SIAM J. SCI. COMPUT. © 2015 Society for Industrial and Applied Mathematics
Vol. 37, No. 2, pp. A970–A992

CONVERGENCE ANALYSIS FOR THREE PARAREAL SOLVERS*

SHU-LIN WU† AND TAO ZHOU‡

Abstract. We analyze in this paper the convergence properties of the parareal algorithm for the symmetric positive definite problem $\mathbf{u}' + A\mathbf{u} = g$. The coarse propagator $\mathcal{G}$ is fixed to the backward-Euler method and three time integrators are chosen for the $\mathcal{F}$-propagator: the trapezoidal rule, the third-order diagonal implicit Runge–Kutta (RK) (DIRK) method, and the fourth-order Gauss RK method. It is well known that the Parareal-Euler algorithm using the backward-Euler method for $\mathcal{F}$ and $\mathcal{G}$ converges rapidly, but less is known when one uses for $\mathcal{F}$ the trapezoidal rule, or the fourth-order Gauss RK method, especially when the mesh ratio $J(=\Delta T/\Delta t)$ is small. We show that for a specified $\lambda_{\max}$(the maximal eigenvalue of A or its upper bound), there exists some critical $J_{\rm cri}$ such that the parareal solvers derived from these three choices of $\mathcal{F}$ converge as fast as Parareal-Euler, provided $J \geq J_{\rm cri}$. Precisely, for $\mathcal{F}$ the trapezoidal rule and the fourth-order Gauss RK method, $J_{\rm cri}$ depends on $\Delta T, \Delta t$, and $\lambda_{\max}$ and we present concise formulas to calculate $J_{\rm cri}$, while for $\mathcal{F}$ the third-order DIRK method, $J_{\rm cri} = 4$, independently of these parameters. Numerical examples with applications in fractional PDEs and uncertainty quantification are presented to support the theoretical predictions.

Key words. parareal algorithm, SPD problems, time-dependent PDEs, trapezoidal rule, third-order DIRK method, fourth-order Gauss Runge–Kutta method

AMS subject classifications. 65R20, 45L05, 65L20, 68Q60

DOI. 10.1137/140970756

1. Introduction.

The parareal algorithm, proposed by Lions, Maday, and Turinici in [18], received considerable attention in recent years, because of its potential for parallel-in-time computation. It permits the computation of the solution later in time, before having fully accurate approximations at earlier times, while the global accuracy of the iterative process after a few iterations is comparable to that given by a sequential numerical method used on a fine discretization in time. Nowadays, this algorithm, as well as some other relevant algorithms [5, 7, 9, 10, 11, 22, 29, 21], have been used in many fields by many researchers, such as molecular-dynamics simulations [2], morphological transformation simulations [15], structural (fluid) dynamics simulations [5, 10], optimal control [6, 20, 24, 23], Hamiltonian simulations [8, 14], simulations of turbulent plasmas [26, 27], fast computation of wave equations [7] and Volterra integral equations [17], etc. For a survey of parallel-in-time algorithms, see [16].

The algorithm is defined by two time propagators, namely, $\mathcal{F}$ and $\mathcal{G}$, which are associated with the small step size Δt and the big step size ΔT. The two step sizes satisfy $\Delta T = J\Delta t$, where $J \geq 2$ is an integer. Since the backward-Euler method is stable for all possible choices of ΔT and the coarse propagator is often forced to take large step sizes, choosing for $\mathcal{G}$ the backward-Euler method makes sense. (We can also consider other implicit methods, but they are all more expensive than the backward-Euler method.) Throughout this paper, the $\mathcal{G}$-propagator is fixed to the

*Submitted to the journal's Methods and Algorithms for Scientific Computing section May 28, 2014; accepted for publication (in revised form) February 2, 2015; published electronically April 21, 2015. This work was supported by the NSF of China (11301362, 11371157, 91130003) and the NSF of Technology & Education of Sichuan Province (2013QZY01,2014JQ0035).

http://www.siam.org/journals/sisc/37-2/97075.html

†Corresponding Author. School of Science, Sichuan University of Science and Engineering, Zigong, Sichuan 643000, China (wushulin_ylp@163.com).

