ON PARAREAL ALGORITHMS FOR SEMILINEAR PARABOLIC STOCHASTIC PDEs*

CHARLES-EDOUARD BREHIER† AND XU WANG‡

Abstract. Parareal algorithms are studied for semilinear parabolic stochastic partial differential equations to achieve a "parallel-in-real-time" implementation. These algorithms proceed as two-level integrators, with the fine integrator being given by the exponential Euler scheme in this work. Two choices for the coarse integrator are considered: the linear implicit Euler scheme and the exponential Euler scheme. It is proved that as the number of iterations increases, the order of convergence is limited by the regularity of the noise, whereas for the exponential Euler case, the order of convergence always increases. The influences on the performance of the parareal algorithms, of the choice of the coarse integrator, of the regularity of the noise, and of the number of parareal iterations are also illustrated by extensive numerical experiments.

Key words. parareal algorithm, parabolic stochastic partial differential equations, strong convergence, implicit Euler scheme, exponential Euler scheme

AMS subject classifications. 60H35, 65M12, 65W05

DOI. 10.1137/19M1251011

1. Introduction. In the last two decades, numerical methods for stochastic partial differential equations (SPDEs) have been extensively studied; see, for instance, the monographs [13], [15], and [18] and references therein. The rate of convergence of the schemes used for temporal and spatial discretization is related to the regularity of the noise, which may be arbitrarily low, and in such a situation, effective numerical methods are difficult to construct. For instance, for a one-dimensional semilinear parabolic SPDE, driven by Gaussian space-time white noise, trajectories are only α-Hölder continuous in time and 2α-Hölder continuous in space for $\alpha < \frac{1}{4}$, and standard Euler and finite difference schemes thus have a (strong) rate of convergence equal to α and 2α, respectively.

In this article, we will only focus on the temporal discretization and consider a semilinear parabolic SPDE of the type

$$\frac{\partial u(t,x)}{\partial t} = \frac{\partial^2 u(t,x)}{\partial x^2} + f(u(t,x)) + \dot{W}^Q(t,x), \quad (t,x) \in \mathbb{R}_+ \times (0,1),$$
$$u(0,x) = u_0(x), \quad x \in (0,1), \quad u(t,0) = u(t,1) = 0,$$

that is, a one-dimensional semilinear heat equation with homogeneous Dirichlet boundary conditions and with additive Gaussian noise, which is white in time and colored in space. A rigorous interpretation as a stochastic evolution equation (in the sense of [5]), driven by a Wiener process, with values in an infinite-dimensional Hilbert space is given by

$$du(t) = Au(t)dt + F(u(t))dt + dW^Q(t)$$

†Univ Lyon, CNRS, Université Claude Bernard Lyon 1, UMR5208, Institut Camille Jordan, F-69622 Villeurbanne, France (brehier@math.univ-lyon1.fr).
‡Department of Mathematics, Purdue University, West Lafayette, IN 47907 (wang4191@purdue.edu).

(see also section 2), where precise assumptions for the linear operator A, the nonlinear operator F, and the covariance operator Q are stated. It is well known that the rate of convergence (studied only in the strong sense in this article) of the error depends on the properties of the covariance operator Q of the Wiener process and can be arbitrarily small. In particular, in the case of space-time white noise (Q is the identity), the order of convergence of Euler schemes is essentially $\frac{1}{4}$. If the noise is of trace-class (Q has finite trace), the order of convergence is essentially $\frac{1}{2}$.

The objective of this article is to study the applicability of so-called parareal algorithms to improve the performance of temporal discretization schemes. Implementing parallel algorithms to solve time-dependent evolution equations is not natural, and parallel-in-time integration methods have been extensively studied, based on multigrid or multiple shooting techniques. We refer to the pioneering contributions of [3, 4, 12, 14, 19, 22, 23] and the review in [6]. The parareal algorithm has been introduced in [16]; see also [21]. It is a *parallel-in-real-time* technique, based on the use of two integrators with two different time-step sizes: a coarse integrator, denoted by $\mathcal{G}$, with coarse time-step size ΔT and a fine integrator, denoted by $\mathcal{F}$, using J steps with time-step size δt such that $\Delta T = J\delta t$. In practice, $\mathcal{G}$ may be less accurate but cheaper than $\mathcal{F}$. The parareal algorithm is an iterative method, using a predictor-corrector strategy, in which computations of the fine integrator at each iteration are performed in parallel: The recursion is given by

$$u_{n+1}^{(k+1)} = \mathcal{G}(u_n^{(k+1)}, t_n, t_{n+1}) + \mathcal{F}(u_n^{(k)}, t_n, t_{n+1}) - \mathcal{G}(u_n^{(k)}, t_n, t_{n+1}),$$
$$u_0^{(k)} = u_0,$$

where k represents the index for parareal iterations. For $k = 0$, the scheme is initialized using the coarse integrator. We refer to (5) and to section 3 below for a detailed presentation of the algorithm.

Since the pioneering work of [16], where the parareal algorithm has been introduced for a class of ordinary differential equations, several extensions have been considered; see, for instance, [1, 7, 8, 9, 10, 20]. In particular, [1] deals with parabolic PDEs and studies the stability and convergence properties, which may require regularity properties, depending on the choice of integrators. The application of the parareal algorithm for stochastic systems has been considered first in [2] and more recently in [11] for stochastic Schrödinger PDEs and in [24] for a class of stochastic differential equations. More precisely, in [11], parareal algorithms for the stochastic Schrödinger equation with damping are studied with $\mathcal{F}$ being the exact solver and $\mathcal{G}$ being the exponential-θ scheme. The long-term convergence is obtained for the case $\theta > \frac{1}{2}$ or sufficiently large α, which ensures sufficient exponential decay of the coarse integrator.

Let us now describe the contributions of this article. The parareal algorithm is applied to the SPDE above. The exponential Euler scheme is chosen as the fine integrator. The main contribution of this article is to reveal that the parareal algorithm behaves differently, depending on the choice of the coarse integrator, when applied to semilinear parabolic SPDEs. Such results, based on both theoretical analysis and numerical experiments, have not been reported before to the best of our knowledge. The error of the parareal algorithm, considered in this article, is the distance between the solution computed by the proposed algorithm and a reference solution generated by the fine integrator (which is computationally expensive and is not computed in practice). The parareal algorithm is useful to reduce computational cost only if the order of convergence of this error (with respect to the coarse time-step size ΔT) is strictly larger for some $k \geq 1$ than for $k = 0$.

To state the main results of this article (see Theorems 4.1, 5.1, and 5.2 below for precise statements), let us assume that, for some $\alpha > 0$, the covariance operator Q satisfies a condition of the type

$$\|(-A)^{\alpha-\frac{1}{2}}Q^{\frac{1}{2}}\|_{\mathcal{L}_2(H)} < \infty,$$

where $\|\cdot\|_{\mathcal{L}_2(H)}$ denotes the Hilbert–Schmidt norm for operators from H to H.

First, assume that the linear implicit Euler scheme is chosen as the coarse integrator; i.e., $e^{\Delta TA}$ is approximated by $(I - \Delta TA)^{-1}$. In addition, assume that $F = 0$. Then the order of convergence of the parareal algorithm, with k iterations, is essentially $\min(\alpha, k+1)$ and saturates at α when k increases; see Theorem 4.1 and numerical experiments in section 4.2. In particular, if $\alpha \leq 1$, which includes the space-time white noise case with essentially $\alpha = \frac{1}{4}$, the application of the parareal algorithm is useless since this order of convergence does not depend on k. The way the error behaves in terms of k thus depends a lot on the regularity of the noise.

Second, assume that the exponential Euler scheme is chosen as the coarse integrator. Then it is proved that the order of convergence of the parareal algorithm, with k iterations, is at least essentially of size $(k+1)\alpha$ if $\alpha \in (0, \frac{1}{2}]$; see Theorem 5.1. Contrary to the first case, this order is linear in k; thus, applying the parareal algorithm always reduces the computational cost whatever the regularity of the noise. Numerical experiments (see section 5.2) reveal that the order of convergence is sharp for $k = 0$ (the parareal algorithm is not applied) and $k = 1$ (one iteration is applied) but is larger when $k \geq 2$. In fact, the choice $k = 1$ is optimal when considering the final computational cost. Theorem 5.2 proves that for $\alpha \in (0, \frac{1}{4}]$ and $k \geq 2$, the order of convergence is at least of size $2k\alpha$, which is indeed larger than $(k+1)\alpha$.

In conclusion, the parareal algorithm may offer an effective strategy to reduce computational cost for the simulation of trajectories of SPDEs. Several questions remain open: For instance, generalizations in higher dimension, algorithms for equations with multiplicative noise, or using other integrators are left for future works.

This article is organized as follows. Precise assumptions on the operators A and F and on the covariance operator Q are provided in section 2. Section 3 is devoted to introducing the parareal algorithm, presenting the possible choices of coarse integrators, and defining the error. The study of the behavior of the parareal algorithm when the linear implicit Euler scheme is chosen as the coarse integrator is provided in section 4: More precisely, see Theorem 4.1 for the theoretical error estimates and section 4.2 for the numerical experiments. The study of the behavior of the parareal algorithm when the exponential Euler scheme is chosen as the coarse integrator is provided in section 5: More precisely, see Theorems 5.1 and 5.2 for the theoretical error estimates and section 5.2 for the numerical experiments.

2. Setting. Let $H = L^2(0,1)$, with norm and inner product denoted by $|\cdot|$ and $\langle\cdot,\cdot\rangle$, respectively. In this work, stochastic evolution equations, with additive noise, of the following type are considered (see [5, 17]): Given an initial condition $u(0) = u_0 \in H$,

$$(1) \qquad\qquad du(t) = Au(t)dt + F(u(t))dt + dW^Q(t),$$

where the solution $(u(t))_{t\geq 0}$ is a stochastic process with values in H. The operator A is a linear parabolic differential operator with homogeneous Dirichlet boundary conditions (section 2.1); the operator F is a globally Lipschitz, nonlinear operator (section 2.2); and $(W^Q(t))_{t\geq 0}$ is a Q-Wiener process. Appropriate assumptions to justify the global well-posedness of (1) are introduced below.

In the following, the initial condition u_0 is assumed to be deterministic; however, the extension to a random initial condition u_0 (independent of the Wiener process and satisfying appropriate moment conditions) is straightforward by a conditioning argument.

2.1. Linear operator A. The linear operator A is defined as the unbounded linear operator on $H = L^2(0, 1)$ such that

$$\begin{cases} Dom(A) = \mathbb{H}^2(0,1) \cap \mathbb{H}_0^1(0,1), \\ Au = u'', \ u \in Dom(A). \end{cases}$$

In other words, A is the Laplace operator with homogeneous Dirichlet boundary conditions. Recall that A is an unbounded, self-adjoint, linear operator and that $Ae_p = -\lambda_p e_p$ for all $p \in \mathbb{N}$, where the eigenvalues are given by $\lambda_p = (\pi p)^2$ and the eigenfunctions $e_p = \sqrt{2}\sin(p\pi\cdot)$ form a complete orthonormal system of H.

The linear operator A generates an analytic and strongly continuous semigroup on H, denoted by $(e^{tA})_{t \geq 0}$. Note that for all $u \in H$, one has

$$e^{tA}u = \sum_{p \in \mathbb{N}} e^{-\lambda_p t}\langle u, e_p\rangle e_p.$$

For any $\alpha \in [0, 1]$ and $u \in H$, let

$$|u|_\alpha^2 = \sum_{p \in \mathbb{N}} \lambda_p^{2\alpha}\langle u, e_p\rangle^2 \in [0, \infty].$$

For $u \in Dom((-A)^\alpha) = \{u \in H : |u|_\alpha < \infty\}$, set $(-A)^\alpha u = \sum_{p \in \mathbb{N}} \lambda_p^\alpha \langle u, e_p\rangle e_p \in H$, and note that $|u|_\alpha = |(-A)^\alpha u|$. In addition, for any $\alpha \in [0, 1]$ and $u \in H$, let $(-A)^{-\alpha}u = \sum_{p \in \mathbb{N}} \lambda_p^{-\alpha}\langle u, e_p\rangle e_p \in H$ and $|u|_{-\alpha}^2 = \sum_{p \in \mathbb{N}} \lambda_p^{-2\alpha}\langle u, e_p\rangle^2$.

