$U_m = S(t_m)u_0 + \sum_{j=0}^{m-1}\int_{t_j}^{t_{j+1}} S(t_m-\sigma)\,\mathrm{d}\sigma\, F(U_j) + \int_0^{t_m} S(t_m-\sigma)\,\mathrm{d}W(\sigma), \tag{3.2}$$

where $U_j \approx u(t_j)$, $j = 0, 1, \ldots, M$, and where the convolution containing the nonlinear term is approximated but the linear terms, including the stochastic convolution integral, are computed exactly; see section 4 for details.

For spatial discretization, we define finite-dimensional subspaces H_N of H by $H_N = \mathrm{span}\{\varphi_1, \varphi_2, \ldots, \varphi_N\}$, where $\{\varphi_k\}_{k=1}^{\infty}$ are the eigenvectors of A, (2.1). Then we define the projector

$$\mathcal{P}_N \colon H \to H_N, \quad \mathcal{P}_N v = \sum_{k=1}^{N} (v, \varphi_k)\varphi_k, \quad v \in H. \tag{3.3}$$

We also define the operator

$$A_N \colon H_N \to H_N, \quad A_N = A\mathcal{P}_N, \tag{3.4}$$

which generates a family of resolvent operators $\{S_N(t)\}_{t\geq 0}$ in H_N. It is known that

$$S_N(t)\mathcal{P}_N = S(t)\mathcal{P}_N, \tag{3.5}$$

$$\|A^{-\nu}(I-\mathcal{P}_N)\| = \sup_{k\geq N+1} \lambda_k^{-\nu} = \lambda_{N+1}^{-\nu}, \quad \nu \geq 0. \tag{3.6}$$

The representation of S_N, similar to (2.5), is given by

$$S_N(t)v = \sum_{k=1}^{N} s_k(t)(v, \varphi_k)\varphi_k.$$

Therefore, the smoothing properties (2.7)–(2.9) also hold for S_N with constants independent of N.

Hence, the fully discrete approximation of (1.1), based on the temporal approximation (3.2), is given by

$$\begin{aligned} U_m^N = S_N(t_m)\mathcal{P}_N u_0 + \sum_{j=0}^{m-1}\int_{t_j}^{t_{j+1}} S_N(t_m-\sigma)\,\mathrm{d}\sigma\,\mathcal{P}_N F\left(U_j^N\right) \\ + \int_0^{t_m} S_N(t_m-\sigma)\mathcal{P}_N\,\mathrm{d}W(\sigma), \end{aligned} \tag{3.7}$$

with initial value $U_0^N = \mathcal{P}_N u_0$. Now we state and prove the main theorem, which shows the strong rate of convergence.

THEOREM 1. *Suppose that Assumptions* 1, 2, *and* 3 *hold and* $\|u_0\|_{L_4(\Omega;\dot{H}^\gamma)} \le K$. *Then, for a constant* $C = C(B,K,L,T,\beta,\rho,\gamma)$, *we have*

$$\sup_{t_m\in[0,T]} \|u(t_m) - U_m^N\|_{L_2(\Omega;H)} \le C\left(\lambda_N^{-\frac{\gamma}{2}} + \Delta t^{\gamma\rho}\right).$$

Proof. By subtracting (3.7) from (3.1), we get

$$\begin{aligned} u(t_m) - U_m^N &= S(t_m)u_0 - S_N(t_m)\mathcal{P}_N u_0 \\ &\quad + \sum_{j=0}^{m-1}\int_{t_j}^{t_{j+1}} \left\{S(t_m-\sigma)F(u(\sigma)) - S_N(t_m-\sigma)\mathcal{P}_N F\left(U_j^N\right)\right\}\mathrm{d}\sigma \\ &\quad + \int_0^{t_m}\left\{S(t_m-\sigma) - S_N(t_m-\sigma)\mathcal{P}_N\right\}\mathrm{d}W(\sigma). \end{aligned}$$

By recalling (3.5) and taking norms, we obtain

$$\begin{aligned} \left\|u(t_m) - U_m^N\right\|_{L_2(\Omega;H)} &\le \|S(t_m)(I-\mathcal{P}_N)u_0\|_{L_2(\Omega;H)} \\ &\quad + \left\|\int_0^{t_m} S(t_m-\sigma)(I-\mathcal{P}_N)F(u(\sigma))\,\mathrm{d}\sigma\right\|_{L_2(\Omega;H)} \\ &\quad + \left\|\sum_{j=0}^{m-1}\int_{t_j}^{t_{j+1}} S(t_m-\sigma)\mathcal{P}_N\big(F(u(\sigma)) - F(U_j^N)\big)\,\mathrm{d}\sigma\right\|_{L_2(\Omega;H)} \\ &\quad + \left\|\int_0^{t_m} S(t_m-\sigma)(I-\mathcal{P}_N)\,\mathrm{d}W(\sigma)\right\|_{L_2(\Omega;H)} = \sum_{l=1}^{4} I_l. \end{aligned} \tag{3.8}$$

We note that I_1, I_2, and I_4 correspond to the *spatial discretization error*, while I_3 corresponds to the *temporal error*.

1. Spatial error. The estimate of I_1 is a consequence of (2.7) with $s=0$ and (3.6), as

$$\begin{aligned} I_1 &\le \|S(t_m)\|_{\mathcal{L}}\|A^{-\frac{\gamma}{2}}(I-\mathcal{P}_N)A^{\frac{\gamma}{2}}u_0\|_{L_2(\Omega;H)} \\ &\le C\lambda_{N+1}^{-\frac{\gamma}{2}}\|u_0\|_{L_2(\Omega;\dot{H}^\gamma)} \le C\lambda_{N+1}^{-\frac{\gamma}{2}}. \end{aligned} \tag{3.9}$$

For I_2, by using (2.7) and (3.6), we have

$$\begin{aligned} I_2 &\le \int_0^{t_m} \|A^{\gamma} S(t_m-\sigma)\|_{\mathcal{L}} \\ &\qquad \times \|A^{-\gamma}(I-\mathcal{P}_N)\|_{\mathcal{L}} \|F(u(\sigma))\|_{L_2(\Omega;H)}\,\mathrm{d}\sigma \\ &\le C\int_0^{t_m} (t_m-\sigma)^{-\gamma\rho}\lambda_{N+1}^{-\gamma}\|F(u(\sigma))\|_{L_2(\Omega;H)}\,\mathrm{d}\sigma \le C\lambda_{N+1}^{-\gamma}, \end{aligned} \tag{3.10}$$

where we recall that $\gamma\rho<1$ and use (2.11) and (2.14) with $p=2, \gamma=0$.

