



**HAL**  
open science

# On parareal algorithms for semilinear parabolic Stochastic PDEs

Charles-Edouard Bréhier, Xu Wang

► **To cite this version:**

Charles-Edouard Bréhier, Xu Wang. On parareal algorithms for semilinear parabolic Stochastic PDEs. SIAM Journal on Mathematical Analysis, 2020, 10.1137/19M1251011 . hal-02022432

**HAL Id: hal-02022432**

**<https://hal.science/hal-02022432>**

Submitted on 17 Feb 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On parareal algorithms for semilinear parabolic Stochastic PDEs

Charles-Edouard Bréhier and Xu Wang

ABSTRACT. Parareal algorithms are studied for semilinear parabolic stochastic partial differential equations. These algorithms proceed as two-level integrators, with fine and coarse schemes, and have been designed to achieve a “parallel in real time” implementation. In this work, the fine integrator is given by the exponential Euler scheme. Two choices for the coarse integrator are considered: the linear implicit Euler scheme, and the exponential Euler scheme.

The influence on the performance of the parareal algorithm, of the choice of the coarse integrator, of the regularity of the noise, and of the number of parareal iterations, is investigated, with theoretical analysis results and with extensive numerical experiments.

## 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], [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 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  $\alpha$ -Hölder continuous in time and  $2\alpha$ -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  $\alpha$  and  $2\alpha$  respectively.

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

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

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 (1), see also Section 2, where precise assumptions for 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 [3, 4, 12, 14, 19, 22, 23], and the review [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  $\Delta T$ , and a fine integrator, denoted by  $\mathcal{F}$ , using  $J$  steps with time-step size  $\delta t$ , 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

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

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

Since the pioneering work [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 stochastic Schrödinger equation with damping are studied with  $\mathcal{F}$  being the exact solver and  $\mathcal{G}$  being the exponential- $\theta$  scheme. The longterm convergence is obtained for the case  $\theta > \frac{1}{2}$  or sufficient large  $\alpha$ , 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, up to 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 not 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  $\Delta T$ ) is strictly larger for some  $k \geq 1$  than for  $k = 0$ .

To state the main results of this article (see Theorems 4.1, 5.2 and 5.3 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.*  $e^{\Delta T A}$  is approximated by  $(I - \Delta T A)^{-1}$ . 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  $\alpha$  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 5.2. 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 5.3 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, to presenting the possible choices of coarse integrators, and to 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 5.2 and 5.3 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  $u(0) = u_0 \in H$ ,

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

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, non-linear 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 sequel, 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

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

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  $Ae_p = -\lambda_p$  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  $u \in D((-A)^\alpha) = \{u \in H : |u|_\alpha < \infty\}$ , 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  $t \in (0, \infty)$ ,*

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

**2.2. Nonlinear operator  $F$ .** The analysis of the rate of convergence for 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  $\gamma$ -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 owing to 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  $F : H \rightarrow H$  is globally Lipschitz continuous, and is twice Fréchet differentiable, with bounded first and second order derivatives. 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  $u \in D((-A)^{\alpha+\kappa})$  and  $h \in H$ ,*

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

and for all  $u, h \in D((-A)^\alpha)$ ,

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

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

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

Let us now recall that a Nemytskii operator  $F : H \rightarrow H$  is defined such that  $F(u) = f \circ u$  for all  $u \in H$ , where  $f : \mathbb{R} \rightarrow \mathbb{R}$  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 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  $\left( (\beta_p(t))_{t \geq 0} \right)_{p \in \mathbb{N}}$  denote 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

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

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.*,  $\text{Tr}(Q) = \|Q^{\frac{1}{2}}\|_{\mathcal{L}_2(H)}^2 = \sum_{p \in \mathbb{N}} \gamma_p < \infty$ .

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  $\bar{\alpha}$  as follows:

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

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

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  $\gamma_p = \lambda_p^{\frac{1}{2}-2\bar{\alpha}}$  (observe that in that case the notation is consistent with the definition of  $\bar{\alpha}$  in the general case).

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

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

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

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

Under the assumptions stated above, this problem is globally well-posed. We quote without proof the following standard result.

**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*

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

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

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

### 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  $\Delta T$  and  $\delta t$ . It is assumed that  $T = N\Delta T$  and  $\Delta T = J\delta t$ , where  $N$  and  $J$  are integers. For all  $n \in \{0, \dots, N\}$  and  $j \in \{0, \dots, J\}$ , 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. 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 Wiener 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  $\Delta T$ . In this work, it is defined as follows: for all  $n \in \{0, \dots, N-1\}$  and all  $u \in H$ , let

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

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  $\delta t$ . 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  $n \in \{0, \dots, N-1\}$ , all  $j \in \{0, \dots, J-1\}$  and all  $u \in H$ ,

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

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

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

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

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

for  $j \in \{0, \dots, J-1\}$  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  $n \in \{0, \dots, N-1\}$

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

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

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

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:

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

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

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



This may be rewritten as follows:  $u_n^{\text{ref}} = v_n^{\text{ref}}$ , where for all  $\ell \in \{0, \dots, NJ - 1\}$ ,

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

In addition, note that

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

To conclude this section, we state without proof two standard results concerning, first, the qualitative properties of the reference solution (moment estimates), second, the rate of convergence of the error  $u_n^{\text{ref}} - u(t_n)$ .