Regularization properties of the semigroup $(e^{tA})_{t \geq 0}$ are stated in Proposition 2.1 below. The following notation is used. First, $\mathcal{L}(H)$ is the space of linear bounded operators from H to H, with the operator norm denoted by $\|\cdot\|_{\mathcal{L}(H)}$. Second, $\mathcal{L}_2(H)$ is the space of Hilbert–Schmidt operators from H to H, with the Hilbert–Schmidt norm denoted by $\|\cdot\|_{\mathcal{L}_2(H)}$.

PROPOSITION 2.1. *For all $t \geq 0$, $\|e^{tA}\|_{\mathcal{L}(H)} \leq e^{-\lambda_1 t}$. Moreover, for all $\alpha \in [0,1]$, there exists $C_\alpha \in (0, \infty)$ such that, for all $t \in (0, \infty)$,*

$$\|(-A)^\alpha e^{tA}\|_{\mathcal{L}(H)} \leq C_\alpha \min(t, 1)^{-\alpha}, \ \|(-A)^{-\alpha}(e^{tA} - I)\|_{\mathcal{L}(H)} \leq C_\alpha \min(t, 1)^\alpha.$$

2.2. Nonlinear operator F. The analysis of the rate of convergence for the parareal algorithm below proceeds in a simplified, abstract framework, whereas numerical experiments are performed in the more natural framework of Nemytskii operators. The abstract framework does not encompass this case. Indeed, the treatment of Nemytskii would require the introduction of further concepts (such as γ-Radonifying operators in order to work in Banach spaces $L^p(0,1)$). Instead of increasing the technical level of the presentation, the choice made for this article is to study the main features of the parareal algorithms applied to SPDEs in a more pedagogical way because of the simplified framework.

Let us first state the assumptions on the nonlinear operator F which are employed for the theoretical analysis.

Assumption 1. The nonlinear operator $F : H \to H$ is globally Lipschitz continuous and is twice Fréchet differentiable, with bounded first- and second-order deriv-

atives. Moreover, for any $\alpha \in (0, \frac{1}{2})$ and any arbitrarily small $\kappa \in (0, \frac{1}{2} - \alpha)$, there exists $C_{F,\alpha,\kappa} \in (0, \infty)$ such that, for all $u \in Dom((-A)^{\alpha+\kappa})$ and $h \in H$,

$$|DF(u).h|_{-\alpha} \leq C_{F,\alpha,\kappa}\big(1 + |u|_{\alpha+\kappa}\big)|h|_{-\alpha}$$

and for all $u, h \in Dom((-A)^{\alpha})$,

$$|DF(u).h|_{\alpha-\kappa} \leq C_{F,\alpha,\kappa}\big(1 + |u|_{\alpha}\big)|h|_{\alpha},$$

where the linear operator $DF(u)$ stands for the Fréchet derivative of F at u and $DF(u).h$ stands for the derivative of F at u in the direction h.

Finally, for all $u_1, u_2, h \in Dom((-A)^{\alpha+\kappa})$,

$$\big|\big(DF(u_2) - DF(u_1)\big).h\big|_{-\alpha} \leq C_{F,\alpha,\kappa}\big(1 + |u_1|_{\alpha+\kappa} + |u_2|_{\alpha+\kappa}\big)|u_2 - u_1|_{-\alpha}|h|_{\alpha+\kappa}.$$

Let us now recall that a Nemytskii operator $F : H \to H$ is defined such that $F(u) = f \circ u$ for all $u \in H$, where $f : \mathbb{R} \to \mathbb{R}$ is a real-valued mapping, assumed to be at least globally Lipschitz continuous. As explained above, even if f is assumed of class $\mathcal{C}^2$ with bounded first- and second-order derivatives, the associated nonlinear Nemytskii operator F does not satisfy the conditions of Assumption 1. Indeed, the appropriate generalization requires estimates in $L^p(0,1)$ spaces for $p \in (2, \infty)$ (using Hölder's inequality).

2.3. Wiener process. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfying the usual conditions. The expectation operator is denoted by $\mathbb{E}[\cdot]$.

Let $((\beta_p(t))_{t\geq 0})_{p\in\mathbb{N}}$ be a sequence of independent standard real-valued Wiener processes, and let $(\epsilon_p)_{p\in\mathbb{N}}$ be a complete orthonormal system of H and $(\gamma_p)_{p\in\mathbb{N}}$ be a sequence of nonnegative real numbers. The cylindrical Wiener process is defined as

$$W(t) = \sum_{p\in\mathbb{N}} \beta_p(t)\epsilon_p.$$

The Q-Wiener process is then defined as

$$W^Q(t) = \sum_{p\in\mathbb{N}} \sqrt{\gamma_p}\beta_p(t)\epsilon_p$$

and can be written as $W^Q(t) = Q^{\frac{1}{2}}W(t)$, where the linear self-adjoint operators $Q^{\frac{1}{2}}$ and Q satisfy

$$Q^{\frac{1}{2}}u = \sum_{p\in\mathbb{N}} \sqrt{\gamma_p}\langle u, \epsilon_p\rangle\epsilon_p, \ \ Qu = \sum_{p\in\mathbb{N}} \gamma_p\langle u, \epsilon_p\rangle\epsilon_p \quad \forall\, u \in H.$$

Note that the Q-Wiener process $W^Q(t)$ takes values in H if and only if Q is a trace-class linear operator, i.e., $\mathrm{Tr}(Q) = \|Q^{\frac{1}{2}}\|^2_{\mathcal{L}_2(H)} = \sum_{p\in\mathbb{N}} \gamma_p < \infty$.

Assumption 2 states the conditions on Q required to ensure the well-posedness of (1).

Assumption 2. Assume that there exists $\alpha > 0$ such that $\|(-A)^{\alpha-\frac{1}{2}}Q^{\frac{1}{2}}\|_{\mathcal{L}_2(H)} < \infty$.

Define the parameter $\overline{\alpha}$ as follows:

$$\overline{\alpha} = \sup\left\{\alpha \in (0,\infty),\ \|(-A)^{\alpha-\frac{1}{2}}Q^{\frac{1}{2}}\|_{\mathcal{L}_2(H)} < \infty\right\}.$$

Then $\overline{\alpha} > 0$ if and only if Assumption 2 is satisfied. For instance, if $Q = I$ (space-time white noise), then $\overline{\alpha} = \frac{1}{4}$. If Q is a trace-class operator, then $\overline{\alpha} = \frac{1}{2}$. In the article, we are mostly interested in the regime $\overline{\alpha} \in (0, \frac{1}{2}]$.

The numerical experiments are performed with the following example: For all $p \in \mathbb{N}$, $\epsilon_p = e_p$ (thus, the operators A and Q commute), and $\gamma_p = \lambda_p^{\frac{1}{2}-2\overline{\alpha}}$ (observe that in that case, the notation is consistent with the definition of $\overline{\alpha}$ in the general case).

To conclude this section, let us introduce the following notation: If U is an H-valued random variable for all $\alpha \in [0,1]$ and $q \in \mathbb{N}$,

$$\|U\|_q = \left(\mathbb{E}[|U|^q]\right)^{\frac{1}{q}}, \quad \|U\|_{q,\alpha} = \left(\mathbb{E}[|U|_\alpha^q]\right)^{\frac{1}{q}}.$$

2.4. Well-posedness and regularity properties. Solutions of (1) are understood in the mild sense: For all $t \geq 0$,

$$(2) \qquad u(t) = e^{tA}u_0 + \int_0^t e^{(t-s)A}F(u(s))ds + \int_0^t e^{(t-s)A}dW^Q(s).$$

Under the assumptions stated above, this problem is globally well-posed. We quote without proof the following standard result (see, for instance, [5, Chapter 7] and [18, section 10.4] for similar statements).

PROPOSITION 2.2. *Let Assumption 2 be satisfied. For any initial condition $u_0 \in H$, there exists a unique mild solution (2) of the SPDE (1). Moreover, for any $T > 0$ and $q \in \mathbb{N}$, there exists $C_{T,q} \in (0,\infty)$ such that*

$$\sup_{0 \leq t \leq T} \|u(t)\|_q \leq C_{T,q}\left(1 + |u_0|\right).$$

Moreover, for any $\alpha \in (0, \min(\overline{\alpha}, \frac{1}{2}))$, there exists $C_{T,q,\alpha} \in (0,\infty)$ such that

$$\|u(t)\|_{q,\alpha} \leq C_{T,q,\alpha}\left(1 + \min\left(|u_0|_\alpha, t^{-\alpha}|u_0|\right)\right) \ \forall\, t \in (0,T],$$
$$\|u(t) - u(s)\|_q \leq C_{T,q,\alpha}|t-s|^\alpha\left(1 + \min\left(|u_0|_\alpha, \min(t,s)^{-\alpha}|u_0|\right)\right) \ \forall\, t,s \in (0,T].$$

3. Parareal algorithms.

3.1. Fine and coarse integrators. Let $T \in (0,\infty)$ be given. Introduce the so-called coarse and fine time-step sizes ΔT and δt. It is assumed that $T = N\Delta T$ and $\Delta T = J\delta t$, where N and J are integers. For all $n \in \{0,\ldots,N\}$ and $j \in \{0,\ldots,J\}$, let

$$t_n = n\Delta T, \quad t_{n,j} = t_n + j\delta t = (nJ + j)\delta t.$$

Note that the coarse and the fine integrators introduced below are random mappings. More precisely, for all $n \in \mathbb{N}$, the mappings $\mathcal{G}_n = \mathcal{G}(\cdot, t_n, t_{n+1})$ and $\mathcal{F}_n = \mathcal{F}(\cdot, t_n, t_{n+1})$ depend on the increments $(W^Q(t) - W^Q(t_n))_{t_n \leq t \leq t_{n+1}}$.

3.1.1. Coarse integrator. The coarse integrator is a numerical scheme with time-step size ΔT. In this work, it is defined as follows: For all $n \in \{0, \ldots, N-1\}$ and all $u \in H$, let

$$(3) \qquad \mathcal{G}(u, t_n, t_{n+1}) = \hat{S}_{\Delta T} u + \Delta T \hat{S}_{\Delta T} F(u) + \hat{S}_{\Delta T}\big(W^Q(t_{n+1}) - W^Q(t_n)\big)$$

with

- either $\hat{S}_{\Delta T} = e^{\Delta T A}$ (exponential Euler scheme);
- or $\hat{S}_{\Delta T} = S_{\Delta T} = (I - \Delta T A)^{-1}$ (linear implicit Euler scheme).

The notation $\mathcal{G}_{\text{expo}}$ (resp., $\mathcal{G}_{\text{imp}}$) is often used below to refer to the coarse integrator with the exponential Euler scheme (resp., with the linear implicit Euler scheme). As will be seen below, these two coarse integrators have very different behaviors when applied to the SPDEs considered in this article.

3.1.2. Fine integrator. The fine integrator consists of J steps of a numerical scheme with time-step size δt. In this work, this numerical scheme is obtained by the exponential Euler scheme. More precisely, introduce the auxiliary integrator $\mathcal{F}_{\text{aux}}$ For all $n \in \{0, \ldots, N-1\}$, all $j \in \{0, \ldots, J-1\}$, and all $u \in H$,

$$\mathcal{F}_{\text{aux}}(u, t_{n,j}, t_{n,j+1}) = e^{\delta t A} u + \delta t e^{\delta t A} F(u) + e^{\delta t A}\big(W^Q(t_{n,j+1}) - W^Q(t_{n,j})\big).$$

The fine integrator $\mathcal{F}$, at the coarse time scale, is then defined as follows:

$$(4) \qquad \mathcal{F}(\cdot, t_n, t_{n+1}) = \mathcal{F}_{\text{aux}}(\cdot, t_{n,J-1}, t_{n,J}) \circ \cdots \circ \mathcal{F}_{\text{aux}}(\cdot, t_{n,0}, t_{n,1}).$$

In other words, the solution $v_{n,J} = \mathcal{F}(u, t_n, t_{n+1})$. is computed using the recursion formula

$$v_{n,j+1} = e^{\delta t A} v_{n,j} + \delta t e^{\delta t A} F(v_{n,j}) + e^{\delta t A}\big(W^Q(t_{n,j+1}) - W^Q(t_{n,j})\big)$$
$$v_{n,0} = u$$

for $j \in \{0, \ldots, J-1\}$ with $J \delta t = \Delta T$.