Now we estimate I_4. Using the Itô isometry (2.2), we have

$$\begin{aligned} I_4 &\le \left\| \left(\int_0^{t_m} \left\| S(t_m-\sigma)(I-\mathcal{P}_N)Q^{\frac12} \right\|_{\mathrm{HS}}^2 \mathrm{d}\sigma \right)^{\frac12} \right\| \\ &\le \left\| A^{\frac{\beta-\frac1\rho}{2}} Q^{\frac12} \right\|_{\mathrm{HS}} \|A^{-\frac{\gamma}{2}}(I-\mathcal{P}_N)\|_{\mathcal{L}} \left(\int_0^{t_m} \left\| A^{\frac{(\gamma-\beta)+\frac1\rho}{2}} S(t_m-\sigma) \right\|_{\mathrm{HS}}^2 \mathrm{d}\sigma \right)^{\frac12}, \end{aligned}$$

which, by (2.7), (3.6), and since $(\beta-\gamma)\rho\in(0,1)$, implies

$$\begin{aligned} I_4 &\le C\lambda_{N+1}^{-\frac{\gamma}{2}} \left\| A^{\frac{\beta-\frac1\rho}{2}} Q^{\frac12} \right\|_{\mathrm{HS}} \left(\int_0^{t_m} (t_m-\sigma)^{-1+(\beta-\gamma)\rho}\,\mathrm{d}\sigma \right)^{\frac12} \\ &\le C\lambda_{N+1}^{-\frac{\gamma}{2}}. \end{aligned} \tag{3.11}$$

2. Temporal error. Here we estimate I_3, i.e.,

$$I_3 = \left\| \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} S(t_m-\sigma)\mathcal{P}_N\big(F(u(\sigma)) - F\left(U_j^N\right)\big)\,\mathrm{d}\sigma \right\|_{L_2(\Omega;H)}.$$

We use the Taylor expansion

$$F(u(\sigma)) = F(u(t_j)) + F'(u(t_j))\big(u(\sigma)-u(t_j)\big) + R_{F,j}(\sigma),$$

where the remainder is

$$R_{F,j}(\sigma) = \int_0^1 F''\Big(u(t_j)+\gamma\big(u(\sigma)-u(t_j)\big)\Big)\big(u(\sigma)-u(t_j), u(\sigma)-u(t_j)\big)(1-\gamma)\,\mathrm{d}\gamma,$$

to get

$$\begin{aligned} I_3 &\le \left\| \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} S(t_m-\sigma)\mathcal{P}_N\big(F(u(t_j)) - F\big(U_j^N\big)\big)\,\mathrm{d}\sigma \right\|_{L_2(\Omega;H)} \\ &\quad + \left\| \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} S(t_m-\sigma)\mathcal{P}_N F'(u(t_j))\big(u(\sigma)-u(t_j)\big)\,\mathrm{d}\sigma \right\|_{L_2(\Omega;H)} \\ &\quad + \left\| \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} S(t_m-\sigma)\mathcal{P}_N R_{F,j}(\sigma)\,\mathrm{d}\sigma \right\|_{L_2(\Omega;H)}. \end{aligned}$$

By substituting $u(\sigma)$ and $u(t_j)$ from the variation of constants formula (2.10) in the second term, we have

$$I_3 \le \sum_{l=1}^{7} I_{3,l}, \tag{3.12}$$

where

$$I_{3,1} = \left\| \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} S(t_m-\sigma)\mathcal{P}_N\big(F(u(t_j)) - F\left(U_j^N\right)\big)\,\mathrm{d}\sigma \right\|_{L_2(\Omega;H)},$$

$$I_{3,2} = \left\| \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} S(t_m-\sigma)\mathcal{P}_N F'(u(t_j))\big(S(\sigma)-S(t_j)\big)u_0\,\mathrm{d}\sigma \right\|_{L_2(\Omega;H)},$$

$$I_{3,3} = \left\| \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} S(t_m-\sigma)\mathcal{P}_N F'(u(t_j)) \int_{t_j}^{\sigma} S(\sigma-\tau)F(u(\tau))\,\mathrm{d}\tau\,\mathrm{d}\sigma \right\|_{L_2(\Omega;H)},$$

$$\begin{aligned} I_{3,4} = \Bigg\| &\sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} S(t_m-\sigma)\mathcal{P}_N F'(u(t_j)) \\ &\times \int_0^{t_j} \big(S(\sigma-\tau)-S(t_j-\tau)\big)F(u(\tau))\,\mathrm{d}\tau\,\mathrm{d}\sigma \Bigg\|_{L_2(\Omega;H)}, \end{aligned}$$

$$I_{3,5} = \left\| \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} S(t_m-\sigma)\mathcal{P}_N F'(u(t_j)) \int_{t_j}^{\sigma} S(\sigma-\tau)\,\mathrm{d}W(\tau)\,\mathrm{d}\sigma \right\|_{L_2(\Omega;H)},$$

$$\begin{aligned} I_{3,6} = \Bigg\| &\sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} S(t_m-\sigma)\mathcal{P}_N F'(u(t_j)) \\ &\times \int_0^{t_j} \big(S(\sigma-\tau)-S(t_j-\tau)\big)\,\mathrm{d}W(\tau)\,\mathrm{d}\sigma \Bigg\|_{L_2(\Omega;H)}, \end{aligned}$$

and

$$I_{3,7} = \left\| \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} S(t_m-\sigma)\mathcal{P}_N R_{F,j}(\sigma)\,\mathrm{d}\sigma \right\|_{L_2(\Omega;H)}.$$

First, using (2.11) and (2.7) with $s=0$, we have

$$I_{3,1} \le ML\Delta t \sum_{j=0}^{m-1} \left\| u(t_j) - U_j^N \right\|_{L_2(\Omega;H)}. \tag{3.13}$$