‡Institute of Computational Mathematics and Scientific/Engineering Computing, AMSS, the Chinese Academy of Sciences, Beijing, China (tzhou@lsec.cc.ac.cn).

backward-Euler method. In this setting, the accuracy of the converged solution and the convergence rate is determined by the $\mathcal{F}$-propagator. For the problem

$$\mathbf{u}' + A\mathbf{u} = \mathbf{g}, \ A \in \mathbb{R}^{m\times m} \text{ symmetric positive definite (SPD)}, \tag{1.1}$$

it was shown numerically in [12, 28] that the Parareal-Euler algorithm using for $\mathcal{F}$ the backward-Euler method, converges rapidly. This rapid convergence is theoretically explained in a recent work [24, Lemma 4.3], where the authors proved that the convergence factor of the Parareal-Euler algorithm is around 0.298. More recently, we proved in [30] that this result also holds for some second-order time integrators, such as $\mathcal{F}$=DIRK2 (second-order diagonal implicit Runge–Kutta (RK) method) and $\mathcal{F}$=TR/BDF2 (i.e., the `ode23tb` solver for ODEs in MATLAB). However, for $\mathcal{F}$, using the trapezoidal rule—the simplest implicit second-order time integrator, the convergence factor of the corresponding parareal algorithm, which we denote by Parareal-TR, approaches 1, when $\Delta T\lambda_{\max}$ goes to $+\infty$, where $\lambda_{\max}$ denotes the maximal eigenvalue (or its upper bound) of A. This confirms early experiments we performed for testing Parareal-TR for heat equations with a fine spatial mesh size. We obtained similar experiments for the Parareal-Gauss4 algorithm, which consists of using for $\mathcal{F}$ the fourth-order Gauss RK method. We note that Parareal-Euler, Parareal-TR, and Parareal-Gauss4, together with Parareal-DIRK3 (using for $\mathcal{F}$ the third-order DIRK method), represent the simplest implicit parareal solvers with order 1 to 4 for the converged solution.

We study in this paper the three representative parareal algorithms with order 2 to 4 in detail: Parareal-TR, Parareal-DIRK3, and Parareal-Gauss4. Our main finding is that for specified $\lambda_{\max}$ there exists some critical quantity J_{cri}, depending on whether Δt (or ΔT) is fixed or not, such that these three parareal algorithms converge as fast as the Parareal-Euler algori3thm, provided the mesh ratio J satisfies $J \geq J_{\text{cri}}$. We review the basic idea of the parareal algorithm and present some numerical illustration of the convergence factor in section 2. In section 3, we present the main results of this paper (Theorems 3.1 and 3.2), namely, the formulas for calculating J_{cri} and that the convergence factors of these three parareal algorithms are around $\frac{1}{3}$ for $J \geq J_{\text{cri}}$. The detailed proofs are given in section 4 and numerical experiments are given in section 5. We finish this paper in section 6 with some concluding remarks.

2. Review of the parareal algorithm.

In this section, we review the parareal framework and introduce the main result given by Gander and Vandewalle [12], which is the starting point of our work in this paper. For the system of ODEs

$$\begin{cases} u'(t) = f(t, u(t)), & t \in [0, T], \\ u(0) = u_0, & t = 0, \end{cases} \tag{2.1}$$

where the function f: $\mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$ is Lipschitz continuous, the parareal algorithm can be described as follows. First, the whole time interval $[0, T]$ is divided into N large time intervals $[T_n, T_{n+1}]$, $n = 0, 1, \ldots, N-1$. We suppose that all the time intervals are of uniform size, i.e., $T_{n+1} - T_n = \Delta T = \frac{T}{N}$. Second, we divide each large time interval $[T_n, T_{n+1}]$ into $J(\geq 2)$ small time-intervals $[T_{n+\frac{j\Delta T}{J}}, T_{n+\frac{(j+1)\Delta T}{J}}]$, $j = 0, 1, \ldots, J-1$. Then, two numerical propagators $\mathcal{G}$ and $\mathcal{F}$ are assigned to the coarse and fine time grids, where $\mathcal{G}$ is usually a low-order and inexpensive numerical method and $\mathcal{F}$ is usually a higher-order expensive method. We designate by the symbol $\ominus$ the time-sequential parts of the algorithm, and by the symbol $\oplus$ the time-parallel parts. Then, the parareal algorithm proposed in [18] is the following.