3.2. The parareal algorithm. The initialization step of the parareal algorithm consists in applying the coarse integrator: For $n \in \{0, \ldots, N-1\}$,

$$u_{n+1}^{(0)} = \mathcal{G}(u_n^{(0)}, t_n, t_{n+1}),$$
$$u_0^{(0)} = u_0.$$

Let $K \in \mathbb{N}$ denote the number of parareal iterations. Iterations for $k \in \{0, \ldots, K-1\}$ are defined as follows: Given the values $(u_m^{(k)})_{0 \le m \le N}$ at iteration k, compute, for all $n \in \{0, \ldots, N-1\}$,

$$(5) \qquad u_{n+1}^{(k+1)} = \mathcal{G}(u_n^{(k+1)}, t_n, t_{n+1}) + \mathcal{F}(u_n^{(k)}, t_n, t_{n+1}) - \mathcal{G}(u_n^{(k)}, t_n, t_{n+1}),$$
$$u_0^{(k)} = u_0.$$

The core of the approach lies in the ability, at each iteration in k, to perform in parallel the computations in (5) for different values of n, hence the terminology of "parareal algorithms" for "parallelization in real time."

3.3. The reference solution. The reference solution is defined using the fine integrator:

$$
\begin{aligned}
u_{n+1}^{\mathrm{ref}} &= \mathcal{F}(u_n^{\mathrm{ref}}, t_n, t_{n+1}), \\
u_0^{\mathrm{ref}} &= u_0.
\end{aligned}
\tag{6}
$$

Observe that following (4), this reference solution is in fact defined in terms of the integrator $\mathcal{F}_{\mathrm{aux}}$, applied with the fine time-step size δt. Precisely, for all $n \in \{0, \ldots, N-1\}$, one has $u_{n+1}^{\mathrm{ref}} = v_{n,J}^{\mathrm{ref}}$ defined by

$$
\begin{aligned}
v_{n,j+1}^{\mathrm{ref}} &= \mathcal{F}_{\mathrm{aux}}\big(v_{n,j}^{\mathrm{ref}}, t_{n,j}, t_{n,j+1}\big), \ j \in \{0, \ldots, J-1\} \\
v_{n,0}^{\mathrm{ref}} &= u_n^{\mathrm{ref}}.
\end{aligned}
$$

This may be rewritten as follows: $u_n^{\mathrm{ref}} = v_{nJ}^{\mathrm{ref}}$, where for all $\ell \in \{0, \ldots, NJ-1\}$,

$$
v_{\ell+1}^{\mathrm{ref}} = \mathcal{F}_{\mathrm{aux}}\big(v_\ell^{\mathrm{ref}}, \ell\delta t, (\ell+1)\delta t\big).
$$

In addition, note that

$$
v_\ell^{\mathrm{ref}} = e^{\ell\delta t A}u_0 + \delta t \sum_{l=0}^{\ell-1} e^{(\ell-l)\delta t A}F(v_l^{\mathrm{ref}}) + \sum_{l=0}^{\ell-1} e^{(\ell-l)\delta t A}\big(W^Q((l+1)\delta t) - W^Q(l\delta t)\big).
$$

To conclude this section, we state without proof two standard results concerning, first, the qualitative properties of the reference solution (moment estimates) and, second, the rate of convergence of the error $u_n^{\mathrm{ref}} - u(t_n)$. See, for instance, [18, sections 10.6 to 10.9] and references therein for similar statements.

PROPOSITION 3.1. *Let $T > 0$, $\alpha \in (0, \min(\overline{\alpha}, \frac{1}{2}))$, and $q \in \mathbb{N}$. There exists $C_{T,q,\alpha} \in (0, \infty)$ such that, for all $u_0 \in Dom((-A)^\alpha)$,*

$$
\sup_{0 \le n \le N-1} \||u_n^{\mathrm{ref}}\||_{q,\alpha} \le C_{T,q,\alpha}\big(1 + |u_0|_\alpha\big)
$$

and such that the following error estimate holds true:

$$
\sup_{0 \le n \le N-1} \||u_n^{\mathrm{ref}} - u(t_n)\||_q \le C_{T,q,\alpha}\delta t^\alpha\big(1 + |u_0|_\alpha\big).
$$

Remark 3.2. The assumption that $u_0 \in Dom((-A)^\alpha)$ may be weakened using Proposition 2.1.

Remark 3.3. Since noise is additive in the SPDE (1), the order of convergence in Proposition 3.1 may be larger than $\frac{1}{2}$ when $\overline{\alpha}$ is sufficiently large. This type of estimate is not considered in this article.

3.4. Error and residual operators. In the implementation of parareal algorithms, the reference solution u_n^{ref} defined above is not computed in practice. Instead, the quantity $u_n^{(k)}$ defined in (5) is computed. To estimate the error between $u_n^{(k)}$ and $u(t_n)$, due to Proposition 3.1, it is sufficient to study the error between $u_n^{(k)}$ and u_n^{ref}. Observe that the parareal solution $u_n^{(k)}$ may be computed with a lower computational cost than the reference solution u_n^{ref}, using parallel computations in (5) (except for the initialization). Error estimates are required to determine the choice of time-step

sizes ΔT and δt and of the number of parareal iterations K to achieve a given error criterion with minimal computational cost.

For all $n \in \{0, \ldots, N\}$ and $k \in \{0, \ldots, K\}$ (where K is the number of parareal iterations), let the error be defined by

$$\epsilon_n^{(k)} := u_n^{(k)} - u_n^{\mathrm{ref}}. \tag{7}$$

Note that $\epsilon_0^{(k)} = 0$ for all $k \in \mathbb{N}_0$.

Moreover, by construction, one gets $\epsilon_n^{(k)} = 0$ for all $n \leq k$, which indicates that the numerical solution $\{u_n^{(k)}\}_{n=0,\ldots,N}$ will definitely converge to the reference solution $\{u_n^{\mathrm{ref}}\}_{n=0,\ldots,N}$ if the iterated number K is sufficiently large, i.e., $K \geq N$. However, to get a speedup, in practice, the iterated number K will be chosen significantly smaller than N, which is further discussed in section 3.5.

It is convenient to introduce the residual operators defined by

$$\mathcal{R}_n(u) := \mathcal{F}_n(u) - \mathcal{G}_n(u) \tag{8}$$

for all $n \in \{0, \ldots, N-1\}$, where the notation $\mathcal{F}_n(u) := \mathcal{F}(u, t_n, t_{n+1})$ and $\mathcal{G}_n(u) := \mathcal{G}(u, t_n, t_{n+1})$ is used.

Then the error defined by (7) satisfies the recursion formula, where the residual operators $\mathcal{R}_n$ defined by (8) appear as

$$\begin{aligned}
\epsilon_{n+1}^{(k+1)} &= \mathcal{G}_n(u_n^{(k+1)}) + \mathcal{F}_n(u_n^{(k)}) - \mathcal{G}_n(u_n^{(k)}) - \mathcal{F}_n(u_n^{\mathrm{ref}}) \\
&= \mathcal{G}_n(u_n^{(k+1)}) - \mathcal{G}_n(u_n^{\mathrm{ref}}) + \mathcal{R}_n(u_n^{(k)}) - \mathcal{R}_n(u_n^{\mathrm{ref}}) \\
&= \hat{S}_{\Delta T}\epsilon_n^{(k+1)} + \Delta T \hat{S}_{\Delta T}\left[F(u_n^{(k+1)}) - F(u_n^{\mathrm{ref}}) \right] + \mathcal{R}_n(u_n^{(k)}) - \mathcal{R}_n(u_n^{\mathrm{ref}}),
\end{aligned} \tag{9}$$

where the linear operator $\hat{S}_{\Delta T}$ depends on the choice of the coarse integrator; see (3).

Up to this point, the choice of the coarse integrator plays no role in the presentation. The major finding of this article is that the behavior of the parareal algorithm applied for SPDEs (1) differs when choosing the exponential Euler scheme or the linear implicit Euler scheme as the coarse integrator. Indeed, the theoretical results and the numerical experiments reveal that as the number of parareal iterations k increases, the evolution of the order of convergence of the error $\epsilon_n^{(k)}$ has a different behavior, depending on the choice of coarse integrator.

3.5. Analysis of the computational cost. The objective of this section is to compare the costs for computing $u_n^{(k)}$ using the parareal algorithm 5 and for computing the reference solution u_n^{ref}.

The computational advantage of using the parareal algorithm is due to the possibility to compute the quantities $u_n^{(k+1)}$ in parallel for fixed $k \geq 0$; see (5). Let N_{proc} denote the number of available processors.

Let $T \in (0, \infty)$ denote the final time, and consider $n = N$ such that $N\Delta T = T$. Denote by $\tau_{\mathcal{G}}$ the computational time for one evaluation of $\mathcal{G}(u, t_n, t_{n+1})$. It is assumed that $\tau_{\mathcal{G}}$ does not depend on ΔT, $n \in \mathbb{N}_0$ and on $u \in H$. Denote also by $\tau_{\mathcal{F},\mathrm{aux}}$ the computational time for one realization of $\mathcal{F}_{\mathrm{aux}}(u, t_{n,j}, t_{n,j+1})$. Then the computational time for one realization of $\mathcal{F}(\cdot, t_n, t_{n+1})$, denoted by $\tau_{\mathcal{F}}$, is

$$\tau_{\mathcal{F}} = J\tau_{\mathcal{F},\mathrm{aux}} = \frac{\Delta T \tau_{\mathcal{F},\mathrm{aux}}}{\delta t}.$$

It is also assumed that $\tau_{\mathcal{F},\mathrm{aux}}(u, t_{n,j}, t_{n,j+1})$ does not depend on δt, n, j and on u.

3.5.1. Parareal algorithm. For the initialization step, the computational cost is equal to $N\tau_{\mathcal{G}}$ since at this stage no parallelization procedure is applied.

For each iteration of the algorithm, the third term $\mathcal{G}(u_n^{(k)}, t_n, t_{n+1})$ in (5) has already been computed at the previous iteration, and the values of the second term $\mathcal{F}(u_n^{(k)}, t_n, t_{n+1})$ may be computed in parallel. A sequential computation remains to be done; thus, the computational cost of one iteration of the parareal algorithm is

$$N\left(\tau_{\mathcal{G}} + \frac{\tau_{\mathcal{F}}}{N_{\mathrm{proc}}}\right).$$

If K iterations of the parareal algorithm are performed, the associated computational cost is thus equal to

$$\mathrm{Cost}^{\mathrm{parareal}} = (K+1)\frac{T}{\Delta T}\tau_{\mathcal{G}} + K\frac{T}{\delta t}\frac{\tau_{\mathcal{F},\mathrm{aux}}}{N_{\mathrm{proc}}}.$$

3.5.2. Reference solution. The reference solution u_n^{ref} is computed using the fine integrator $\mathcal{F}_{\mathrm{aux}}$ with time-step size δt; see (6). The associated computational cost is equal to

$$\mathrm{Cost}^{\mathrm{ref}} = \frac{T}{\Delta T}\tau_{\mathcal{F}} = \frac{T}{\delta t}\tau_{\mathcal{F},\mathrm{aux}}.$$

3.5.3. Efficiency. The efficiency of the parareal algorithm, compared with a direct simulation using the fine integrator only, is studied in terms of the ratio

$$\mathcal{E} = \frac{\mathrm{Cost}^{\mathrm{ref}}}{\mathrm{Cost}^{\mathrm{parareal}}} = \frac{1}{\frac{K}{N_{\mathrm{proc}}} + (K+1)\frac{\delta t}{\Delta T}\frac{\tau_{\mathcal{G}}}{\tau_{\mathcal{F},\mathrm{aux}}}}.$$

Note that since the efficiency never goes to infinity as ΔT goes to 0, the parareal algorithm does not improve the rate of convergence with respect to the time-step size. Instead, the parareal algorithm may improve the computational efficiency. To go further in the analysis of the efficiency of the parareal algorithm, it is essential to study the rate of convergence of the error $\epsilon_n^{(k)}$ in terms of ΔT.