To estimate $I_{3,2}$, we have

$$\begin{aligned} I_{3,2} &\le \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \left\| A^{\frac{\delta}{2}} S(t_m-\sigma) \right\|_{\mathcal{L}} \left\| A^{-\frac{\delta}{2}} F'(u(t_j))\big(S(\sigma)-S(t_j)\big)u_0 \right\|_{L_2(\Omega;H)} \mathrm{d}\sigma \\ &= \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \left\| A^{\frac{\delta}{2}} S(t_m-\sigma) \right\|_{\mathcal{L}} \left\| A^{-\frac{\delta}{2}} F'(u(t_j)) \int_{t_j}^{\sigma} \dot{S}(\tau)u_0\,\mathrm{d}\tau \right\|_{L_2(\Omega;H)} \mathrm{d}\sigma, \end{aligned}$$

so that, using (2.12), (2.7), and (2.14) with $p=4$, we obtain

$$\begin{aligned}
I_{3,2} &\le C \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} (t_m-\sigma)^{-\frac{\delta\rho}{2}} \big(1+\|u(t_j)\|_{L_4(\Omega;\dot H^\gamma)}\big) \\
&\quad \times \bigg\| \Big\| \int_{t_j}^{\sigma} \dot S(\tau) u_0 \,\mathrm{d}\tau \Big\|_{-\gamma} \bigg\|_{L_4(\Omega;\mathbb{R})} \mathrm{d}\sigma \\
&\le C \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} (t_m-\sigma)^{-\frac{\delta\rho}{2}} \bigg\| \int_{t_j}^{\sigma} \big\|A^{-\gamma} \dot S(\tau) A^{\frac{\gamma}{2}} u_0 \,\mathrm{d}\tau\big\| \bigg\|_{L_4(\Omega;\mathbb{R})} \mathrm{d}\sigma.
\end{aligned}$$

Now, by (2.9), we have

$$\begin{aligned}
I_{3,2} &\le C\|u_0\|_{L_4(\Omega;\dot H^\gamma)} \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} (t_m-\sigma)^{-\frac{\delta\rho}{2}} \int_{t_j}^{\sigma} \tau^{\gamma\rho-1} \,\mathrm{d}\tau \,\mathrm{d}\sigma \\
&\le \frac{C}{\gamma\rho} \big(t_{j+1}^{\gamma\rho} - t_j^{\gamma\rho}\big) \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} (t_m-\sigma)^{-\frac{\delta\rho}{2}} \,\mathrm{d}\sigma \le C\big(t_{j+1}^{\gamma\rho} - t_j^{\gamma\rho}\big),
\end{aligned}$$

and, since $\gamma\rho \in (0,1)$, we consequently have

$$I_{3,2} \le C\Delta t^{\gamma\rho}. \tag{3.14}$$

Now we estimate $I_{3,3}$ in (3.12). Using (2.11) and (2.7) with $s=0$, we have

$$\begin{aligned}
I_{3,3} &\le \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \|S(t_m-\sigma)\|_{\mathcal{L}} \bigg\| F'(u(t_j)) \int_{t_j}^{\sigma} S(\sigma-\tau) F(u(\tau)) \,\mathrm{d}\tau \bigg\|_{L_2(\Omega;H)} \mathrm{d}\sigma \\
&\le L \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \|S(t_m-\sigma)\|_{\mathcal{L}} \int_{t_j}^{\sigma} \|S(\sigma-\tau)\|_{\mathcal{L}} \|F(u(\tau))\|_{L_2(\Omega;H)} \,\mathrm{d}\tau \,\mathrm{d}\sigma \\
&\le C \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \int_{t_j}^{\sigma} \big(1+\|u(\tau)\|_{L_2(\Omega;H)}\big) \,\mathrm{d}\tau \,\mathrm{d}\sigma,
\end{aligned}$$

which, by (2.14) with $p=2, \gamma=0$, implies

$$I_{3,3} \le C\Delta t. \tag{3.15}$$

To estimate $I_{3,4}$ in (3.12), we have

$$\begin{aligned}
I_{3,4} &\le \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \Big\|A^{\frac{\delta}{2}} S(t_m-\sigma)\Big\|_{\mathcal{L}} \\
&\quad \times \bigg\| A^{-\frac{\delta}{2}} F'(u(t_j)) \int_0^{t_j} \big(S(\sigma-\tau) - S(t_j-\tau)\big) F(u(\tau)) \,\mathrm{d}\tau \bigg\|_{L_2(\Omega;H)} \mathrm{d}\sigma \\
&= \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \Big\|A^{\frac{\delta}{2}} S(t_m-\sigma)\Big\|_{\mathcal{L}} \\
&\quad \times \bigg\| \Big\| A^{-\frac{\delta}{2}} F'(u(t_j)) \int_0^{t_j} \int_{t_j}^{\sigma} \dot S(\theta-\tau) \,\mathrm{d}\theta \; F(u(\tau)) \,\mathrm{d}\tau \Big\| \bigg\|_{L_2(\Omega;\mathbb{R})} \mathrm{d}\sigma,
\end{aligned}$$

which, in view of (2.7), (2.12), and (2.14) with $p = 2$, implies

$$\begin{aligned} I_{3,4} &\leq C \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} (t_m - \sigma)^{-\frac{\delta\rho}{2}} \\ &\quad \times \left\| \int_0^{t_j} \int_{t_j}^{\sigma} A^{-\frac{\gamma}{2}} \dot{S}(\theta - \tau) \, \mathrm{d}\theta \; F(u(\tau)) \, \mathrm{d}\tau \right\|_{L_4(\Omega;\mathbb{R})} \mathrm{d}\sigma \\ &\leq C \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} (t_m - \sigma)^{-\frac{\delta\rho}{2}} \\ &\quad \times \int_0^{t_j} \int_{t_j}^{\sigma} \left\| A^{-\frac{\gamma}{2}} \dot{S}(\theta - \tau) \right\|_{\mathcal{L}} \mathrm{d}\theta \, \left\| F(u(\tau)) \right\|_{L_4(\Omega;H)} \mathrm{d}\tau \, \mathrm{d}\sigma. \end{aligned}$$