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

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

and such that the following error estimate holds true:

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

**REMARK 3.2.** *The assumption that  $u_0 \in D((-A)^\alpha)$  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  $\bar{\alpha}$  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  $u_n^{\text{ref}}$  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  $u_n^{\text{ref}}$ . Observe that the parareal solution  $u_n^{(k)}$  may be computed with a lower computational cost than the reference solution  $u_n^{\text{ref}}$ , using parallel computations in (5) (expect for the initialization). Error estimates are required to determine the choice of time-step sizes  $\Delta T$  and  $\delta t$ , and of the number of parareal iterations  $K$ , to achieve a given error criterion, with minimal computational cost.

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

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

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

Moreover, by construction, one get  $\epsilon_n^{(k)} = 0$  for all  $n \leq k$ , which indicates that the numerical solution  $\{u_n^{(k)}\}_{n=0, \dots, N}$  will definitely converge to the reference solution  $\{u_n^{\text{ref}}\}_{n=0, \dots, N}$ , 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

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

for all  $n \in \{0, \dots, N-1\}$ , 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:

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

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  $u_n^{\text{ref}}$ .

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  $N_{\text{proc}}$  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  $\Delta T$ ,  $n \in \mathbb{N}_0$  and on  $u \in H$ . Denote also by  $\tau_{\mathcal{F}, \text{aux}}$  denote the computational time for one realization of  $\mathcal{F}_{\text{aux}}(u, t_{n,j}, t_{n,j+1})$ . Then the computational time for one realization of  $\mathcal{F}(\cdot, t_n, t_{n+1})$ , denoted by  $\tau_{\mathcal{F}}$ , is

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

It is also assumed that  $\tau_{\mathcal{F}, \text{aux}}(u, t_{n,j}, t_{n,j+1})$  does not depend on  $\delta t$ ,  $n, j$  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, observe that in (5), the third term  $\mathcal{G}(u_n^{(k)}, t_n, t_{n+1})$  has already been computed at the previous iteration, and that 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

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

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

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

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

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

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

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

Note that since the efficiency never goes to infinity as  $\Delta T$  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  $\Delta T$ .

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

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

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  $\delta t = o(\Delta T)$  as  $\Delta T \rightarrow 0$ , and the efficiency is then of size  $\frac{N_{\text{proc}}}{K}$ . This means that parallelization ( $N_{\text{proc}} \geq 2$ ) 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  $\Delta T$  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.*  $\mathcal{G} = \mathcal{G}_{\text{imp}}$  with  $\hat{S}_{\Delta T} = S_{\Delta T}$ , then 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  $\bar{\alpha}$ , and in particular saturates when  $k$  increases. Essentially, the order of convergence (in the framework studied below) is equal to  $\min(\bar{\alpha}, k + 1)$ . 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  $\gamma_p = \lambda_p^{\frac{1}{2} - 2\bar{\alpha}}$ .

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

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

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

**THEOREM 4.1.** *Assume that  $\gamma_p = \lambda_p^{\frac{1}{2}-2\bar{\alpha}}$ , with  $\bar{\alpha} > 0$ . Let  $T \in (0, \infty)$  and  $k \in \mathbb{N}_0$ .*

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

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

- *If  $k + 1 \geq \bar{\alpha}$ , then for all  $\alpha \in (0, \bar{\alpha})$ , 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  $\bar{\alpha} = \frac{1}{4}$  ( $Q = I$ , space-time white noise), or  $\bar{\alpha} = \frac{1}{2}$  ( $Q$  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  $p$ -th 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

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

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)

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

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

$$\epsilon_n^{(0)} = \sum_{m=0}^{n-1} (S_{\Delta t}^{n-m} - S((n-m)\Delta T)) (W^Q((m+1)\Delta T) - W^Q(m\Delta T)).$$

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

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

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

$$\begin{aligned} \mathbb{E} |\epsilon_n^{(0)}(p)|^2 &= \gamma_p \Delta T \sum_{m=0}^{n-1} \left( \frac{1}{(1 + \lambda_p \Delta T)^{(n-m)}} - e^{-\lambda_p(n-m)\Delta T} \right)^2 \\ &= \gamma_p \Delta T \sum_{m=1}^n \left( \frac{1}{(1 + \lambda_p \Delta T)^m} - e^{-\lambda_p m \Delta T} \right)^2 \\ &\leq \gamma_p \Delta T \sum_{m=1}^n \left( \frac{1}{(1 + \lambda_p \Delta T)^m} - e^{-\lambda_p m \Delta T} \right) \frac{1}{(1 + \lambda_p \Delta T)^m} \\ &\leq C\gamma_p \Delta T \sum_{m=1}^n \frac{m}{(1 + \lambda_p \Delta T)^{(2m-1)}} (1 \wedge (\lambda_p \Delta T))^2 \\ &\leq C\gamma_p \sum_{m=1}^{\infty} \frac{m\lambda_p \Delta T}{(1 + \lambda_p \Delta T)^{(m-1)}} \frac{(1 \wedge (\lambda_p \Delta T))^2}{\lambda_p(1 + \lambda_p \Delta T)} \\ &\leq C\gamma_p \frac{(1 \wedge (\lambda_p \Delta T))^2}{\lambda_p}. \end{aligned}$$

This concludes the proof of the claim (10). Then, using the expression  $\gamma_p = \lambda_p^{\frac{1}{2}-2\bar{\alpha}}$  of the eigenvalues of the covariance operator  $Q$ , one obtains

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

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

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

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

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

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

This concludes the proof of Theorem 4.1. □

**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  $\gamma_p = \lambda_p^{\frac{1}{2}-2\bar{\alpha}}$ ). 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  $J = 2^j$ ,  $j = 4, \dots, 9$ . An average over  $M = 100$  independent Monte-Carlo samples is used to approximate the expectations.