On the one hand, assume that the rate of convergence does not depend on k. Then, to balance the errors $u_n^{(k)} - u_n^{\mathrm{ref}}$ and $u_n^{\mathrm{ref}} - u(n\Delta T)$, it is necessary to choose ΔT and δt of the same size; hence,

$$\mathcal{E} = \frac{\mathrm{Cost}^{\mathrm{ref}}}{\mathrm{Cost}^{\mathrm{parareal}}} = \frac{1}{\frac{K}{N_{\mathrm{proc}}} + C(K+1)\frac{\tau_{\mathcal{G}}}{\tau_{\mathcal{F},\mathrm{aux}}}}.$$

To maximize the efficiency $\mathcal{E}$ above, the optimal choice is apparently $K = 0$: Parareal iterations increase the computational cost (linearly in K) in spite of the use of parallelization.

On the other hand, assume that the rate of convergence of $\epsilon_n^{(k)}$ is strictly larger than the rate for $\epsilon_n^{(0)}$. Then the time-step sizes are chosen such that $\delta t = \mathrm{o}(\Delta T)$ as $\Delta T \to 0$, and the efficiency is then of size $\frac{N_{\mathrm{proc}}}{K}$. This means that parallelization ($N_{\mathrm{proc}} \geq 2$) reduces the cost and that the optimal choice is $K = 1$.

In sections 4 and 5, the rates of convergence of the error $\epsilon_n^{(k)}$ with respect to ΔT are studied, depending on the choice of the coarse integrator and on the regularity of the noise.

4. Linear-implicit Euler scheme as the coarse integrator. The objective of this section is to prove that when the coarse integrator is chosen as the linear implicit Euler scheme, i.e., $\mathcal{G} = \mathcal{G}_{\mathrm{imp}}$ with $\hat{S}_{\Delta T} = S_{\Delta T}$, the behavior of the parareal algorithm depends a lot on the regularity of the noise. More precisely, it is proved that the order of convergence of the error $\epsilon_n^{(k)}$ to 0 cannot exceed $\overline{\alpha}$ and in particular saturates when k increases. Essentially, the order of convergence (in the framework studied below) is equal to $\min(\overline{\alpha}, k+1)$. The theoretical and numerical results are consistent and show that the rates obtained by the theoretical analysis are sharp.

For the theoretical analysis developed in this section, the framework is as follows. First, the initial condition $u_0 = 0$ and the nonlinear operator $F = 0$ are set equal to 0. Second, the covariance Q of the noise commutes with A, i.e., $Qe_p = \gamma_p e_p$ for all $p \in \mathbb{N}$, and eigenvalues satisfy $\gamma_p = \lambda_p^{\frac{1}{2}-2\overline{\alpha}}$.

In this case, the solution is a Gaussian process, and $u(t) = \int_0^t e^{(t-s)A} dW^Q(s)$. Moreover, the recursion formula (9) for the error yields the equalities

$$\epsilon_n^{(k+1)} = S_{\Delta T}\epsilon_{n-1}^{(k+1)} + \left(e^{\Delta T A} - S_{\Delta T}\right)\epsilon_{n-1}^{(k)} = \sum_{m=0}^{n-1} S_{\Delta T}^{n-1-m}\left(e^{\Delta T A} - S_{\Delta T}\right)\epsilon_m^{(k)}.$$

4.1. Theoretical error estimates. The main theoretical result of this section is Theorem 4.1.

THEOREM 4.1. *Assume that $\gamma_p = \lambda_p^{\frac{1}{2}-2\overline{\alpha}}$ with $\overline{\alpha} > 0$. Let $T \in (0, \infty)$ and $k \in \mathbb{N}_0$.*
- *If $k + 1 < \overline{\alpha}$, then there exists $C_{T,k,\overline{\alpha}} \in (0, \infty)$ such that*

$$\sup_{n\Delta T \leq T} \||\epsilon_n^{(k)}\||_2 \leq C_{T,k,\overline{\alpha}}\Delta T^{k+1}.$$

- *If $k + 1 \geq \overline{\alpha}$, then for all $\alpha \in (0, \overline{\alpha})$, there exists $C_{T,k,\alpha} \in (0, \infty)$ such that*

$$\sup_{n\Delta T \leq T} \||\epsilon_n^{(k)}\||_2 \leq C_{T,k,\alpha}\Delta T^{\alpha}.$$

In particular, for $\overline{\alpha} = \frac{1}{4}$ ($Q = I$, space-time white noise) or $\overline{\alpha} = \frac{1}{2}$ (Q is trace-class), the rate of convergence does not depend on k, and performing parareal iterations does not increase the order of convergence; see section 3.5.

Proof. For all $p \in \mathbb{N}$, let $\epsilon_n^{(k)}(p) = \langle \epsilon_n^{(k)}, e_p \rangle$ denote the pth component of the error $\epsilon_n^{(k)}$. Then the expression above is rewritten as

$$\epsilon_n^{(k+1)}(p) = \sum_{m=0}^{n-1} V(-\lambda_p\Delta T)^{n-1-m}R(-\lambda_p\Delta T)\epsilon_m^{(k)}(p),$$

where, for $z \in (-\infty, 0]$, one has $V(z) := \frac{1}{1-z}$ and $R(z) := e^z - \frac{1}{1-z}$. The inequality $|R(z)| \leq 1 \wedge |z|^2$ yields

$$\left(\mathbb{E}|\epsilon_n^{(k+1)}(p)|^2\right)^{\frac{1}{2}} \leq C\left(1 \wedge (\lambda_p\Delta T)^2\right)\left(\sum_{m=0}^{n-1}\frac{1}{(1+\lambda_p\Delta T)^{n-1-m}}\right)\sup_{0\leq m\leq n}\left(\mathbb{E}|\epsilon_m^{(k)}(p)|^2\right)^{\frac{1}{2}}.$$

Observe that

$$\sum_{m=0}^{n-1}\frac{1}{(1+\lambda_p\Delta T)^{n-1-m}} \leq \sum_{m=0}^{\infty}\frac{1}{(1+\lambda_p\Delta T)^m} = \frac{1+\lambda_p\Delta T}{\lambda_p\Delta T}.$$

Thus, for all n and all k, one obtains (using a recursion argument)

$$\left(\mathbb{E}|\epsilon_n^{(k)}(p)|^2\right)^{\frac{1}{2}} \leq C\left(1 \wedge (\lambda_p \Delta T)\right) \sup_{0 \leq m \leq n} \left(\mathbb{E}|\epsilon_m^{(k-1)}(p)|^2\right)^{\frac{1}{2}}$$

$$\leq C^k\left(1 \wedge (\lambda_p \Delta T)\right)^k \sup_{0 \leq m \leq n} \left(\mathbb{E}|\epsilon_m^{(0)}(p)|^2\right)^{\frac{1}{2}}.$$

It remains to study the error $\epsilon_n^{(0)}$ at the initialization step. One has the identity

$$\epsilon_n^{(0)} = \sum_{m=0}^{n-1} \left(S_{\Delta T}^{n-m} - e^{(n-m)\Delta T A}\right)\left(W^Q((m+1)\Delta T) - W^Q(m\Delta T)\right).$$

Let us prove the following claim: There exists $C \in (0, \infty)$ such that

(10) $$\sup_{n \in \mathbb{N}} \mathbb{E}|\epsilon_n^{(0)}(p)|^2 \leq \frac{C\gamma_p}{\lambda_p}\left(1 \wedge (\lambda_p \Delta T)\right)^2.$$

Using the Itô isometry formula and the fact $b^m - a^m \leq mb^{m-1}(b-a)$ for all $0 \leq a \leq b$ and $m \in \mathbb{N}$ (with $b = V(-\lambda_p \Delta T)$ and $a = e^{-\lambda_p \Delta T}$), one obtains

$$\mathbb{E}|\epsilon_n^{(0)}(p)|^2 = \gamma_p \Delta T \sum_{m=0}^{n-1} \left(\frac{1}{(1+\lambda_p \Delta T)^{(n-m)}} - e^{-\lambda_p(n-m)\Delta T}\right)^2$$

$$= \gamma_p \Delta T \sum_{m=1}^{n} \left(\frac{1}{(1+\lambda_p \Delta T)^m} - e^{-\lambda_p m \Delta T}\right)^2$$

$$\leq \frac{\gamma_p \Delta T}{1 + \lambda_p \Delta T}\left(\frac{1}{(1+\lambda_p \Delta T)} - e^{-\lambda_p \Delta T}\right) + \gamma_p \Delta T \sum_{m=2}^{n} \frac{m^2\left(1 \wedge (\lambda_p \Delta T)^2\right)^2}{(1+\lambda_p \Delta T)^{2(m-1)}}$$

$$\leq \frac{\gamma_p\left(1 \wedge (\lambda_p \Delta T)^2\right)}{\lambda_p}\left(1 + \phi(\lambda_p \Delta T)\right)$$

using the inequality $|R(z)| \leq 1 \wedge |z|^2$ and where, for all $x \geq 0$, one has

$$\phi(x) = \sum_{m=2}^{\infty} \frac{m^2 x\left(1 \wedge x^2\right)}{(1+x)^{2(m-1)}}.$$

It remains to check that $\sup_{x \in (0,\infty)} \phi(x) < \infty$. On the one hand, for all $x \in [1, \infty)$, one has

$$\phi(x) \leq \sum_{m=2}^{\infty} \frac{x}{(1+x)^2} \frac{m^2}{(1+x)^{2(m-2)}} \leq 2^4 \sum_{m=2}^{\infty} m^2 4^{-m} < \infty.$$

On the other hand, for all $x \in (0, 1)$, one has $(1+x)^{m-1} \geq \frac{m(m-1)}{2} x^2 \geq C^{-1} m^2 x^2$ if $m \geq 2$; thus,

$$\phi(x) \leq \sum_{m=2}^{\infty} \frac{x}{(1+x)^{m-1}} \frac{m^2 x^2}{(1+x)^{m-1}} \leq C \sum_{m=1}^{\infty} \frac{x}{(1+x)^{m-1}} = C.$$

This concludes the proof of the claim (10). Then, using the expression $\gamma_p = \lambda_p^{\frac{1}{2} - 2\overline{\alpha}}$

of the eigenvalues of the covariance operator Q, one obtains

$$
\begin{aligned}
\||\epsilon_n^{(k)}\||_2 &= \left(\sum_{p=1}^{\infty} \mathbb{E}|\epsilon_n^{(k)}(p)|^2 \right)^{\frac{1}{2}} \\
&\leq C_k \left(\sum_{p=1}^{\infty} \frac{1}{\lambda_p^{\frac{1}{2}+2\overline{\alpha}}} \left(1 \wedge (\lambda_p \Delta T) \right)^{2(k+1)} \right)^{\frac{1}{2}}.
\end{aligned}
$$

It remains to identify the orders of convergence. On the one hand, assume that $k + 1 < \overline{\alpha}$. Using $(1 \wedge (\lambda_p \Delta T))^{2(k+1)} \leq (\lambda_p \Delta T)^{2(k+1)}$ yields

$$
\||\epsilon_n^{(k)}\||_2 \leq C \Delta T^{k+1} \left(\sum_{p=1}^{\infty} \lambda_p^{-\frac{1}{2} - 2(\overline{\alpha} - k - 1)} \right)^{\frac{1}{2}} \leq C_{k,\overline{\alpha}} \Delta T^{k+1}
$$

since $\overline{\alpha} - k - 1 > 0$.

On the other hand, assume that $k + 1 \geq \overline{\alpha}$, and let $\alpha \in (0, \overline{\alpha})$. Using the inequality $(1 \wedge (\lambda_p \Delta T))^{2(k+1)} \leq (\lambda_p \Delta T)^{2\alpha}$ yields

$$
\||\epsilon_n^{(k)}\||_2 \leq C \Delta T^{\alpha} \left(\sum_{p=1}^{\infty} \lambda_p^{-\frac{1}{2} - 2\overline{\alpha} + 2\alpha} \right)^{\frac{1}{2}} \leq C_{k,\alpha} \Delta T^{\alpha}.
$$

This concludes the proof of Theorem 4.1. $\qquad\Box$

4.2. Numerical experiments. The objective of this section is to demonstrate that the orders of convergence obtained in Theorem 4.1 are sharp. In addition, experiments in the semilinear case ($F \neq 0$) are also provided.