Now, by (2.9) and (2.11), we have

$$\begin{aligned} I_{3,4} &\leq C \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} (t_m - \sigma)^{-\frac{\delta\rho}{2}} \\ &\quad \times \int_0^{t_j} \int_{t_j}^{\sigma} (\theta - \tau)^{\frac{\gamma\rho}{2} - 1} \, \mathrm{d}\theta \, \left(1 + \|u(\tau)\|_{L_4(\Omega;H)}\right) \mathrm{d}\tau \, \mathrm{d}\sigma, \end{aligned}$$

which, together with (2.14) with $p = 4, \gamma = 0$, implies

$$I_{3,4} \leq C \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} (t_m - \sigma)^{-\frac{\delta\rho}{2}} \int_0^{t_j} \int_{t_j}^{\sigma} (\theta - \tau)^{\frac{\gamma\rho}{2} - 1} \, \mathrm{d}\theta \; \mathrm{d}\tau \, \mathrm{d}\sigma.$$

Then, computing the double integral as

$$\begin{aligned} \int_0^{t_j} \int_{t_j}^{\sigma} (\theta - \tau)^{\frac{\gamma\rho}{2} - 1} \, \mathrm{d}\theta \, \mathrm{d}\tau &= \int_{t_j}^{\sigma} \int_0^{t_j} (\theta - \tau)^{\frac{\gamma\rho}{2} - 1} \, \mathrm{d}\tau \, \mathrm{d}\theta \\ &= \frac{2}{\gamma\rho} \int_{t_j}^{\sigma} \left(\theta^{\frac{\gamma\rho}{2}} - (\theta - t_j)^{\frac{\gamma\rho}{2}} \right) \mathrm{d}\theta \leq \frac{2}{\gamma\rho} t_j^{\frac{\gamma\rho}{2}} \Delta t, \end{aligned}$$

due to $\frac{\gamma\rho}{2} \in (0, \frac{1}{2})$, we conclude the estimate

$$I_{3,4} \leq C\Delta t. \tag{3.16}$$

We now estimate the terms in (3.12), which are affected by the noise. For $I_{3,5}$, using the fact that the expected value of independent processes is zero, and then the Cauchy–Schwarz inequality, we have

$$\begin{aligned} I_{3,5}^2 &= \mathbb{E} \left\| \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} S(t_m - \sigma) \mathcal{P}_N F'(u(t_j)) \int_{t_j}^{\sigma} S(\sigma - \tau) \, \mathrm{d}W(\tau) \, \mathrm{d}\sigma \right\|^2 \\ &= \sum_{j=0}^{m-1} \mathbb{E} \left\| \int_{t_j}^{t_{j+1}} \int_{t_j}^{\sigma} S(t_m - \sigma) \mathcal{P}_N F'(u(t_j)) S(\sigma - \tau) \, \mathrm{d}W(\tau) \, \mathrm{d}\sigma \right\|^2 \\ &\leq \Delta t \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \mathbb{E} \left\| \int_{t_j}^{\sigma} S(t_m - \sigma) \mathcal{P}_N F'(u(t_j)) S(\sigma - \tau) \, \mathrm{d}W(\tau) \right\|^2 \mathrm{d}\sigma. \end{aligned}$$

Then, by the Itô isometry (2.2), (2.11) and (2.14) with $p=2,\gamma=0$, we have

$$I_{3,5}^2 \le \Delta t \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \int_{t_j}^{\sigma} \left\| S(t_m-\sigma)\mathcal{P}_N F'(u(t_j)) S(\sigma-\tau) Q^{\frac12} \right\|_{\mathrm{HS}}^2 \mathrm{d}\tau\,\mathrm{d}\sigma$$
$$\le C\Delta t \left\| A^{\frac{\beta-\frac1\rho}{2}} Q^{\frac12}\right\|_{\mathrm{HS}}^2 \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \int_{t_j}^{\sigma} \left\| S(t_m-\sigma)\right\|_{\mathcal{L}}^2 \left\| A^{\frac{-\beta+\frac1\rho}{2}} S(\sigma-\tau)\right\|_{\mathcal{L}}^2 \mathrm{d}\tau\,\mathrm{d}\sigma.$$

Now, using (2.4) and (2.7), we obtain

$$I_{3,5}^2 \le C\Delta t \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \int_{t_j}^{\sigma} (\sigma-\tau)^{\beta\rho-1}\,\mathrm{d}\tau\,\mathrm{d}\sigma \le C\Delta t^{1+\beta\rho},$$

and therefore, we conclude the estimate

$$I_{3,5} \le C\Delta t^{\frac{1+\beta\rho}{2}}. \tag{3.17}$$

Now we estimate $I_{3,6}$. To this end, having

$$I_{3,6} \le \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \left\| A^{\frac{\delta}{2}} S(t_m-\sigma)\right\|_{\mathcal{L}}$$
$$\times \left\| A^{-\frac{\delta}{2}} F'(u(t_j)) \int_0^{t_j} \big(S(\sigma-\tau)-S(t_j-\tau)\big)\,\mathrm{d}W(\tau)\right\|_{L_2(\Omega;H)} \mathrm{d}\sigma$$
$$= \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \left\| A^{\frac{\delta}{2}} S(t_m-\sigma)\right\|_{\mathcal{L}}$$
$$\times \left\|\left\| A^{-\frac{\delta}{2}} F'(u(t_j)) \int_0^{t_j} \int_{t_j}^{\sigma} \dot S(\theta-\tau)\,\mathrm{d}\theta\,\mathrm{d}W(\tau)\right\|\right\|_{L_2(\Omega;H)} \mathrm{d}\sigma,$$

and using (2.7) and (2.12), we obtain

$$I_{3,6} \le C\left(1+\sup_{t\in[0,T]} \|u(t)\|_{L_4(\Omega;\dot H^\gamma)}\right)$$
$$\times \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} (t_m-\sigma)^{-\frac{\delta\rho}{2}} \left\| \int_0^{t_j} \int_{t_j}^{\sigma} A^{-\frac{\gamma}{2}} \dot S(\theta-\tau)\,\mathrm{d}\theta\,\mathrm{d}W(\tau)\right\|_{L_4(\Omega;H)} \mathrm{d}\sigma.$$