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

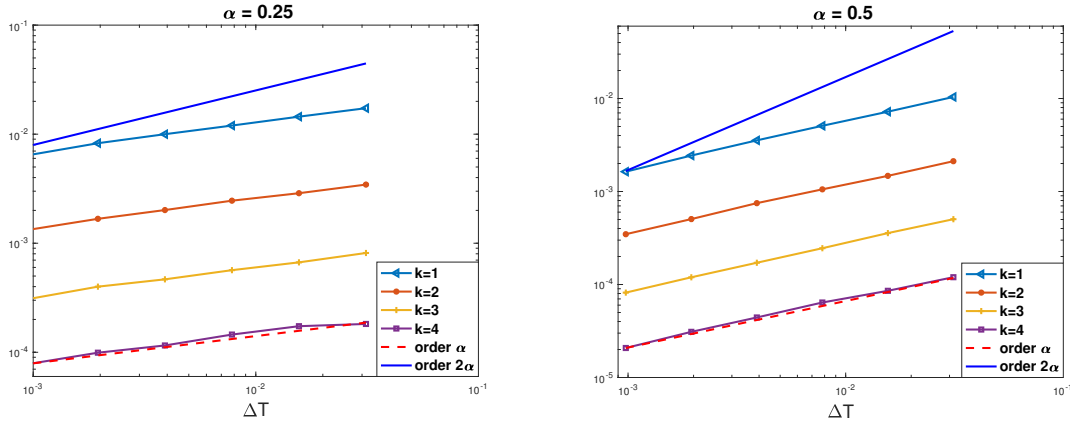


FIGURE 1. Orders of convergence of the error with respect to  $\Delta T$ , for  $\bar{\alpha} = 0.25$  (left) and  $\bar{\alpha} = 0.5$  (right), for different values of  $k \in \{1, 2, 3, 4\}$ , in the linear implicit Euler scheme case.

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

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  $\bar{\alpha} = 4$  (on the left) is limited to  $\frac{3}{2}$  when  $k \geq 2$ , which is the same as the deterministic case (on the right).

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

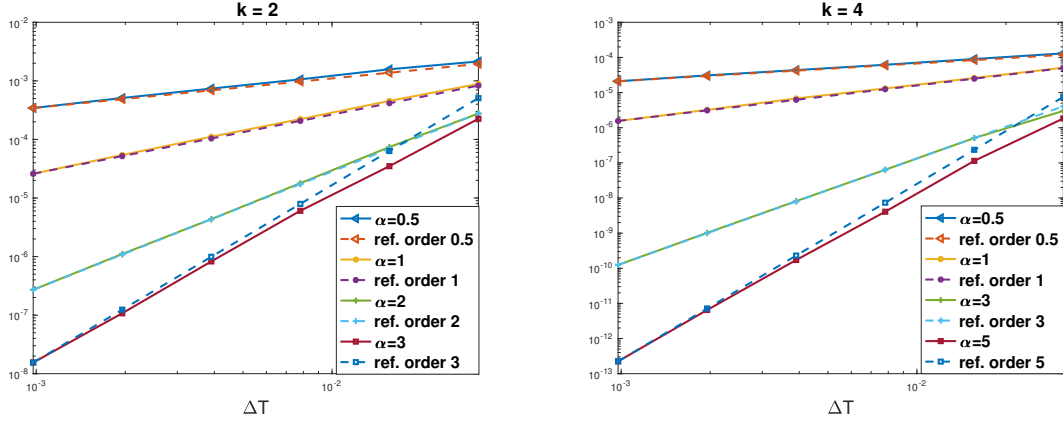


FIGURE 2. Orders of convergence of the error with respect to  $\Delta T$ , for  $k = 2$  (left) and  $k = 4$  (right), for different values of  $\bar{\alpha}$ , in the linear implicit Euler scheme case.

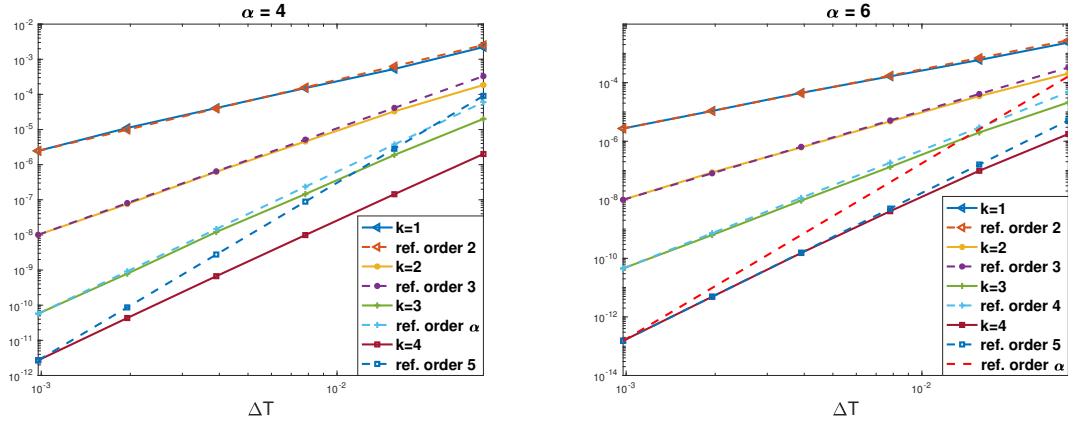


FIGURE 3. Orders of convergence of the error with respect to  $\Delta T$ , for  $\bar{\alpha} = 4$  (left) and  $\bar{\alpha} = 6$  (right), for different values of  $k \in \{1, 2, 3, 4\}$ , in the linear implicit Euler scheme case.