First, the SPDE $du(t) = Au(t) + dW^Q(t)$, with $u(0) = 0$, is considered, where the covariance Q is given as above ($Qe_p = \gamma_p e_p$, with $\gamma_p = \lambda_p^{\frac{1}{2} - 2\overline{\alpha}}$). Spatial discretization is performed using finite differences, with mesh size $h = 0.01$. In addition, the noise is truncated; i.e., the Q-Wiener process $W^Q(t)$ is replaced by $\sum_{p=1}^{P} \gamma_p^{\frac{1}{2}} \beta_p(t) e_p$ with $P = 100$. Numerical parameters are chosen as follows: The final time is $T = 1$, the fine time-step size is $\delta t = 2^{-13}$, and the coarse time-step size is $\Delta T = J\delta t$ with $J = 2^j$, $j = 4, \ldots, 9$. An average over $M = 100$ independent Monte Carlo samples is used to approximate the expectations.

First, Figure 1 reports numerical simulations for $\overline{\alpha} = 0.25$ (space-time white noise) and $\overline{\alpha} = 0.5$ (trace-class noise). For all values of k, the order of convergence in the $\|| \cdot \||_2$-norm sense is equal to $\overline{\alpha}$, as predicted by Theorem 4.1.

Second, Figures 2 (two fixed values of k and $\overline{\alpha}$ varies) and 3 (two fixed values of $\overline{\alpha}$ and k varies) allow us to check that the orders of convergence in Theorem 4.1 are sharp: The order is indeed equal to $\min(\overline{\alpha}, k + 1)$.

To conclude this section, we report numerical simulations in the semilinear case.

Figures 4 and 5 show the order for semilinear equation (1) with $F(u) = \cos(u)$ and $F(u) = 5\cos(u)$, respectively. The order for the additive noise case with $\overline{\alpha} = 4$ (on the left) is limited to $\frac{3}{2}$ when $k \geq 2$, which is the same as the deterministic case (on the right).

Remark 4.2. According to the numerical simulations in both linear and semilinear cases, for different scales of α (see Figures 2 and 4), the error decreases as the iterated number k increases, which implies the convergence of the algorithm in terms of k for fixed coarse step size ΔT.
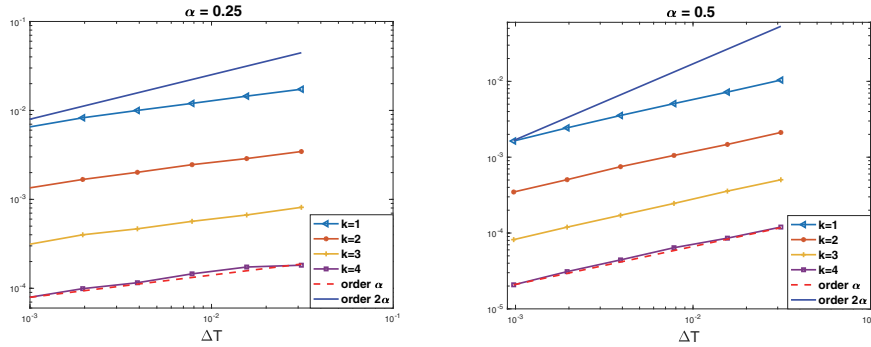
FIG. 1. *Orders of convergence of the error in the $\|\| \cdot \|\|_2$-norm sense with respect to ΔT, for $\overline{\alpha} = 0.25$ (left) and $\overline{\alpha} = 0.5$ (right), for different values of $k \in \{1, 2, 3, 4\}$, in the linear implicit Euler scheme case.*
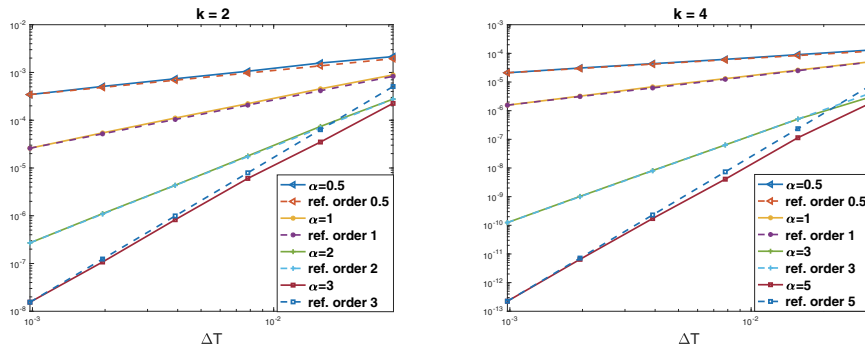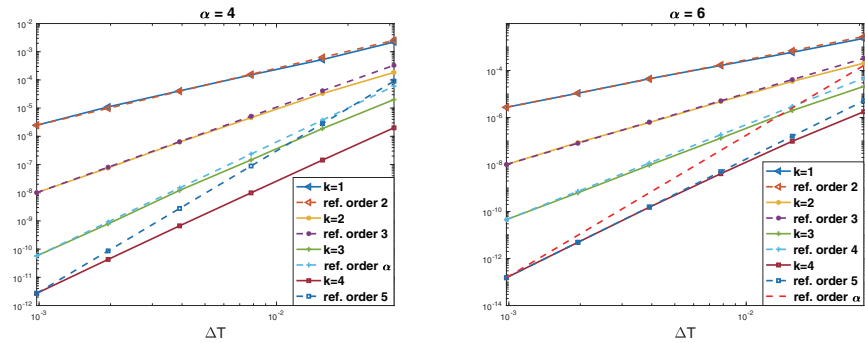

FIG. 2. *Orders of convergence of the error in the $\|\| \cdot \|\|_2$-norm sense with respect to ΔT, for $k = 2$ (left) and $k = 4$ (right), for different values of $\alpha = \overline{\alpha}$, in the linear implicit Euler scheme case.*

FIG. 3. *Orders of convergence of the error in the $\|\| \cdot \|\|_2$-norm sense with respect to ΔT, for $\overline{\alpha} = 4$ (left) and $\overline{\alpha} = 6$ (right), for different values of $k \in \{1, 2, 3, 4\}$, in the linear implicit Euler scheme case.*

5. Exponential Euler scheme as the coarse integrator. The objective of this section is to prove that when the exponential Euler scheme is chosen as the coarse integrator, i.e., $\mathcal{G} = \mathcal{G}_{\text{expo}}$ with $\hat{S}_{\Delta T} = e^{\Delta T A}$, then parareal iterations improve the
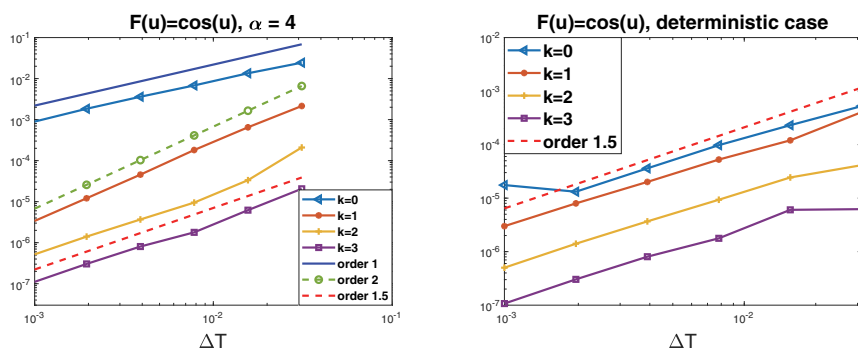
FIG. 4. *Orders of convergence of the error in the $\VERT \cdot \VERT_2$-norm sense with respect to ΔT, for the additive noise case with $\overline{\alpha} = 4$ (left) and the deterministic case (right), for nonlinear term $F(u) = \cos(u)$ and different values of $k \in \{0, 1, 2, 3\}$, in the linear implicit Euler scheme case.*
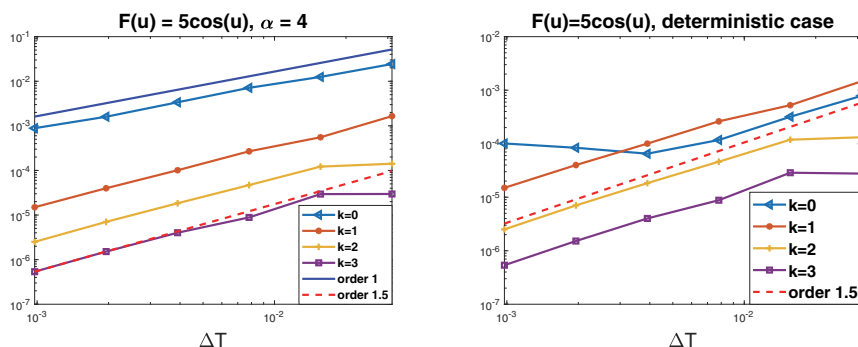

FIG. 5. *Orders of convergence of the error in the $\VERT \cdot \VERT_2$-norm sense with respect to ΔT, for the additive noise case with $\overline{\alpha} = 4$ (left) and the deterministic case (right), for nonlinear term $F(u) = 5\cos(u)$ and different values of $k \in \{0, 1, 2, 3\}$, in the linear implicit Euler scheme case.*

rate of convergence of the error $\epsilon_n^{(k)}$ to 0 in terms of ΔT. Contrary to the situation of section 4, this effect holds true without restrictions on the regularity parameter $\overline{\alpha}$, in particular for $\overline{\alpha} = \frac{1}{4}$ (space-time white noise).

The analysis in this section is performed for the SPDE (1), with the nonlinear coefficient F satisfying Assumption 1. Let also Assumption 2 be satisfied. The main results in this section are the error estimates stated in Theorems 5.1—which gives an order of convergence $(k + 1) \min(\overline{\alpha}, \frac{1}{2})$ for all $k \in \mathbb{N}_0$—and 5.2—which gives an improved order of convergence when $k \geq 2$; see section 5.1. Numerical experiments in section 5.2 illustrate that the result in Theorem 5.1 is sharp when $k = 0$ and $k = 1$ and that indeed better convergence rates are obtained for $k \geq 2$. Based on some a priori estimates given in section 5.3, proofs of the main results are provided in section 5.4.

Observe that when the coarse integrator is the exponential Euler scheme, the recursion formula (9) for the error yields the equality

$$
\begin{aligned}
\epsilon_n^{(k+1)} = \Delta T \sum_{m=0}^{n-1} e^{(n-m)\Delta TA} \left[F(u_m^{(k+1)}) - F(u_m^{\mathrm{ref}}) \right] \\
+ \sum_{m=0}^{n-1} e^{(n-1-m)\Delta TA} \left[\mathcal{R}_m(u_m^{(k)}) - \mathcal{R}_m(u_m^{\mathrm{ref}}) \right],
\end{aligned}
$$

(11)

where we recall that the residual operator $\mathcal{R}_n$ is defined by (8). As will be clear below, if $F = 0$, then $\epsilon_n^{(k)} = 0$ for all n as soon as $k \geq 1$. This property reveals why the choice of the exponential Euler scheme as the coarse and the fine integrator provides better results.

Note also that the fact that the noise is additive in (1) is fundamental in the analysis. In addition, since noise is additive, it is expected that the order of convergence for $k = 0$ may be equal to 1 (instead of $\frac{1}{2}$) if $\overline{\alpha}$ is sufficiently large (at least larger than $\frac{1}{2}$). This effect is not considered below since the objective is mainly to study the increase in the order of convergence produced by parareal iterations and in particular in situations where the noise is not very regular, i.e., for space-time white noise, with $\overline{\alpha} = \frac{1}{4}$.

5.1. Theoretical error estimates. The first error estimate is stated in Theorem 5.1, which may be interpreted as follows: Each parareal iteration improves the rate of convergence, proportionally to $\min(\overline{\alpha}, \frac{1}{2})$. Let us stress that Theorem 5.1 is optimal for $k = 0$ and $k = 1$, as illustrated by the numerical experiments reported in section 5.2.

THEOREM 5.1. *Let $T \in (0, \infty)$, $k \in \mathbb{N}_0$, $q \in \mathbb{N}$, $\alpha \in [0, \min(\overline{\alpha}, \frac{1}{2}))$, and $\kappa > 0$ be arbitrarily small. For all $u_0 \in Dom((-A)^\alpha)$, there exists $C_{T,k,q,\alpha,\kappa}(u_0) \in (0, \infty)$ such that, for all $\Delta T \in (0, 1)$,*

$$\sup_{n\Delta T \leq T} \|\!|\epsilon_n^{(k)}|\!\|_q \leq C_{T,k,q,\alpha,\kappa}(u_0) \Delta T^{(k+1)(\alpha-\kappa)}.$$

Following the discussion in section 3.5 concerning the computational cost, Theorem 5.1 shows that applying the parareal algorithm may be used to reduce the computational cost if the coarse integrator is the exponential Euler scheme, whatever the regularity of the noise.