Then, by (2.14) with $p=4$ and the Burkholder–Davis–Gundy inequality (2.3),

$$I_{3,6} \le C \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} (t_m-\sigma)^{-\frac{\delta\rho}{2}}$$
$$\times \left\| \left(\int_0^{t_j} \left\| \int_{t_j}^{\sigma} A^{-\frac{\gamma}{2}} \dot S(\theta-\tau)\,\mathrm{d}\theta\; Q^{\frac12}\right\|_{\mathrm{HS}}^2 \mathrm{d}\tau\right)^{\frac12} \right\|_{L_4(\Omega;\mathbb{R})} \mathrm{d}\sigma$$
$$\le C \left\| A^{\frac{\beta-\frac1\rho}{2}} Q^{\frac12}\right\|_{\mathrm{HS}} \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} (t_m-\sigma)^{-\frac{\delta\rho}{2}}$$
$$\times \left\| \left(\int_0^{t_j} \left(\int_{t_j}^{\sigma} \left\| A^{\frac{-(\gamma+\beta)+\frac1\rho}{2}} \dot S(\theta-\tau)\right\|_{\mathcal{L}} \mathrm{d}\theta \right)^2 \mathrm{d}\tau \right)^{\frac12} \right\|,$$

which, using (2.4) and (2.9), implies

$$I_{3,6} \leq C \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} (t_m - \sigma)^{-\frac{\delta\rho}{2}} \left\| \left(\int_0^{t_j} \left(\int_{t_j}^{\sigma} (\theta - \tau)^{\frac{(\gamma+\beta)\rho}{2} - \frac{3}{2}} \, \mathrm{d}\theta \right)^2 \mathrm{d}\tau \right)^{\frac{1}{2}} \right\|.$$

From this and

$$\begin{aligned}
&\int_0^{t_j} \left(\int_{t_j}^{\sigma} (\theta - \tau)^{\frac{(\gamma+\beta)\rho}{2} - \frac{3}{2}} \, \mathrm{d}\theta \right)^2 \mathrm{d}\tau \\
&\quad = C \int_0^{t_j} \left((\sigma - \tau)^{\gamma\rho - \frac{1}{2} + \frac{(\beta-\gamma)\rho}{2}} - (t_j - \tau)^{\gamma\rho - \frac{1}{2} + \frac{(\beta-\gamma)\rho}{2}} \right)^2 \mathrm{d}\tau \\
&\quad = C \int_0^{t_j} \left((\sigma - \tau)^{\gamma\rho} (\sigma - \tau)^{-\frac{1}{2} + \frac{(\beta-\gamma)\rho}{2}} - (t_j - \tau)^{\gamma\rho} (t_j - \tau)^{-\frac{1}{2} + \frac{(\beta-\gamma)\rho}{2}} \right)^2 \mathrm{d}\tau \\
&\quad \leq C \int_0^{t_j} \left((\sigma - \tau)^{\gamma\rho} (t_j - \tau)^{-\frac{1}{2} + \frac{(\beta-\gamma)\rho}{2}} - (t_j - \tau)^{\gamma\rho} (t_j - \tau)^{-\frac{1}{2} + \frac{(\beta-\gamma)\rho}{2}} \right)^2 \mathrm{d}\tau \\
&\quad = C \int_0^{t_j} (t_j - \tau)^{-1 + (\beta-\gamma)\rho} \left((\sigma - \tau)^{\gamma\rho} - (t_j - \tau)^{\gamma\rho} \right)^2 \mathrm{d}\tau \\
&\quad \leq C \Delta t^{2\gamma\rho} \int_0^{t_j} (t_j - \tau)^{-1 + (\beta-\gamma)\rho} \, \mathrm{d}\tau \\
&\quad = C t_j^{(\beta-\gamma)\rho} \Delta t^{2\gamma\rho},
\end{aligned}$$

we conclude the estimate

$$I_{3,6} \leq C \Delta t^{\gamma\rho}. \tag{3.18}$$

To estimate $I_{3,7}$, the last term in (3.12), we have

$$I_{3,7} \leq \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} \left\| A^{\frac{\eta}{2}} S(t_m - \sigma) \right\|_{\mathcal{L}} \left\| \left\| A^{-\frac{\eta}{2}} R_{F,j}(\sigma) \right\| \right\|_{L_2(\Omega;H)} \mathrm{d}\sigma.$$

By (2.7) and (2.13), this implies

$$I_{3,7} \leq C \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} (t_m - \sigma)^{-\frac{\eta\rho}{2}} \| u(\sigma) - u(t_j) \|^2_{L_4(\Omega;H)} \, \mathrm{d}\sigma,$$

which, considering the fact that [1, Proposition 3.2]

$$\| u(\sigma) - u(t_j) \|_{L_4(\Omega;H)} \leq C (\sigma - t_j)^{\frac{\gamma\rho}{2}},$$

we conclude the estimate

$$I_{3,7} \leq C \Delta t^{\gamma\rho}. \tag{3.19}$$

Finally, inserting (3.9)–(3.11) and (3.13)–(3.19) into (3.12), we have

$$\left\| u(t_m) - U_m^N \right\|_{L_2(\Omega;H)} \leq C \left(\Delta t^{\gamma\rho} + \lambda_{N+1}^{-\frac{\gamma}{2}} \right) + C \Delta t \sum_{j=0}^{m-1} \| u(t_j) - U_j \|_{L_2(\Omega;H)},$$

which, by the discrete Gronwall lemma, completes the proof. ∎

Remark 3. We note that the temporal strong rate of convergence is (almost) twice the rate of the backward Euler method combined with the first order Lubich convolution quadrature used in [1, 11]. In particular, when Q is of trace class we almost recover the deterministic order $O(\Delta t)$ in time (cf. Remark 4).