Algorithm 2.1. Parareal algorithm.

⊖ **Initialization**: Compute sequentially $u_{n+1}^0 = \mathcal{G}(T_n, u_n^0, \Delta T)$ with $u_0^0 = u_0$, $n = 0, 1, \ldots, N-1$;

For $k = 0, 1, \ldots$

⊕ **Step 1** On each subinterval $[T_n, T_{n+1}]$, compute $\tilde{u}_{n+\frac{j+1}{J}} = \mathcal{F}(T_{n+\frac{j}{J}}, \tilde{u}_{n+\frac{j}{J}}, \frac{\Delta T}{J})$ with initial value $\tilde{u}_n = u_n^k$, where $T_{n+\frac{j}{J}} = T_n + \frac{j\Delta T}{J}$ and $j = 0, 1, \ldots, J-1$;

⊖ **Step 2** Perform sequential corrections

$$u_{n+1}^{k+1} = \mathcal{G}\left(T_n, u_n^{k+1}, \Delta T\right) + \tilde{u}_{n+1} - \mathcal{G}\left(T_n, u_n^k, \Delta T\right),$$

where $u_0^{k+1} = u_0$, $n = 0, 1, \ldots, N-1$;

⊖ **Step 3** If $\{u_n^{k+1}\}_{n=1}^N$ satisfies the stopping criterion, terminate the iteration; otherwise go to **Step 1**.

Clearly, the argument $\tilde{u}_{n+1}$ can be written as $\tilde{u}_{n+1} = \mathcal{F}^J(T_n, u_n^k, \Delta t)$ and therefore Algorithm 2.1 can be written compactly as

$$u_{n+1}^{k+1} = \mathcal{G}\left(T_n, u_n^{k+1}, \Delta T\right) + \mathcal{F}^J\left(T_n, u_n^k, \Delta t\right) - \mathcal{G}\left(T_n, u_n^k, \Delta T\right), \tag{2.2}$$

where $\Delta t = \frac{\Delta T}{J}$ and $\mathcal{F}^J(T_n, u_n^k, \Delta t)$ denotes a value calculated by running J steps of the fine propagator $\mathcal{F}$ with initial value u_n^k and the fine step size Δt.

For the linear system of ODEs (1.1) with $A \in \mathbb{R}^{m\times m}$ an SPD matrix (which therefore can be diagonalized and all the eigenvalues are positive real numbers), the following result concerning the convergence of the parareal algorithm on sufficiently long time intervals holds for any choice of the fine propagators and can be obtained directly from the analysis given by Gander and Vandewalle [12].

Theorem 2.1 (general result given in [12]). *Let $\mathcal{F}$ be a time integrator with stability function $\mathcal{R}_f(z)$ and $\sigma(A) = \{\lambda_1, \ldots, \lambda_m\}$ be the set of eigenvalues of the matrix A in* (1.1). *Then, the errors $\{e_n^k\}$ of the parareal algorithm at the kth iteration using the backward-Euler method as the coarse propagator satisfy*

$$\sup_n \|V e_n^k\|_\infty \leq \rho^k \sup_n \|V e_n^0\|_\infty, \ \rho = \max_{\lambda\in\sigma(A)} \mathcal{K}(\Delta T\lambda, J), \tag{2.3}$$

where ρ is the convergence factor of the parareal algorithm, $k \geq 1$ is the iteration index, and $V \in \mathbb{R}^{m\times m}$ consists of the eigenvectors of A (i.e., $V^{-1}AV = diag(\lambda_1, \ldots, \lambda_m)$). The argument $\mathcal{K}$, which we call the convergence factor corresponding to a single eigenvalue (or in short "contraction factor" hereafter), is defined by

$$\mathcal{K}(z, J) = \frac{\left|\mathcal{R}_f^J(z/J) - \frac{1}{1+z}\right|}{1 - \left|\frac{1}{1+z}\right|}. \tag{2.4}$$

Candidates for the $\mathcal{F}$-propagator studied in this paper are the trapezoidal rule, the third-order DIRK method, and the fourth-order Gauss RK method, which, together