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 content of this section is organized as follows. The main results of this sections are the error estimates stated in Theorems 5.2 – which gives an order of convergence  $(k+1) \min(\bar{\alpha}, \frac{1}{2})$  for all  $k \in \mathbb{N}_0$  – and 5.3 – 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 5.2 is sharp when  $k = 0$  and  $k = 1$ , and that indeed better convergence rates are obtained for  $k \geq 2$ . Proofs of the results are provided in Section 5.3, based on auxiliary results which are proved in Section 5.4.

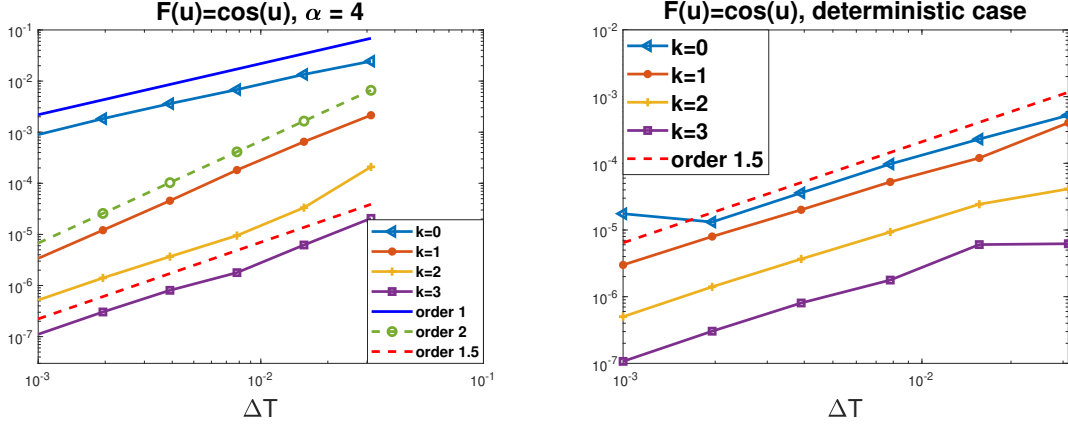


FIGURE 4. Orders of convergence of the error with respect to  $\Delta T$ , for additive noise case with  $\bar{\alpha} = 4$  (left) and deterministic case (right), for nonlinear term  $F(u) = \cos(u)$  and different values of  $k \in \{0, 1, 2, 3\}$ , in the linear implicit Euler scheme case.

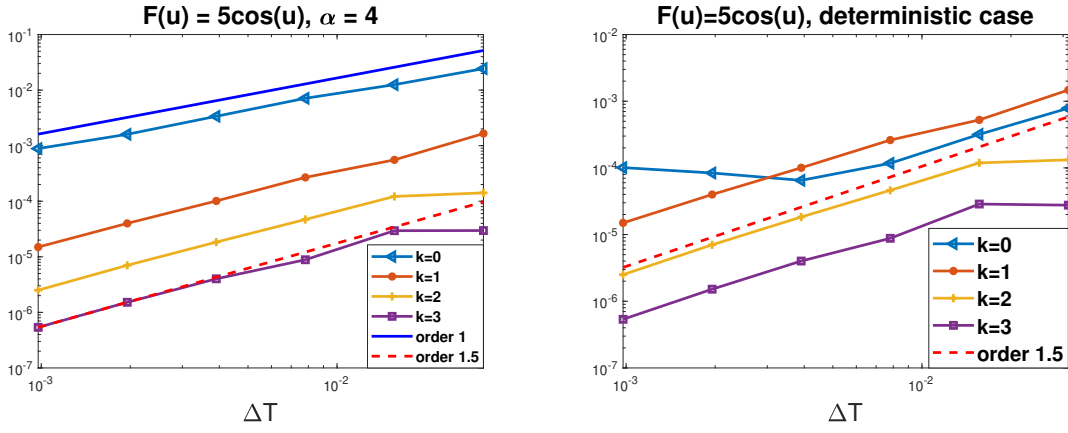


FIGURE 5. Orders of convergence of the error with respect to  $\Delta T$ , for additive noise case with  $\bar{\alpha} = 4$  (left) and deterministic case (right), for nonlinear term  $F(u) = 5 \cos(u)$  and different values of  $k \in \{0, 1, 2, 3\}$ , in the linear implicit Euler scheme case.

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

$$(11) \quad \epsilon_n^{(k+1)} = \Delta T \sum_{m=0}^{n-1} e^{(n-m)\Delta T A} [F(u_m^{(k+1)}) - F(u_m^{\text{ref}})] + \sum_{m=0}^{n-1} e^{(n-1-m)\Delta T A} [\mathcal{R}_m(u_m^{(k)}) - \mathcal{R}_m(u_m^{\text{ref}})],$$

where we recall that the residual operator  $\mathcal{R}_n$  are 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  $\bar{\alpha}$  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  $\bar{\alpha} = \frac{1}{4}$ .