Remark 4. For the deterministic form of the model problem (1.1), i.e., with $W = 0$, the rate is therefore $O\big(\Delta t + \lambda_{N+1}^{-\gamma}\big)$, as expected. Indeed, recalling (3.9) and Remark 2, we have

$$I_1 \le \|S(t_m)\|_{\mathcal{L}} \|A^{-\gamma}(I - \mathcal{P}_N)A^{\gamma}u_0\| \le C\lambda_{N+1}^{-\gamma}\|u_0\|_{2\gamma}.$$

We also recall (3.10), for which we have, in this case,

$$I_2 \le C\lambda_{N+1}^{-\gamma}(1 + \|u_0\|).$$

Remark 5. To avoid the restrictive assumption that one has access to the eigenvalues and eigenfunctions of A, in theory, one may discretize (1.1) in space by other means, such as the finite element method. Indeed, when A is minus the Dirichlet Laplacian in $L_2(\mathcal{D})$, then one has nonsmooth data error estimates for the finite element method, at least for the main example (1.3) (see [15]), and the error analysis in the present paper can be performed with a slight increase in technicality using these nonsmooth data estimates. However, the corresponding algorithm would be difficult to implement in practice. Indeed, if S_h denotes the finite element approximation of S, where h is the finite element mesh size, one would have to simulate a Gaussian random variable with covariance operator

$$\int_0^{t_n} S_h(t)P_hQP_hS_h(t)\,\mathrm{d}t$$

which, even in the simplest case $Q = I$, is not practically feasible unless one has access to the eigenvalues and eigenfunctions of the discrete Laplacian A_h. Therefore, we have chosen to analyze the spectral Galerkin method for which the scheme is easily implementable, but even in this case, this is true only when A and Q commute.

Remark 6. The feasibility of the proposed numerical scheme relies heavily on whether one knows the scalar functions s_k from (2.6). This is the case for the Riesz kernel (see section 4) or for the tempered Riesz kernel, but in general this leads to the additional difficulty of solving (2.6), for example, by numerically inverting a Laplace transform.

4. Numerical implementation.

In this section we present the explicit form of the solution of (2.6) in terms of the Mittag–Leffler functions. Then, we illustrate the temporal strong order of convergence, to confirm the proposed rate in Theorem 1.

4.1. Explicit representation of the solution.

First, we derive an explicit representation of the resolvent family in terms of the Mittag–Leffler functions when b is the Riesz kernel.

Recall that the one parameter Mittag–Leffler function $E_a(z)$, $a > 0$, is defined as

$$E_a(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(ak+1)}, \quad z \in \mathbb{C},$$

where obviously $E_1(z) = \exp(z)$. By taking the Laplace transform of (2.6), when $b(t) = t^{\rho-2}/\Gamma(\rho-1)$, we have $\widehat{b}(z) = z^{1-\rho}$ and

$$\widehat{s_k}(z) = \frac{1}{z + \lambda_k z^{1-\rho}} = \frac{z^{\rho-1}}{z^{\rho} + \lambda_k},$$

which implies

$$s_k(t) = E_\rho(-\lambda_k t^\rho).$$

Thus, the resolvent family is given by

$$S(t)v = \sum_{k=1}^{\infty} E_\rho(-\lambda_k t^\rho)(v, \varphi_k)\varphi_k.$$

To explain the computer implementation of the fully discrete method (3.7), we note that

$$S_N(t_m) = \sum_{k=1}^{N} E_\rho(-t_m^\rho \lambda_k)(v, \varphi_k)\varphi_k.$$

Suppose that Q has the same eigenfunctions as A, so that $Qv = \sum_{k=1}^{\infty} \mu_k(v, \varphi_k)\varphi_k$. Then, for each time step $m = 1, \dots, M$, the approximation U_m^N defined by (3.7) is given by $U_m^N = \sum_{k=1}^{N} U_{m,k}^N \varphi_k$, where for $k = 1, \dots, N$,

$$\begin{aligned} U_{m,k}^N = E_\rho\left(-\lambda_k t_m^\rho\right) u_{0,k} &+ \sum_{j=0}^{m-1} \int_{t_j}^{t_{j+1}} E_\rho(-\lambda_k (t_m - \sigma)^\rho)\, \mathrm{d}\sigma \; F_k\left(U_j^N\right) \\ &+ \int_0^{t_m} E_\rho(-\lambda_k (t_m - \sigma)^\rho) \mu_k^{\frac{1}{2}} \, \mathrm{d}\beta_k(\sigma) \end{aligned} \tag{4.1}$$

and where $u_{0,k} = (u_0, \varphi_k)$, $F_k(\cdot) = (F(\cdot), \varphi_k)$, and β_k, $k = 1, \dots, N$, are mutually independent standard Brownian motions.

We note that the integrals of the Mittag–Leffler functions are computable, e.g., by means of a simple quadrature, say the trapezoidal method. For evaluating the Mittag–Leffler function we use `mlf.m` from [19]. What one has to be careful with is how to simulate, for fixed k, the random variables

$$\mathcal{O}(t_m) := \int_0^{t_m} E_\rho(-\lambda_k (t_m - \sigma)^\rho) \mu_k^{\frac{1}{2}} \, \mathrm{d}\beta_k(\sigma), \quad m = 1, \dots, M.$$

Observe that the $\mathbb{R}^M$-valued random variable

$$\mathcal{N} := (\mathcal{O}(t_1), \dots, \mathcal{O}(t_M))$$

is a 0-mean Gaussian random variable with covariance matrix

$$(M)_{i,j} = \int_0^{\min(t_i, t_j)} E_\rho(-\lambda_k (t_i - \sigma)^\rho) E_\rho(-\lambda_k (t_j - \sigma)^\rho) \mu_k \, \mathrm{d}\sigma.$$

Thus $\mathcal{N} = L\xi$, where $LL^T = M$, and ξ is an $\mathbb{R}^M$-valued random variable with independent standard Gaussian coordinates. This difficulty does not arise in the memoryless case as there one can exploit the semigroup property of the solution operator. In that case one only has to simulate the independent Gaussian random variables $\xi_i := \int_{t_{i-1}}^{t_i} \exp(-\lambda_k (t_i - \sigma)) \mu_k^{\frac{1}{2}} \, \mathrm{d}\beta_k(\sigma)$, $i = 1, \dots, M$, and then take $\mathcal{O}(t_m) = \sum_{i=1}^{m} \exp(-\lambda_k (t_m - t_i)) \xi_i$, $m = 1, \dots, M$.

4.2. Numerical experiments. Since the main contribution of this paper is the temporal approximation, we only present a simplified numerical experiment with uncoupled eigenmodes. More precisely, let $u = \sum_{k=1}^{\infty} u_k \varphi_k$ and define the nonlinear operator

$$F(u) = \sum_{k=1}^{\infty} f(u_k)\varphi_k. \tag{4.2}$$

We simulate various coordinates of the numerical approximation; that is, we simulate the random variables $U_{M,k}^N$ in (4.1) for various values of k, with $\mu_k = 1$ and $\lambda_k = k^2\pi^2$ being the eigenvalues of the Dirichlet Laplacian in one dimension on $\mathcal{D} = [0,1]$.