The second result, Theorem 5.2, states that the estimate from Theorem 5.2 can be improved when $k \geq 2$. The practical relevance of this result is questionable: It requires $k \geq 2$, whereas it is expected (see section 3.5) that choosing $k = 1$ is optimal. Nevertheless, the study of the phenomenon stated in Theorem 5.2 is motivated by the numerical experiments reported in section 5.2, which exhibit that indeed the order of convergence is larger than $(k + 1)\min(\overline{\alpha}, \frac{1}{2})$; namely, it is at least equal to $(k - 1)\min(2\overline{\alpha}, \frac{1}{2}) + 2\alpha \leq k\min(2\overline{\alpha}, \frac{1}{2})$.

THEOREM 5.2. *Let $T \in (0, \infty)$, $k \in \mathbb{N} \setminus \{1\}$, $q \in \mathbb{N}$, $\alpha \in (0, \min(\overline{\alpha}, \frac{1}{2}))$, and $\kappa > 0$ be arbitrarily small. For all $u_0 \in Dom((-A)^\alpha)$, There exists $C_{T,k,q,\alpha,\kappa}(u_0) \in (0, \infty)$ such that, for all $\Delta T \in (0, 1)$,*

$$\sup_{n\Delta T \leq T} \|\!|\epsilon_n^{(k)}|\!\|_q \leq C_{T,k,q,\alpha,\kappa}(u_0) \Delta T^{(k-1)(\min(2\alpha,\frac{1}{2})-\kappa)+2\alpha}.$$

The proofs of Theorems 5.1 and 5.2 are postponed to section 5.4.

To simplify the exposition, the way the constants $C_{T,k,q,\alpha,\kappa}(u_0)$ above depend on $|u_0|_\alpha$ is not made precise.

Remark 5.3. Theorem 5.1 (and Lemma 5.1–(i)) can be proved under less restrictive conditions on the linear operator F instead of Assumption 1: There exists $\eta \in (0, \frac{1}{2})$ such that one has estimates of the type

$$|DF(u).h|_{-\eta-\alpha} \leq C_{F,\eta,\alpha,\kappa}(1 + |u|_{\alpha+\kappa})|h|_{-\alpha}$$
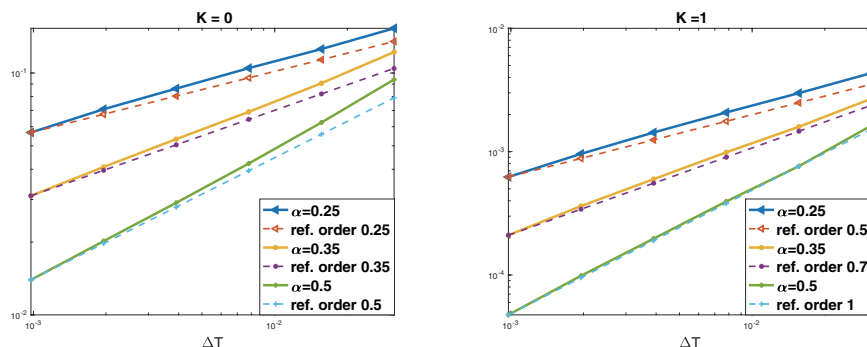
FIG. 6. *Orders of convergence of the error in the $\| \cdot \|_2$-norm sense with respect to ΔT, for $K = 0$ (left) and $K = 1$ (right), for different values of $\alpha = \overline{\alpha} \in \{0.25, 0.35, 0.5\}$, in the exponential Euler scheme case.*
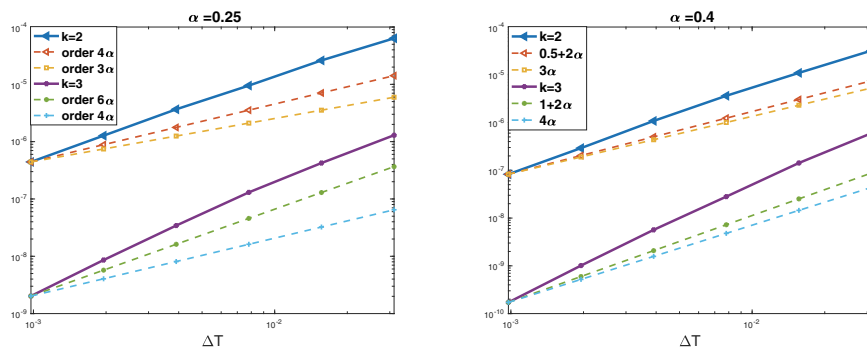

FIG. 7. *Orders of convergence of the error in the $\| \cdot \|_2$-norm sense with respect to ΔT, for $\alpha = \overline{\alpha} = 0.25$ (left) and $\alpha = \overline{\alpha} = 0.4$ (right), for different values of $k \in \{2, 3\}$, in the exponential Euler scheme case.*

and

$$|D^2 F(u).(h_1, h_2)|_{-\eta} \leq C_{F,\eta} |h_1| |h_2|.$$

This setting encompasses the case of Nemytskii operators, with $\eta \in (\frac{1}{4}, \frac{1}{2})$.

5.2. Numerical experiments. The objective of this section is to illustrate Theorems 5.1 and 5.2.

The SPDE (1), with nonlinear operator $F(u) = 5\cos(u)$ and initial condition $u(0) = 0$, is considered, with covariance operator given by $Qe_p = \gamma_p e_p$, with $\gamma_p = \lambda_p^{\frac{1}{2} - 2\overline{\alpha}}$. Spatial discretization is performed using finite differences, with mesh size $h = 0.01$. In addition, the noise is truncated; i.e., the Q-Wiener process $W^Q(t)$ is replaced by $\sum_{p=1}^{P} \gamma_p^{\frac{1}{2}} \beta_p(t) e_p$, with $P = 100$. Numerical parameters are chosen as follows: The final time is $T = 1$, the fine time-step size is $\delta t = 2^{-15}$, and the coarse time-step size is $\Delta T = J\delta t$ with $J = 2^j$, $j = 5, \ldots, 10$. An average over $M = 100$ independent Monte Carlo samples is used to approximate the expectations.

Figure 6 demonstrates that the orders of convergence in the $\| \cdot \|_2$-norm sense proved in Theorem 5.1 are sharp when $k = 0$ and $k = 1$ for different values of $\overline{\alpha}$.

Figure 7 then illustrates that for $k \geq 2$, the orders of convergence in Theorem 5.1 are not sharp for $\overline{\alpha} = 0.25$ and $\overline{\alpha} = 0.4$. In fact, even the improved error estimates

from Theorem 5.2 seem to be suboptimal in the numerical experiment, especially when $\overline{\alpha} = 0.4$ (figure on the right)—in this case $\min(2\alpha, \frac{1}{2}) = \frac{1}{2}$. When $\overline{\alpha} = 0.25$ (figure on the left), the improved theoretical rate in Theorem 5.2 is closer to the observed numerical rate.

5.3. A priori estimates. The objective of this section is to give some a priori estimates which will be used in the proofs of the main theorems. The following Lemma concerns Lipschitz continuity properties, in appropriate norms, of the residual operators.

LEMMA 5.1. *Let Assumptions 1 and 2 hold, and $\alpha \in [0, 1)$.*
(i) *There exists $C_{F,\alpha} \in (0, \infty)$ such that, for all $\Delta T \in (0, 1]$ and all $u_1, u_2 \in H$, one has*

$$\sup_{n\Delta T \leq T} |\mathcal{R}_n(u_2) - \mathcal{R}_n(u_1)|_\alpha \leq C_{F,q,\alpha} \Delta T^{1-\alpha} |u_2 - u_1|.$$

(ii) *For all $q \in \mathbb{N}$, $\alpha \in (0, \min(\overline{\alpha}, \frac{1}{2}))$, $\beta \in [0, \alpha)$, and arbitrarily small $\kappa \in (0, \min(\alpha - \beta, 1 - \alpha - \beta))$, there exists $C_{T,q,\alpha,\beta,\kappa} \in (0, \infty)$ such that, for all $\Delta T \in (0, 1)$ and all $u_1, u_2 \in Dom((-A)^\alpha)$, one has*

$$\||(-A)^{-\alpha}\big(\mathcal{R}_n(u_2) - \mathcal{R}_n(u_1)\big)\||_q$$
$$\leq C_{T,q,\alpha,\beta,\kappa} \Delta T^{1+\min\left(\alpha+\beta, \frac{1}{2}\right)-\kappa} \big(1 + |u_1|_\alpha^2 + |u_2|_\alpha^2\big) |u_2 - u_1|_\beta.$$

Proof. The proof consists in proving bounds on the derivative

$$D\mathcal{R}_n(u).h = D\mathcal{F}_n(u).h - DG_n(u).h,$$

where the residual operator $\mathcal{R}_n$ is defined by (8). On the one hand, due to (3), it is straightforward to compute

$$DG_n(u).h = e^{\Delta TA}h + \Delta T e^{\Delta TA} DF(u).h.$$

On the other hand, the derivative of $\mathcal{F}_n$ is computed as follows: One has $D\mathcal{F}_n(u).h = \eta_{n,J}^h$, where $J\delta t = \Delta T$, and one has the following recursion formulaes:

$$v_{n,j+1} = e^{\delta tA}v_{n,j} + \delta t e^{\delta tA}F(v_{n,j}) + e^{\delta tA}\delta_{n,j}W^Q , \quad v_{n,0} = u,$$
$$\eta_{n,j+1}^h = e^{\delta tA}\eta_{n,j}^h + \delta t e^{\delta tA}DF(v_{n,j}).\eta_{n,j}^h , \quad \eta_{n,0}^h = h.$$

Then a straightforward computation yields

$$\eta_{n,J}^h = e^{\Delta TA}h + \delta t \sum_{j=0}^{J-1} e^{(J-j)\delta tA} DF(v_{n,j}).\eta_{n,j}^h.$$

Observe that the same term $e^{\Delta TA}h$ appears in expressions of the derivatives $D\mathcal{F}_n(u).h$ and $DG_n(u).h$. This term thus does not appear in $D\mathcal{R}_n(u)$. This is a crucial property (which is not satisfied when the coarse integrator is the linear implicit Euler scheme). Observe also that $\mathcal{R}_n = 0$ if $F = 0$ since the linear part is solved exactly.

Proof of (i). Since F is globally Lipschitz continuous, applying the discrete Gronwall lemma yields the following (almost sure) inequality:

$$\sup_{0 \leq j \leq J} |\eta_{n,j}^h| \leq C|h|.$$

Due to Proposition 2.1 and using the global Lipschitz continuity property of F, it is straightforward to check that

$$
\begin{aligned}
\left|D\mathcal{R}_n(u).h\right|_\alpha = \left|D\mathcal{F}_n(u).h - D\mathcal{G}_n(u).h\right|_\alpha & \\
&\leq \left|\Delta T e^{\Delta T A} DF(u).h\right|_\alpha + \delta t \sum_{j=0}^{J-1} \left|e^{(J-j)\delta t A} DF(v_{n,j}).\eta_{n,j}^h\right|_\alpha \\
&\leq \Delta T \|(-A)^\alpha e^{\Delta T A}\|_{\mathcal{L}(H)} |h| + \delta t \sum_{j=0}^{J-1} \|(-A)^\alpha e^{(J-j)\delta t A}\|_{\mathcal{L}(H)} |\eta_{n,j}^h| \\
&\leq C_\alpha \Delta T^{1-\alpha} |h|
\end{aligned}
$$

using the inequality $\delta t \sum_{j=0}^{J-1} \frac{1}{(J-j)\delta t)^\alpha} \leq C_\alpha \Delta T^{1-\alpha}$. This concludes the proof. $\quad\square$

Observe that in the proof of (i), only the Lipschitz continuity property for F was employed. In the proof of (ii), the conditions stated in Assumption 1 are crucial.