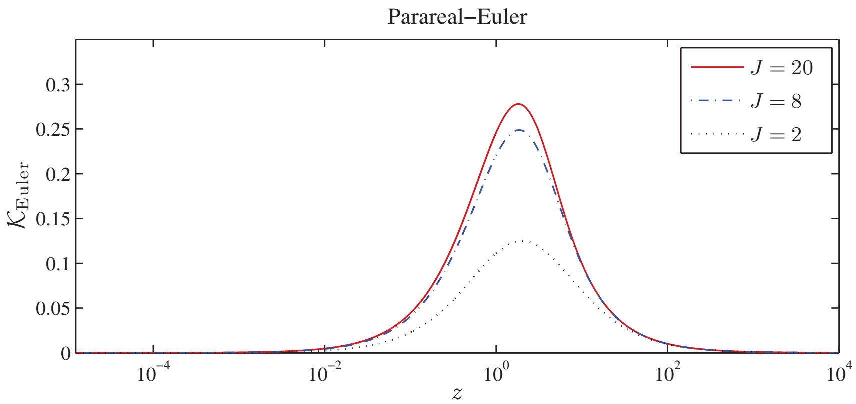

FIG. 1. *The contraction factor $\mathcal{K}_{\mathrm{Euler}}(z, J)$ as a function of z.*

with their stability functions (denoted by $\mathcal{R}_{\mathrm{TR}}$, $\mathcal{R}_{\mathrm{DIRK3}}$, and $\mathcal{R}_{\mathrm{Gauss4}}$), are given by

(2.5)

- trapezoidal rule (i.e., second-order Gauss RK); $\begin{array}{c|c} \frac{1}{2} & \frac{1}{2} \\ \hline & 1 \end{array}$, $\mathcal{R}_{\mathrm{TR}}(z) := \dfrac{1 - \frac{z}{2}}{1 + \frac{z}{2}}$;
- third-order DIRK : $\begin{array}{r|cc} \gamma & \gamma & 0 \\ 1-\gamma & \frac{-1}{\sqrt{3}} & \gamma \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$, $\mathcal{R}_{\mathrm{DIRK3}}(z) := \dfrac{\sqrt{3} + z - \gamma z^2}{\sqrt{3}(z\gamma + 1)^2}$, $\gamma = \dfrac{3 + \sqrt{3}}{6}$;
- fourth-order Gauss RK : $\begin{array}{c|cc} \frac{3-\sqrt{3}}{6} & \frac{1}{4} & \frac{3-2\sqrt{3}}{12} \\ \frac{3+\sqrt{3}}{6} & \frac{3+2\sqrt{3}}{12} & \frac{1}{4} \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$, $\mathcal{R}_{\mathrm{Gauss4}}(z) := \dfrac{1 - \frac{z}{2} + \frac{z^2}{12}}{1 + \frac{z}{2} + \frac{z^2}{12}}$.

The parareal algorithms using the backward-Euler method as the $\mathcal{G}$-propagator and the above three methods as the $\mathcal{F}$-propagator are denoted by Parareal-TR, Parareal-DIRK3, and Parareal-Gauss4, respectively. The contraction factor $\mathcal{K}$ associated with these three algorithms are denoted by $\mathcal{K}_{\mathrm{TR}}$, $\mathcal{K}_{\mathrm{DIRK3}}$, and $\mathcal{K}_{\mathrm{Gauss4}}$.

For given J and $z_{\max} := \Delta T \lambda_{\max}$, the convergence factor $\rho(J)$ defined by $\rho(J) := \max_{z \in [0, z_{\max}]} \mathcal{K}(z, J)$ determines the behavior of the parareal algorithm over long time intervals; the smaller $\rho(J)$ is, the faster the algorithm converges. The quantity $\lambda_{\max}$ denotes the maximal eigenvalue (or an upper bound), of the coefficient matrix A in (1.1). Qualitative analysis of the convergence factor is difficult, even for a rigorous mathematical proof of $\rho(J) < 1$. This is principally because the contraction factor $\mathcal{K}(z, J)$ is very complicated. For the Parareal-Euler algorithm, a rigorous analysis can be found in a recent paper by Mathew, Sarkis, and Schaerer [24], where the authors proved that

$$\max_{z \in [0, z_{\max}]} \mathcal{K}_{\mathrm{Euler}}(z, J) \approx 0.298 (< 1/3) \quad \forall z_{\max} > 0, \tag{2.6}$$

where $\mathcal{K}_{\mathrm{Euler}}(z, J)$ denotes the contraction factor of the Parareal-Euler algorithm, defined by inserting $\mathcal{R}^J(z/J) = (\frac{1}{1+z/J})^J$ into (2.4). More recently, we proved in [30] that the Parareal-TR/BDF2 and Parareal-DIRK2 algorithms, using the TR/BDF2 and second-order DIRK methods as the $\mathcal{F}$-propagator, also satisfy (2.6). In Figure 1 we show an illustration for the Parareal-Euler algorithm (the plot for Parareal-TR/BDF2 and Parareal-DIRK2 looks similar).