Since we are simulating scalar problems, the noise is trace class and we expect the rate of convergence of the MLEI to be almost 1 according to the theory. We also compare the performance of the MLEI to the backward Euler based convolution quadrature (BE) proposed and analyzed in [11, 12] in the linear case and in [1] in the semilinear setting. For BE the theory in [1] predicts a strong rate of almost 0.5 for trace class noise, although the conditions on f there are somewhat different to the present setting and hence the rate could be higher for smooth additive noise. In fact, in the scalar case we have $\|A^t Q^{1/2}\|_{\mathrm{HS}} = \lambda_k^t < \infty$ for any t so that we have smoothness of any order. But when λ_k is large, this quantity is also large, making the error constant large, and then we do not expect to see a higher rate than corresponding to $t = 0$; that is, the trace class noise case (this is seen in Figure 1). This also explains why, for smaller λ_k, we might occasionally observe a better rate than 0.5 for BE (see Figure 2 with λ_2). Nevertheless, in all experiments the MLEI outperforms BE by far and we experimentally see rate 1 for MLEI in all experiments.

We use 100 sample paths in all experiments. The computed solution is compared to a reference solution with much smaller mesh size. We use the functions $f(u) = \sin(u)$ and $f(u) = 5(1-u)/(1+u^2)$ and different values $\rho = 1.2$ and $\rho = 1.75$ (and for Figure 3 also $\rho = 1.5$); the first being closer to the heat equation with solutions dying out quickly (Figure 4), while the latter produces a more pronounced wave phenomenon (Figure 5). For functions with larger Lipschitz constant, such as $f(u) = 100(1-u)/(1+u^2)$, we still get similar convergence behavior as shown in Figure 3.

The figure legends and captions explain the settings for the various experiments.

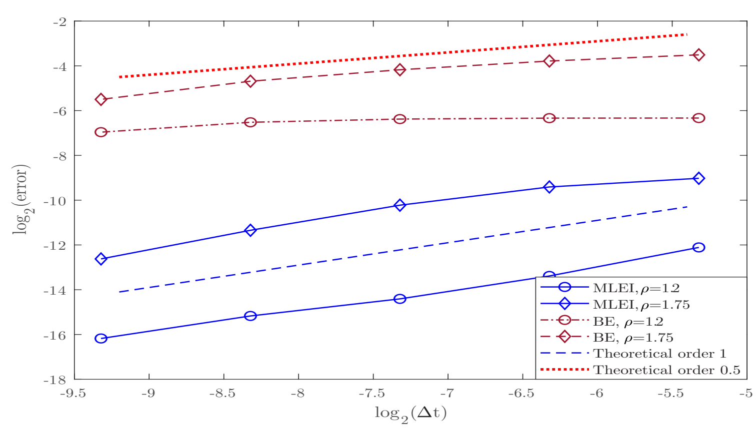

FIG. 1. *Comparison of BE and MLEI temporal rate of convergence with* $f(u) = \sin(u)$, $\rho = 1.2$ *and* $\rho = 1.75$, *and* $\lambda_{30} = 900\pi^2$.

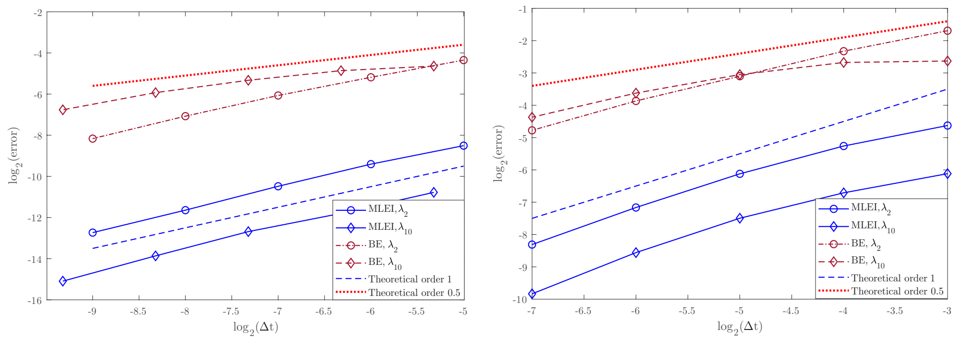

FIG. 2. *Comparison of BE and MLEI temporal rate of convergence with* $f(u) = \sin(u)$ *and* $\lambda_2 = 4\pi^2$*, respectively,* $\lambda_{10} = 100\pi^2$*. Left: with* $\rho = 1.2$*. Right:* $\rho = 1.75$*.*

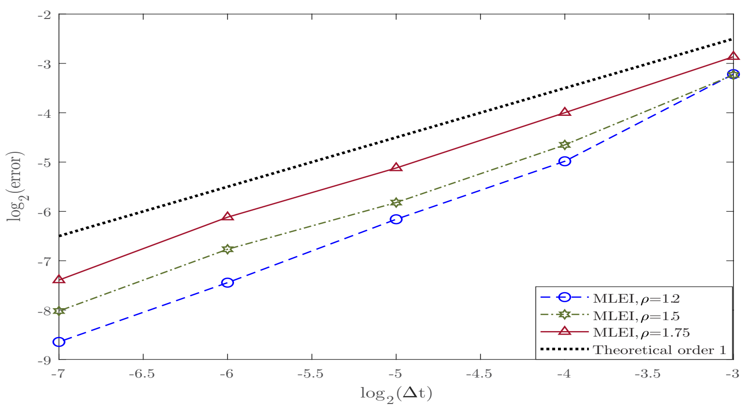

FIG. 3. *Temporal rate of convergence for MLEI with* $f(u) = 5(1-u)/(1+u^2)$*,* $\rho = 1.2, 1.5,$ *and* 1.75*, and* $\lambda_2 = 4\pi^2$*.*