Proof of (ii). Using the notation above and expressions of $D\mathcal{F}_n(u).h$ and of $D\mathcal{G}_n(u).h$, the derivative $D\mathcal{R}_n(u).h = D\mathcal{F}_n(u).h - D\mathcal{G}_n(u).h$ is decomposed as follows:

$$
\begin{aligned}
D\mathcal{R}_n(u).h = \left(\sum_{j=0}^{J-1} \delta t e^{(J-j)\delta t A} - \Delta T e^{\Delta T A}\right) DF(u).h \\
+ \delta t \sum_{j=0}^{J-1} e^{(J-j)\delta t A} DF(u).\left(\eta_{n,j}^h - h\right) \\
+ \delta t \sum_{j=0}^{J-1} e^{(J-j)\delta t A} \left(DF(v_{n,j}) - DF(u)\right).\eta_{n,j}^h.
\end{aligned}
$$

The first term is treated as follows: Using Assumption 1,

$$
\begin{aligned}
\left|(-A)^{-\alpha} \left(\sum_{j=0}^{J-1} \delta t e^{(J-j)\delta t A} - \Delta T e^{\Delta T A}\right) DF(u).h\right| & \\
&\leq C\delta t \sum_{j=0}^{J-1} \|(-A)^{-\alpha-\beta+\kappa}\left(e^{J\delta t A} - e^{(J-j)\delta t A}\right)\|_{\mathcal{L}(H)} |(-A)^{\beta-\kappa} DF(u).h| \\
&\leq C_{\alpha,\beta,\kappa} \delta t \sum_{j=0}^{J-1} (j\delta t)^{\alpha+\beta-\kappa}(1+|u|_\beta)|h|_\beta \\
&\leq C_{\alpha,\beta,\kappa} \Delta T^{1+\alpha+\beta-\kappa}(1+|u|_\alpha)|h|_\beta
\end{aligned}
$$

since $j\delta t \leq J\delta t = \Delta T$ and $\alpha + \beta \leq 2\alpha \leq 1$.

The second term is treated as follows: Observe that $\eta_{n,0}^h = h$; thus, using Assumption 1,

$$
\begin{aligned}
\left|(-A)^{-\alpha+\kappa}\delta t \sum_{j=0}^{J-1} e^{(J-j)\delta t A} DF(u).\left(\eta_{n,j}^h - h\right)\right| & \\
&\leq C_{\alpha,\kappa}\delta t \sum_{j=0}^{J-1} (1+|u|_\alpha)\left|(-A)^{-\alpha+\kappa}\left(\eta_{n,j}^h - \eta_{n,0}^h\right)\right|.
\end{aligned}
$$

Note that the increment $\eta_{n,j}^h - \eta_{n,0}^h$ satisfies

$$\eta_{n,j}^h - \eta_{n,0}^h = \big(e^{j\delta t A} - I\big)h + \delta t \sum_{\ell=0}^{j-1} e^{(j-\ell)\delta t A} D(v_{n,\ell}).\eta_{n,\ell}^h,$$

with the inequality

$$\big|(-A)^{-\alpha+\kappa}\big(e^{j\delta t A} - I\big)h\big| \leq C_{\alpha,\beta,\kappa}(j\delta t)^{\alpha+\beta-\kappa}|h|_\beta.$$

Using the inequality $|\eta_{n,j}^h| \leq C|h|$ and using that κ is chosen such that $\alpha + \beta + \kappa < 1$, one obtains

$$\left|(-A)^{-\alpha}\delta t \sum_{j=0}^{J-1} e^{(J-j)\delta t A} DF(u).\big(\eta_{n,j}^h - h\big)\right| \leq C_{\alpha,\beta,\kappa}\Delta T^{1+\alpha+\beta-\kappa}(1+|u|_\alpha)|h|_\beta.$$

It remains to treat the third term: Using $\beta + \kappa \leq \alpha$, the last inequality from Assumption 1 yields

$$\left|(-A)^{-\alpha}\delta t \sum_{j=0}^{J-1} e^{(J-j)\delta t A}\big(DF(v_{n,j}) - DF(u)\big).\eta_{n,j}^h\right|$$

$$\leq C_{\alpha,\beta,\kappa}\delta t \sum_{j=0}^{J-1} \left|(-A)^{-\beta-\kappa}\big(DF(v_{n,j}) - DF(v_{n,0})\big).\eta_{n,j}^h\right|$$

$$\leq C_{\alpha,\beta,\kappa}\delta t \sum_{j=0}^{J-1} \big(1 + |v_{n,0}|_\beta + |v_{n,j}|_\beta\big)|\eta_{n,j}^h|_\beta |v_{n,j} - v_{n,0}|_{-\beta-\kappa},$$

with $|\eta_{n,j}^h|_\beta \leq C|h|_\beta$. Note that the following moment bound holds true: There exists $C_{q,\beta} \in (0,\infty)$ such that $\||v_{n,j}\||_{q,\beta} \leq C_{q,\beta}$. Moreover, the proof of the following regularity result is straightforward:

$$\||(-A)^{-\beta-\kappa}(v_{n,j} - v_{n,0})\||_q \leq C_{q,\alpha,\beta,\kappa}(j\delta t)^{\min\left(\alpha+\beta+\kappa,\frac{1}{2}\right)}(1+|u|_\alpha).$$

See Proposition 2.2 for a similar estimate when $\beta = 0$ for the process $(u(t))_{t\geq 0}$. Then applying Hölder inequality yields

$$\left\|\left|(-A)^{-\alpha}\delta t \sum_{j=0}^{J-1} e^{(J-j)\delta t A}\big(DF(v_{n,j}) - DF(u)\big).\eta_{n,j}^h\right|\right\|_q$$

$$\leq C_{q,\alpha,\beta}\Delta T^{1+\min\left(\alpha+\beta+\kappa,\frac{1}{2}\right)}|h|_\beta(1+|u|_\alpha^2).$$

Gathering the estimates then concludes the proof. □

Observe that statement (i) (in the case $\alpha = 0$) is not sufficient to exhibit a positive rate of convergence γ for the error $\epsilon_n^{(k)}$ in terms of ΔT. Indeed, that would require the Lipschitz constant of the residual operator to be of size $\Delta T^{1+\gamma}$. Statement (i), with $\alpha = 0$, is used only to establish stability properties of the algorithm, i.e., moment bounds for $u_n^{(k)}$ or to apply a Gronwall lemma. With $\alpha > 0$, this statement is employed to establish moment bounds in $|\cdot|_\alpha$ norms and to deal with some terms for which regularization properties of the semigroup cannot be used.

Statement (ii) is the key result in order to get the rates of convergence given in Theorems 5.1 and 5.2. It requires choosing appropriate norms to analyze the Lipschitz constant of the residual operators: estimation in a weaker norm for $\alpha > 0$ or with higher regularity for $\beta > 0$. More precisely, Theorem 5.1 follows from statement (ii) with $\beta = 0$, and then Theorem 5.2 follows from statement (ii) with $\beta > 0$.

Finally, observe that the need to choose appropriate norms is specific to the infinite-dimensional situation. Moreover, Lemma 5.1 is not satisfied if the linear implicit Euler scheme is chosen as the coarse integrator, even if $F = 0$: Indeed, in that case, $\mathcal{R}_n(u_2) - \mathcal{R}_n(u_1) = (S_{\Delta T} - e^{\Delta T})(u_2 - u_1)$, and $\|S_{\Delta T} - e^{\Delta T}\|_{\mathcal{L}(H)}$ does not converge to 0 as $\Delta T \to 0$. This observation explains why the behaviors and the analysis of the parareal algorithm depends a lot on the choice of the coarse integrator for parabolic semilinear SPDEs.

Remark 5.4. Note that the estimates stated in Lemma 5.1 require u_1 and u_2 to be deterministic; however, below they are applied to random elements, which are measurable with respect to $\sigma(W(t); t \leq t_n)$, whereas the noise component in the residual operator $\mathcal{R}_n$ is measurable with respect to $\sigma(W(t); t_n \leq t \leq t_{n+1})$. Applying a straightforward conditioning argument and interpreting the expectation in Lemma 5.1 as a conditional expectation, yields the required estimates below.

Before giving the proofs of Theorems 5.1 and 5.2, it is also important to state moment bounds for the solution $u_n^{(k)}$, which are similar to Propositions 2.2 and 3.1 for the exact solution $u(t_n)$ and the reference solution u_n^{ref}, respectively.

PROPOSITION 5.5. *For* $T \in (0, \infty)$, $k \in \mathbb{N}_0$, *and* $q \in \mathbb{N}$, *there exists* $C_{T,k,q} \in (0, \infty)$ *such that, for all* $u_0 \in H$ *and* $\Delta T \in (0, 1)$,

$$\sup_{n\Delta T \leq T} \||u_n^{(k)}\||_q \leq C_{T,k,q}(1 + |u_0|).$$

Moreover, let $\alpha \in [0, \min(\overline{\alpha}, \frac{1}{2}))$. *There exists* $C_{T,k,q,\alpha} \in (0, \infty)$ *such that, for all* $u_0 \in Dom((-A)^\alpha)$ *and* $\Delta T \in (0, 1)$,

$$\sup_{n\Delta T \leq T} \||u_n^{(k)}\||_{q,\alpha} \leq C_{T,k,q,\alpha}(1 + |u_0|_\alpha).$$

Proof. First, moment bounds for the reference solution u_n^{ref} are provided by Proposition 3.1. In addition, in this section, the initialization step of the parareal algorithm consists in applying the fine integrator (which is the exponential Euler scheme) with time-step size ΔT; thus, the result of Proposition 3.1 also applies to obtain moment bounds for $u_n^{(0)}$.

Since $u_n^{(k)} = \epsilon_n^{(k)} + u_n^{\text{ref}}$, it thus only remains to prove moment bounds for $\epsilon_n^{(k)}$ when $k \geq 1$.

Using the expression of the error (11) and the global Lipschitz continuity of F, moments are treated as follows: For all $q \in \mathbb{N}$ and $\alpha \in [0, 1)$,

$$\||\epsilon_n^{(k+1)}\||_{q,\alpha} \leq \Delta T \sum_{m=0}^{n-1} \|(-A)^\alpha e^{(n-m)\Delta T A}\|_{\mathcal{L}(H)} \||\epsilon_m^{(k+1)}\||_{q,0}$$

$$+ \sum_{m=0}^{n-2} \|(-A)^\alpha e^{(n-1-m)\Delta T A}\|_{\mathcal{L}(H)} \||\mathcal{R}_m(u_m^{(k)}) - \mathcal{R}_m(u_m^{\text{ref}})\||_{q,0}$$

$$+ \||(-A)^\alpha \big[\mathcal{R}_{n-1}(u_{n-1}^{(k)}) - \mathcal{R}_{n-1}(u_{n-1}^{\text{ref}})\big]\||_{q,0}.$$

Assume first that $\alpha = 0$. Then the claim follows from the application of Lemma 5.1(i), the use of the discrete Gronwall lemma, and the use of a recursion argument with respect to k. When $\alpha > 0$, it remains to apply Lemma 5.1(i) and to apply the regularization estimate from Proposition 2.1 to conclude the proof. $\square$

5.4. Proofs of Theorems 5.1 and 5.2. Now we are in the position to prove the main theorems stated in section 5.1.

Proof of Theorem 5.1. First, assume that $k = 0$. The claim then follows from Proposition 3.1 since both the coarse and the fine integrators are given by the exponential Euler scheme; thus, $\epsilon_n^{(k)} = (u_n^{(k)} - u(t_n)) - (u_n^{\mathrm{ref}} - u(t_n))$, and the initialization of the parareal algorithm consists in applying the coarse integrator.