**5.1. Statement of error estimates.** For the analysis, it is important to first 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  $u_n^{\text{ref}}$  respectively.

PROPOSITION 5.1. *For  $T \in (0, \infty)$ ,  $k \in \mathbb{N}_0$ ,  $q \in \mathbb{N}$ . 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  $\alpha \in [0, \min(\bar{\alpha}, \frac{1}{2})]$ . There exists  $C_{T,k,q,\alpha} \in (0, \infty)$  such that, for all  $u_0 \in D((-A)^\alpha)$  and  $\Delta T \in (0, 1)$ ,

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

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

THEOREM 5.2. *Let  $T \in (0, \infty)$ ,  $k \in \mathbb{N}_0$ ,  $q \in \mathbb{N}$ ,  $\alpha \in [0, \min(\bar{\alpha}, \frac{1}{2})]$ , and arbitrarily small  $\kappa > 0$ . For all  $u_0 \in D((-A)^\alpha)$ , 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 5.2 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 5.3, states that the estimate from Theorem 5.3 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 5.3 is motivated by the numerical experiments reported in Section 5.2, which exhibit that indeed the order of convergence is larger than  $(k+1) \min(\bar{\alpha}, \frac{1}{2})$ , namely it is equal at least equal to  $(k-1) \min(2\bar{\alpha}, \frac{1}{2}) + 2\alpha \leq k \min(2\bar{\alpha}, \frac{1}{2})$ .

THEOREM 5.3. *Let  $T \in (0, \infty)$ ,  $k \in \mathbb{N} \setminus \{1\}$ ,  $q \in \mathbb{N}$ ,  $\alpha \in (0, \min(\bar{\alpha}, \frac{1}{2}))$ , and arbitrarily small  $\kappa > 0$ . For all  $u_0 \in D((-A)^\alpha)$ , 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 Proposition 5.1 and Theorems 5.2 and 5.3 are postponed to Section 5.3.

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 5.4. *Theorem 5.2 (and Lemma 5.5–(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*

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

and

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

*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 Theorem 5.2 and 5.3.

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  $\gamma_p = \lambda_p^{\frac{1}{2}-2\bar{\alpha}}$ . 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  $J = 2^j$ ,  $j = 5, \dots, 10$ . An average over  $M = 100$  independent Monte-Carlo samples is used to approximate the expectations.

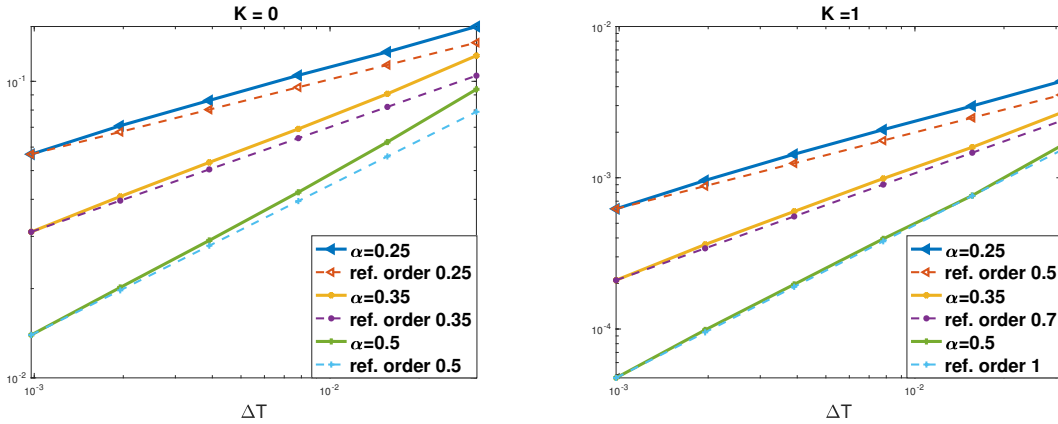


FIGURE 6. Orders of convergence of the error with respect to  $\Delta T$ , for  $K = 0$  (left) and  $K = 1$  (right), for different values of  $\alpha \in \{0.25, 0.35, 0.5\}$ , in the exponential Euler scheme case.

Figure 6 demonstrates that the orders of convergence in Theorem 5.2 are sharp when  $k = 0$  and  $k = 1$ , for different values of  $\bar{\alpha}$ .

Figure 7 then illustrates that for  $k \geq 2$ , the orders of convergence in Theorem 5.2 are not sharp for  $\bar{\alpha} = 0.25$  and  $\bar{\alpha} = 0.4$ . In fact, even the improved error estimates from Theorem 5.3 seem to be sub-optimal in the numerical experiment, especially when  $\bar{\alpha} = 0.4$  (figure on the right) – in this case  $\min(2\alpha, \frac{1}{2}) = \frac{1}{2}$ . When  $\bar{\alpha} = 0.25$  (figure on the left), the improved theoretical rate in Theorem 5.3 is closer to the observed numerical rate.