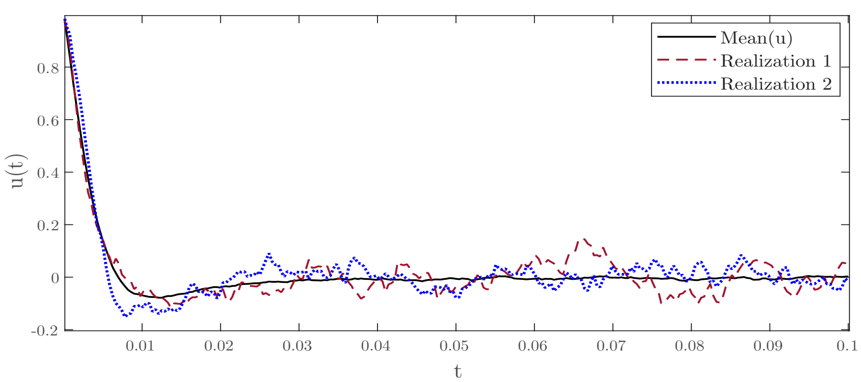

FIG. 4. *Solution behavior with* $f(u) = \sin(u)$*,* $\rho = 1.2$*, and* $\lambda_{10} = 100\pi^2$*.*

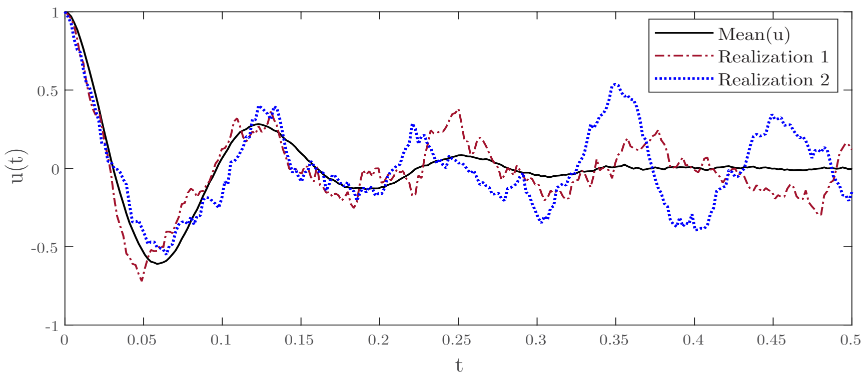

FIG. 5. *Solution behavior with* $f(u) = \sin(u)$, $\rho = 1.75$, *and* $\lambda_{10} = 100\pi^2$.

Acknowledgment. We would like to thank the anonymous referees for constructive comments that helped us improve the manuscript.

REFERENCES

[1] A. Andersson, M. Kovács, and S. Larsson, *Weak error analysis for semilinear stochastic Volterra equations with additive noise*, J. Math. Anal. Appl., 437 (2016), pp. 1283–1304.

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

[3] B. Baeumer, M. Geissert, and M. Kovács, *Existence, uniqueness and regularity for a class of semilinear stochastic Volterra equations with multiplicative noise*, J. Differential Equations, 258 (2015), pp. 535–554.

[4] S. Becker, A. Jentzen, and P. Kloeden, *An exponential Wagner-Platen type scheme for SPDEs*, SIAM J. Numer. Anal., 54 (2016), pp. 2389–2426.

[5] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Encyclopedia Math. Appl. 44, Cambridge University Press, Cambridge, UK, 1992.

[6] M. Gunzburger, B. Li, and J. Wang, *Sharp convergence rates of time discretization for stochastic time-fractional PDEs subject to additive space-time white noise*, Math. Comp., 88 (2019), pp. 1715–1741.

[7] M. Gunzburger, B. Li, and J. Wang, *Convergence of finite element solutions of stochastic partial integro-differential equations driven by white noise*, Numer. Math., 141 (2019), pp. 1043–1077.

[8] M. Hochbruck and A. Ostermann, *Exponential integrators*, Acta Numer., 19 (2010), pp. 209–286.

[9] A. Jentzen, P. Kloeden, and G. Winkel, *Efficient simulation of nonlinear parabolic SPDEs with additive noise*, Ann. Appl. Probab., 21 (2011), pp. 908–950.

[10] M. Kovács, S. Larsson, and F. Lindgren, *Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise* II*: Fully discrete schemes*, BIT, 53 (2013), pp. 497–525.

[11] M. Kovács and J. Printems, *Strong order convergence of a fully discrete approximation of a linear stochastic Volterra equation*, Math. Comp., 83 (2014), pp. 2325–2346.

[12] M. Kovács and J. Printems, *Weak convergence of a fully discrete approximation of a linear stochastic evolution equation with a positive-type memory term*, J. Math. Anal. Appl., 413 (2014), pp. 939–952.

[13] S. Larsson and M. Molteni, *Numerical solution of parabolic problems based on a weak space-time formulation*, Comput. Methods Appl. Math., 17 (2017), pp. 65–84.

[14] G. J. Lord and A. Tambue, *Stochastic exponential integrators for the finite element discretization of SPDEs for multiplicative and additive noise*, IMA J. Numer. Anal., 33 (2013), pp. 515–543.

[15] Ch. Lubich, I. Sloan, and V. Thomée, *Nonsmooth data error estimates for approximations of an evolution equation with a positive-type memory term*, Math. Comp., 65 (1996), pp. 1–17.

[16] F. Mainardi and P. Paradisi, *Fractional diffusive waves*, J. Comput. Acoust., 9 (2001), pp. 1417–1436.
[17] W. McLean and V. Thomée, *Numerical solution of an evolution equation with positive memory term*, J. Austral. Math. Soc. Ser. B, 35 (1993), pp. 23–70.
[18] I. Podlubny, *Fractional Differential Equations*, Academic Press, San Diego, 1999.
[19] I. Podlubny and M. Kacenak, *MATLAB MLF code*, File 8738, MATLAB Central File Exchange, 2009.
[20] J. Prüss, *Evolutionary Integral Equations and Applications*, Monogr. Math. 87, Birkhäuser, Basel, 1993.
[21] X. Wang, *Strong convergence rates of the linear implicit Euler method for the finite element discretization of SPDEs with additive noise*, IMA J. Numer. Anal., 37 (2017), pp. 956–984.
[22] X. Wang, S. Gan, and J. Tang, *Higher order strong approximations of semilinear stochastic wave equation with additive space-time white noise*, SIAM J. Sci. Comput., 36 (2014), pp. A2611–A2632.
[23] X. Wang and R. Qi, *A note on an accelerated exponential Euler method for parabolic SPDEs with additive noise*, Appl. Math. Lett., 46 (2015), pp. 31–37.