Let now $k \geq 1$. Due to the expression of the error (11), applying Lemma 5.1(ii) with $\beta = 0$ and Hölder inequality (see Remark 5.4 for the conditioning argument), one obtains

$$\||\epsilon_n^{(k)}\||_q \leq C\Delta T \sum_{m=0}^{n-1} \||\epsilon_m^{(k)}\||_q$$

$$+ \sum_{m=0}^{n-2} \|(-A)^\alpha e^{(n-1-m)\Delta T A}\|_{\mathcal{L}(H)} \||(-A)^{-\alpha}[\mathcal{R}(u_m^{(k-1)}) - \mathcal{R}(u_m^{\mathrm{ref}})]\||_q$$

$$+ \||\mathcal{R}_{n-1}(u_{n-1}^{(k-1)}) - \mathcal{R}_{n-1}(u_{n-1}^{\mathrm{ref}})\||_q$$

$$\leq C\Delta T \sum_{m=0}^{n-1} \||\epsilon_m^{(k)}\||_q + C\Delta T \||\epsilon_{n-1}^{(k-1)}\||_q$$

$$+ C\Delta T^{1+\alpha-\kappa} \sum_{m=0}^{n-2} \frac{\||\epsilon_m^{(k-1)}\||_{2q} \left(1 + \||u_m^{(k-1)}\||_{4q,\alpha}^2 + \||u_m^{\mathrm{ref}}\||_{4q,\alpha}^2\right)}{((n-m-1)\Delta T)^\alpha}.$$

Due to the moment estimates of Proposition 5.5, the discrete Gronwall lemma yields

$$\sup_{n\Delta T \leq T} \||\epsilon_n^{(k)}\||_q \leq C_{T,q,\alpha,\kappa}(u_0)\Delta T^{\alpha-\kappa} \sup_{n\Delta T \leq T} \||\epsilon_n^{(k-1)}\||_{2q}.$$

Then a straightforward recursion argument yields

$$\sup_{n\Delta T \leq T} \||\epsilon_n^{(k)}\||_q \leq C_{T,k,q,\alpha,\kappa}(u_0)\Delta T^{k(\alpha-\kappa)} \sup_{n\Delta T \leq T} \||\epsilon_n^{(0)}\||_{2^k q}$$

$$\leq C_{T,k,q,\alpha,\kappa}(u_0)\Delta T^{(k+1)(\alpha-\kappa)}$$

utilizing the result when $k = 0$. This concludes the proof of Theorem 5.1. $\square$

Proof of Theorem 5.2. Note that the error $\epsilon_n^{(k)}$ in (11) satisfies

$$\||\epsilon_n^{(k)}\||_{q,\alpha} \leq \Delta T \sum_{m=0}^{n-1} \|(-A)^\alpha e^{(n-m)\Delta T A}\|_{\mathcal{L}(H)} \||\epsilon_m^{(k)}\||_q$$

$$+ \sum_{m=0}^{n-2} \|(-A)^{2\alpha} e^{(n-1-m)\Delta T A}\|_{\mathcal{L}(H)} \||(-A)^{-\alpha}[\mathcal{R}_m(u_m^{(k-1)}) - \mathcal{R}_m(u_m^{\mathrm{ref}})]\||_q$$

$$+ \||\mathcal{R}_{n-1}(u_{n-1}^{(k-1)}) - \mathcal{R}_{n-1}(u_{n-1}^{\mathrm{ref}})\||_{q,\alpha}.$$

We split the proof into two steps.

Step 1. If $k \geq 1$, the error evaluated in the $|\cdot|_\alpha$ norm is also of the order $(k+1)(\alpha - \kappa)$:

$$\sup_{n\Delta T \leq T} \||\epsilon_n^{(k)}\||_{q,\alpha} \leq C_{T,k,q,\alpha,\kappa}(u_0)\Delta T^{(k+1)(\alpha-\kappa)}.$$

Indeed, utilizing Proposition 2.1 and Lemma 5.1, one obtains

$$\||\epsilon_n^{(k)}\||_{q,\alpha} \leq C \left(\Delta T \sum_{m=0}^{n-1} \frac{1}{((n-m)\Delta T)^\alpha} \right) \sup_{m\leq n} \||\epsilon_m^{(k)}\||_q$$

$$+ C \left(\Delta T \sum_{m=0}^{n-1} \frac{1}{((n-1-m)\Delta T)^{2\alpha}} \right) \Delta T^{\alpha-\kappa} \sup_{m\leq n} \||\epsilon_m^{(k-1)}\||_q$$

$$+ \Delta T^{1-\alpha} \||\epsilon_n^{(k-1)}\||_q$$

$$\leq C\Delta T^{(k+1)(\alpha-\kappa)} + C\Delta T^{1-\alpha+k(\alpha-\kappa)},$$

where Theorem 5.1 is used in the last step. With the assumption $\alpha \leq \frac{1}{2}$, it is straightforward to check that $1 - \alpha + k(\alpha - \kappa) \geq (k+1)(\alpha - \kappa)$. This concludes Step 1.

Step 2. It remains to establish the error estimate of Theorem 5.2. Similar as above, Proposition 2.1 and Lemma 5.1, with $\beta = \alpha - \kappa$, lead to

$$\||\epsilon_n^{(k)}\||_{q,\alpha} \leq C\Delta T \sum_{m=0}^{n-1} \frac{1}{((n-m)\Delta T)^\alpha} \||\epsilon_m^{(k)}\||_{q,\alpha}$$

$$+ C\Delta T^{\min(2\alpha,\frac{1}{2})-2\kappa} \left(\Delta T \sum_{m=0}^{n-2} \frac{1}{((n-1-m)\Delta T)^{2\alpha}} \right) \sup_{m\leq n-2} \||\epsilon_m^{(k-1)}\||_{2q,\alpha}$$

$$+ C\Delta T^{1-\alpha} \||\epsilon_{n-1}^{(k-1)}\||_q$$

$$\leq C\Delta T \sum_{m=0}^{n-1} \frac{1}{((n-m)\Delta T)^\alpha} \||\epsilon_m^{(k)}\||_{q,\alpha}$$

$$+ C\Delta T^{\min(2\alpha,\frac{1}{2})-2\kappa} \sup_{m\Delta T \leq T} \||\epsilon_m^{(k-1)}\||_{2q,\alpha},$$

where in the last step we use the fact that $1 - \alpha \geq \frac{1}{2} \geq \min(2\alpha, \frac{1}{2}) - 2\kappa$.

Applying the discrete Gronwall lemma and a recursion argument,

$$\||\epsilon_n^{(k)}\||_{q,\alpha} \leq C\Delta T^{(k-1)(\min(2\alpha,\frac{1}{2})-\kappa)} \||\epsilon_n^{(1)}\||_{2^k q,\alpha}$$

$$\leq C\Delta T^{(k-1)(\min(2\alpha,\frac{1}{2})-\kappa)+2\alpha}$$

due to Step 1 with $k = 1$. This concludes the proof of Theorem 5.2.   □

Remark 5.6. The claim in Step 1 in the proof of Theorem 5.2 requires $k \geq 1$ and is not correct when $k = 0$. In fact, when $k \geq 1$, in the expression of the error, contributions of stochastic terms, which would have low regularity properties, do not appear explicitly since noise is additive.

The application of the claim of Step 2 explains why Theorem 5.2 holds true only if $k \geq 2$.

REFERENCES

[1] G. BAL, *On the convergence and the stability of the parareal algorithm to solve partial differential equations*, in Domain Decomposition Methods in Science and Engineering, Lect. Notes Comput. Sci. Eng. 40, Springer, Berlin, 2005, pp. 425–432, https://doi.org/10.1007/3-540-26825-1_43.

[2] G. BAL, *Parallelization in Time of (Stochastic) Ordinary Differential Equations*, preprint, 2006.

[3] A. BELLEN AND M. ZENNARO, *Parallel algorithms for initial value problems for difference and differential equations*, J. Comput. Appl. Math., 25 (1989), pp. 341–350, https://doi.org/10.1016/0377-0427(89)90037-X.

[4] P. CHARTIER AND B. PHILIPPE, *A parallel shooting technique for solving dissipative ODEs*, Computing, 51 (1993), pp. 209–236, https://doi.org/10.1007/BF02238534.

[5] G. DA PRATO AND J. ZABCZYK, *Stochastic Equations in Infinite Dimensions*, 2nd ed., Encyclopedia Math. Appl. 152, Cambridge University Press, Cambridge, 2014, https://doi.org/10.1017/CBO9781107295513.

[6] M. J. GANDER, *50 years of time parallel time integration*, in Multiple Shooting and Time Domain Decomposition Methods, Contrib. Math. Comput. Sci. 9, Springer, Cham, Switzerland, 2015, pp. 69–113.

[7] M. J. GANDER AND E. HAIRER, *Nonlinear convergence analysis for the parareal algorithm*, in Domain Decomposition Methods in Science and Engineering XVII, Lect. Notes Comput. Sci. Eng. 60, Springer, Berlin, 2008, pp. 45–56, https://doi.org/10.1007/978-3-540-75199-1_4.

[8] M. J. GANDER AND E. HAIRER, *Analysis for parareal algorithms applied to Hamiltonian differential equations*, J. Comput. Appl. Math., 259 (2014), pp. 2–13, https://doi.org/10.1016/j.cam.2013.01.011.

[9] M. J. GANDER AND S. VANDEWALLE, *Analysis of the parareal time-parallel time-integration method*, SIAM J. Sci. Comput., 29 (2007), pp. 556–578, https://doi.org/10.1137/05064607X.

[10] M. J. GANDER AND S. VANDEWALLE, *On the superlinear and linear convergence of the parareal algorithm*, in Domain Decomposition Methods in Science and Engineering XVI, Lect. Notes Comput. Sci. Eng. 55, Springer, Berlin, 2007, pp. 291–298, https://doi.org/10.1007/978-3-540-34469-8_34.

[11] J. HONG, X. WANG, AND L. ZHANG, *Parareal exponential θ-scheme for longtime simulation of stochastic Schrödinger equations with weak damping*, SIAM J. Sci. Comput., 41 (2019), pp. B1155–B1177, http://doi.org/10.1137/18M1176749.

[12] G. HORTON AND S. VANDEWALLE, *A space-time multigrid method for parabolic partial differential equations*, SIAM J. Sci. Comput., 16 (1995), pp. 848–864, https://doi.org/10.1137/0916050.

[13] A. JENTZEN AND P. E. KLOEDEN, *Taylor Approximations for Stochastic Partial Differential Equations*, CBMS-NSF Regional Conf. Ser. Appl. Math. 93, SIAM, Philadelphia, 2011, http://doi.org/10.1137/1.9781611972016.

[14] M. KIEHL, *Parallel multiple shooting for the solution of initial value problems*, Parallel Comput., 20 (1994), pp. 275–295, https://doi.org/10.1016/0167-8191(94)90087-6.

[15] R. KRUSE, *Strong and Weak Approximation of Semilinear Stochastic Evolution Equations*, Lect. Notes Math. 2093, Springer, Cham, Switzerland, 2014, http://doi.org/10.1007/978-3-319-02231-4.

[16] J.-L. LIONS, Y. MADAY, AND G. TURINICI, *Résolution d'EDP par un schéma en temps "pararéel,"* C. R. Acad. Sci. Paris Sér. I Math., 332 (2001), pp. 661–668, https://doi.org/10.1016/S0764-4442(00)01793-6.

[17] W. LIU AND M. RÖCKNER, *Stochastic Partial Differential Equations: An Introduction*, Universitext, Springer, Cham, Switzerland, 2015, https://doi.org/10.1007/978-3-319-22354-4.

[18] G. J. LORD, C. E. POWELL, AND T. SHARDLOW, *An Introduction to Computational Stochastic PDEs*, Cambridge Texts Appl. Math., Cambridge University Press, New York, 2014, http://doi.org/10.1017/CBO9781139017329.

[19] C. LUBICH AND A. OSTERMANN, *Multigrid dynamic iteration for parabolic equations*, BIT, 27 (1987), pp. 216–234, https://doi.org/10.1007/BF01934186.

[20] Y. MADAY AND G. TURINICI, *A parareal in time procedure for the control of partial differential equations*, C. R. Math. Acad. Sci. Paris, 335 (2002), pp. 387–392, https://doi.org/10.1016/S1631-073X(02)02467-6.

[21] Y. MADAY AND G. TURINICI, *The parareal in time iterative solver: A further direction to parallel implementation*, in Domain Decomposition Methods in Science and Engineering, Lect. Notes Comput. Sci. Eng. 40, Springer, Berlin, 2005, pp. 441–448, https://doi.org/10.1007/3-540-26825-1_45.

[22] W. L. MIRANKER AND W. LINIGER, *Parallel methods for the numerical integration of ordinary differential equations*, Math. Comp., 21 (1967), pp. 303–320, https://doi.org/10.2307/2003233.

[23] J. NIEVERGELT, *Parallel methods for integrating ordinary differential equations*, Comm. ACM, 7 (1964), pp. 731–733, https://doi.org/10.1145/355588.365137.

[24] L. ZHANG, W. ZHOU, AND L. JI, *Parareal algorithms applied to stochastic differential equations with conserved quantities*, J. Comput. Math., 37 (2019), pp. 48–60, https://doi.org/10.4208/jcm.1708-m2017-0089.