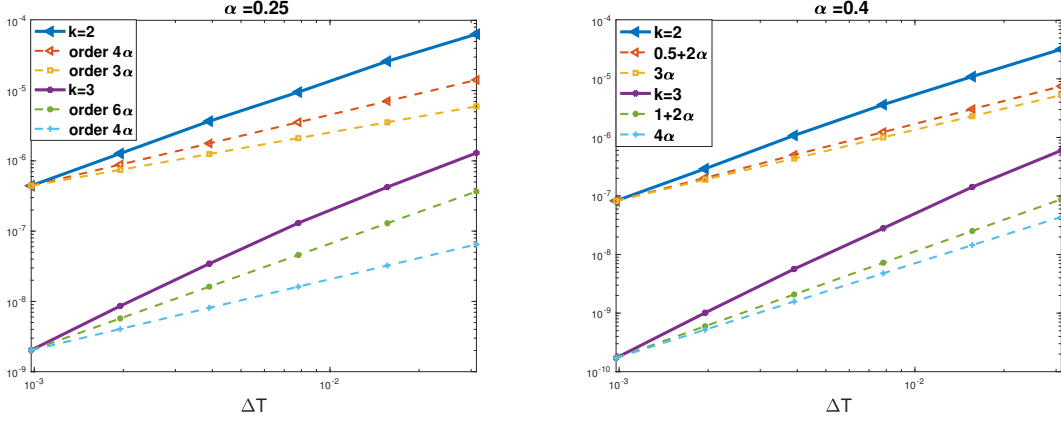


FIGURE 7. Orders of convergence of the error with respect to  $\Delta T$ , for  $\alpha = 0.25$  (left) and  $\alpha = 0.4$  (right), for different values of  $k \in \{2, 3, 4\}$ , in the exponential Euler scheme case.

**5.3. Proof of error estimates.** The objective of this section is to provide proofs for the results stated in Section 5.1. In fact, they are all based on the following Lemma concerning Lipschitz continuity properties, in appropriate norms, of the residual operators.

LEMMA 5.5. *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  $q \in \mathbb{N}$ ,  $\alpha \in (0, \min(\bar{\alpha}, \frac{1}{2}))$ ,  $\beta \in [0, \alpha)$  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  $u_1, u_2 \in D((-A)^\alpha)$ , one has*

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

The proof of Lemma 5.5 is postponed to Section 5.4.

Observe that Statement (i) (in the case  $\alpha = 0$ ) is not sufficient to exhibit a positive rate of convergence  $\gamma$  for the error  $\epsilon_n^{(k)}$ , in terms of  $\Delta T$ . Indeed, that would required 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 Theorem 5.2 and 5.3. It requires to choose 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 5.2 follows from Statement (ii) with  $\beta = 0$ , and then Theorem 5.3 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 5.5 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  $\Delta T \rightarrow 0$ . 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 5.6. *Note that the estimates stated in Lemma 5.5 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 5.5 as a conditional expectation, yield the required estimates below.*

PROOF OF PROPOSITION 5.1. First, moment bounds for the reference solution  $u_n^{\text{ref}}$  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  $\Delta T$ , 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  $\alpha \in [0, 1)$ ,

$$\begin{aligned} \|\epsilon_n^{(k+1)}\|_{q,\alpha} &\leq \Delta T \sum_{m=0}^{n-1} \|(-A)^\alpha e^{(n-m)\Delta T A}\|_{\mathcal{L}(H)} \|\epsilon_m^{(k+1)}\|_{q,0} \\ &\quad + \sum_{m=0}^{n-2} \|(-A)^\alpha e^{(n-1-m)\Delta T A}\|_{\mathcal{L}(H)} \|\mathcal{R}_m(u_m^{(k)}) - \mathcal{R}_m(u_m^{\text{ref}})\|_{q,0} \\ &\quad + \|(-A)^\alpha [\mathcal{R}_{n-1}(u_{n-1}^{(k)}) - \mathcal{R}_{n-1}(u_{n-1}^{\text{ref}})]\|_{q,0}. \end{aligned}$$

Assume first that  $\alpha = 0$ . Then the claim follows from the application of Lemma 5.5–(i), the use of the discrete Gronwall Lemma, and from the use of a recursion argument with respect to  $k$ . When  $\alpha > 0$ , it remains to apply Lemma 5.5–(i), and to apply the regularization estimate from Proposition 2.1 to conclude the proof.  $\square$

PROOF OF THEOREM 5.2. 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  $\epsilon_n^{(k)} = (u_n^{(k)} - u(t_n)) - (u_n^{\text{ref}} - u(t_n))$ , and the initialization of the parareal algorithm consists in applying the coarse integrator.

Let now  $k \geq 1$ . Owing to the expression of the error (11), applying Lemma 5.5–(ii) with  $\beta = 0$  and Hölder inequality (see Remark 5.6 for the conditioning argument), one obtains

$$\begin{aligned} \|\epsilon_n^{(k)}\|_q &\leq C \Delta T \sum_{m=0}^{n-1} \|\epsilon_m^{(k)}\|_q \\ &\quad + \sum_{m=0}^{n-2} \|(-A)^\alpha e^{(n-1-m)\Delta T A}\|_{\mathcal{L}(H)} \|(-A)^{-\alpha} [\mathcal{R}(u_m^{(k-1)}) - \mathcal{R}(u_m^{\text{ref}})]\|_q \\ &\quad + \|\mathcal{R}_{n-1}(u_{n-1}^{(k-1)}) - \mathcal{R}_{n-1}(u_{n-1}^{\text{ref}})\|_q \end{aligned}$$

$$\begin{aligned}
&\leq C\Delta T \sum_{m=0}^{n-1} \|\epsilon_m^{(k)}\|_q + C\Delta T \|\epsilon_{n-1}^{(k-1)}\|_q \\
&+ C\Delta T^{1+\alpha-\kappa} \sum_{m=0}^{n-2} \frac{1}{((n-m-1)\Delta T)^\alpha} \|\epsilon_m^{(k-1)}\|_{2q} (1 + \|u_m^{(k-1)}\|_{4q,\alpha}^2 + \|u_m^{\text{ref}}\|_{4q,\alpha}^2).
\end{aligned}$$

Due to the moment estimates of Proposition 5.1, applying 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

$$\begin{aligned}
\sup_{n\Delta T \leq T} \|\epsilon_n^{(k)}\|_q &\leq C_{T,k,q,\alpha,\kappa}(u_0) \Delta T^{k(\alpha-\kappa)} \sup_{n\Delta T \leq T} \|\epsilon_n^{(0)}\|_{2^k q} \\
&\leq C_{T,k,q,\alpha,\kappa}(u_0) \Delta T^{(k+1)(\alpha-\kappa)},
\end{aligned}$$

owing to the result when  $k = 0$ . This concludes the proof of Theorem 5.2.  $\square$

**PROOF OF THEOREM 5.3.** The proof consists of 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, using the expression of the error (11), Proposition 2.1, and the two statements of Lemma 5.5, one obtains

$$\begin{aligned}
\|\epsilon_n^{(k)}\|_{q,\alpha} &\leq \Delta T \sum_{m=0}^{n-1} \|(-A)^\alpha e^{(n-m)\Delta T A}\|_{\mathcal{L}(H)} \|\epsilon_m^{(k)}\|_q \\
&+ \sum_{m=0}^{n-2} \|(-A)^{2\alpha} e^{(n-1-m)\Delta T A}\|_{\mathcal{L}(H)} \|(-A)^{-\alpha} [\mathcal{R}_m(u_m^{(k-1)}) - \mathcal{R}_m(u_m^{\text{ref}})]\|_q \\
&+ \|(-A)^\alpha [\mathcal{R}_{n-1}(u_{n-1}^{(k-1)}) - \mathcal{R}_{n-1}(u_{n-1}^{\text{ref}})]\|_q \\
&\leq C \left( \Delta T \sum_{m=0}^{n-1} \frac{1}{((n-m)\Delta T)^\alpha} \right) \sup_{m \leq n} \|\epsilon_m^{(k)}\|_q \\
&+ C \left( \Delta T \sum_{m=0}^{n-1} \frac{1}{((n-1-m)\Delta T)^{2\alpha}} \right) \Delta T^{\alpha-\kappa} \sup_{m \leq n} \|\epsilon_m^{(k-1)}\|_q \\
&+ \Delta T^{1-\alpha} \|\epsilon_n^{(k-1)}\|_q \\
&\leq C \Delta T^{(k+1)(\alpha-\kappa)} + C \Delta T^{1-\alpha+k(\alpha-\kappa)},
\end{aligned}$$

owing to Theorem 5.2. With the assumptions  $\alpha \leq \frac{1}{2}$ , 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 5.3.

Using the expression of the error (11), Proposition 2.1, the two statements in Lemma 5.5, with  $\beta = \alpha - \kappa$ , and the moment bounds from Proposition 5.1, one obtains

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

$$\begin{aligned}
& + \sum_{m=0}^{n-2} \|(-A)^{2\alpha} e^{(n-1-m)\Delta T A} \|_{\mathcal{L}(H)} \|(-A)^{-\alpha} [\mathcal{R}_m(u_m^{(k-1)}) - \mathcal{R}_m(u_m^{\text{ref}})]\|_q \\
& + \|\mathcal{R}_{n-1}(u_{n-1}^{(k-1)}) - \mathcal{R}_{n-1}(u_{n-1}^{\text{ref}})\|_{q,\alpha} \\
& \leq C\Delta T \sum_{m=0}^{n-1} \frac{1}{((n-m)\Delta T)^\alpha} \|\epsilon_m^{(k)}\|_{q,\alpha} \\
& + C\Delta T^{\min(2\alpha, \frac{1}{2})-2\kappa} \left( \Delta T \sum_{m=0}^{n-2} \frac{1}{((n-1-m)\Delta T)^{2\alpha}} \right) \sup_{m \leq n-2} \|\epsilon_m^{(k-1)}\|_{2q,\alpha} \\
& + C\Delta T^{1-\alpha} \|\epsilon_{n-1}^{(k-1)}\|_q \\
& \leq C\Delta T \sum_{m=0}^{n-1} \frac{1}{((n-m)\Delta T)^\alpha} \|\epsilon_m^{(k)}\|_{q,\alpha} + C\Delta T^{\min(2\alpha, \frac{1}{2})-2\kappa} \sup_{m\Delta T \leq T} \|\epsilon_m^{(k-1)}\|_{2q,\alpha},
\end{aligned}$$

using 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,

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

owing to Step 1 with  $k = 1$ .

This concludes the proof of Theorem 5.3.  $\square$

**REMARK 5.7.** *The claim in Step 1 in the proof of Theorem 5.3 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 5.3 holds true only if  $k \geq 2$ .*

**5.4. Proof of Lemma 5.5.** The proof of Lemma 5.5 consists in proving bounds on the derivative  $D\mathcal{R}_n(u).h = D\mathcal{F}_n(u).h - D\mathcal{G}_n(u).h$  (recall the residual operator  $\mathcal{R}_n$  is defined by (8)). Let us introduce notation which is used below. On the one hand, due to (3), it is straightforward to compute

$$D\mathcal{G}_n(u).h = e^{\Delta T A} h + \Delta T e^{\Delta T A} DF(u).h.$$

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 formulae

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

Then a straightforward computation yields

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

Observe that the same term  $e^{\Delta T A} h$  appears in expressions of the derivatives  $D\mathcal{F}_n(u).h$  and  $D\mathcal{G}_n(u).h$ . 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 LEMMA 5.5–(i). Since  $F$  is globally Lipschitz continuous, applying the discrete Gronwall Lemma yields the following (almost sure) inequality:

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

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

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

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

Observe that in the proof above, only the Lipschitz continuity property for  $F$  was employed. In the proof below, the estimates stated in Assumption 1 are crucial.

PROOF OF LEMMA 5.5–(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:

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

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

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

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,

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

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

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

with the inequality

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

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

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

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

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

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:

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

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

$$\|(-A)^{-\alpha} \delta t \sum_{j=0}^{J-1} e^{(J-j)\delta t A} (DF(v_{n,j}) - DF(u)) \cdot \eta_{n,j}^h\|_q \leq C_{q,\alpha,\beta} \Delta T^{1+\min(\alpha+\beta+\kappa, \frac{1}{2})} |h|_\beta (1 + |u|_\alpha^2).$$

Gathering the estimates then concludes the proof.  $\square$



## 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*, volume 40 of *Lect. Notes Comput. Sci. Eng.*, pages 425–432. Springer, Berlin, 2005.
- [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(3):341–350, 1989.
- [4] P. Chartier and B. Philippe. A parallel shooting technique for solving dissipative ODEs. *Computing*, 51(3-4):209–236, 1993.
- [5] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014.
- [6] M. J. Gander. 50 years of time parallel time integration. In *Multiple shooting and time domain decomposition methods*, volume 9 of *Contrib. Math. Comput. Sci.*, pages 69–113. Springer, Cham, 2015.
- [7] M. J. Gander and E. Hairer. Nonlinear convergence analysis for the parareal algorithm. In *Domain decomposition methods in science and engineering XVII*, volume 60 of *Lect. Notes Comput. Sci. Eng.*, pages 45–56. Springer, Berlin, 2008.
- [8] M. J. Gander and E. Hairer. Analysis for parareal algorithms applied to Hamiltonian differential equations. *J. Comput. Appl. Math.*, 259(part A):2–13, 2014.
- [9] M. J. Gander and S. Vandewalle. Analysis of the parareal time-parallel time-integration method. *SIAM J. Sci. Comput.*, 29(2):556–578, 2007.
- [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*, volume 55 of *Lect. Notes Comput. Sci. Eng.*, pages 291–298. Springer, Berlin, 2007.
- [11] J Hong, X. Wang, and L. Zhang. Parareal exponential  $\theta$ -scheme for longtime simulation of stochastic Schrödinger equations with weak damping. *arXiv:1803.09188*.
- [12] G. Horton and S. Vandewalle. A space-time multigrid method for parabolic partial differential equations. *SIAM J. Sci. Comput.*, 16(4):848–864, 1995.
- [13] A. Jentzen and P. E. Kloeden. *Taylor approximations for stochastic partial differential equations*, volume 83 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011.
- [14] M. Kiehl. Parallel multiple shooting for the solution of initial value problems. *Parallel Comput.*, 20(3):275–295, 1994.
- [15] R. Kruse. *Strong and weak approximation of semilinear stochastic evolution equations*, volume 2093 of *Lecture Notes in Mathematics*. Springer, Cham, 2014.
- [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(7):661–668, 2001.
- [17] W. Liu and M. Röckner. *Stochastic partial differential equations: an introduction*. Universitext. Springer, Cham, 2015.
- [18] G. J. Lord, C. E. Powell, and T. Shardlow. *An introduction to computational stochastic PDEs*. Cambridge Texts in Applied Mathematics. Cambridge University Press, New York, 2014.
- [19] C. Lubich and A. Ostermann. Multigrid dynamic iteration for parabolic equations. *BIT*, 27(2):216–234, 1987.
- [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(4):387–392, 2002.
- [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*, volume 40 of *Lect. Notes Comput. Sci. Eng.*, pages 441–448. Springer, Berlin, 2005.
- [22] W. L. Miranker and W. Liniger. Parallel methods for the numerical integration of ordinary differential equations. *Math. Comp.*, 21:303–320, 1967.
- [23] J. Nievergelt. Parallel methods for integrating ordinary differential equations. *Comm. ACM*, 7:731–733, 1964.

- [24] L. Zhang, W. Zhou, and L. Ji. Parareal algorithms applied to stochastic differential equations with conserved quantities. *J. Comput. Math.*, 37(1):48–60, 2019.

UNIV LYON, CNRS, UNIVERSITÉ CLAUDE BERNARD LYON 1, UMR5208, INSTITUT CAMILLE JORDAN, F-69622 VILLEURBANNE, FRANCE

*E-mail address:* `brehier@math.univ-lyon1.fr`

DEPARTMENT OF MATHEMATICS, PURDUE UNIVERSITY, 150 NORTH UNIVERSITY STREET, WEST LAFAYETTE, INDIANA 47907, USA

*E-mail address:* `wang4191@purdue.edu`