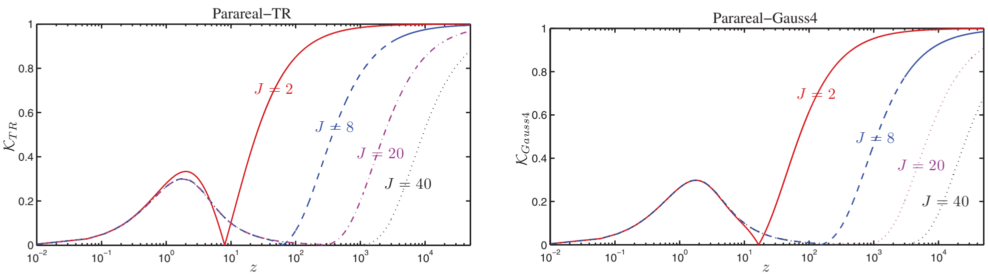

FIG. 2. *The contraction factor $\mathcal{K}(z, J)$ as a function of z, for different choices of J. Left: Parareal-TR ($\mathcal{F}$=trapezoidal rule); Right: Parareal-Gauss4($\mathcal{F}$ = fourth-order Gauss).*

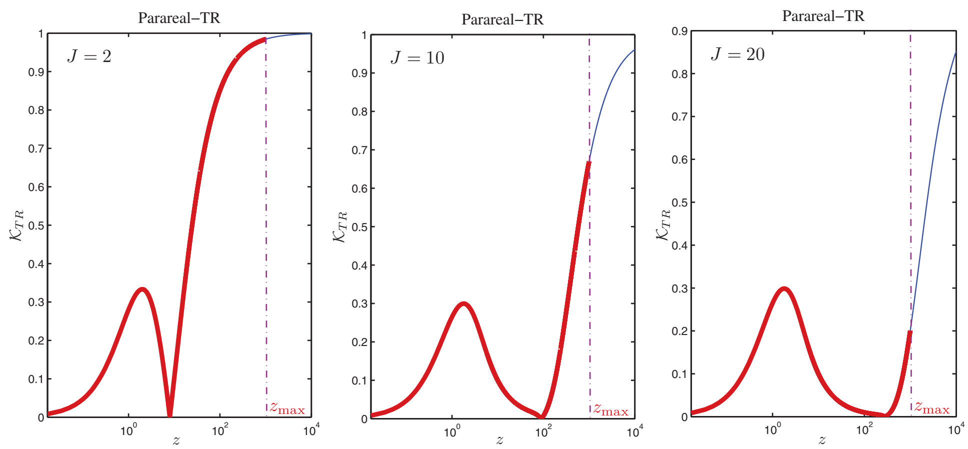

FIG. 3. *For given $z_{\max}$, the convergence factor $\rho(J) := \max_{z\in[0,z_{\max}]} \mathcal{K}^l_{\mathrm{TR}}(z, J)$ tends to a value around $\frac{1}{3}$, as J increases. Here, $z_{\max}$ is a fixed quantity, independently of J; for $z_{\max}$ being connected to J: $z_{\max} = J\tilde{z}_{\max}$ with $\tilde{z}_{\max}$ a fixed quantity, see Figure* 7.

Unfortunately, such a uniform result does not hold for the Parareal-TR and Parareal-Gauss4 algorithms, because

$$\lim_{z\to+\infty} \mathcal{R}_{\mathrm{TR,Gauss4}}(z) = 1 \Rightarrow \lim_{z\to+\infty} \frac{\left|\mathcal{R}^J_{\mathrm{TR,Gauss4}}\left(\frac{z}{J}\right) - \frac{1}{1+z}\right|}{1-\frac{1}{1+z}} = 1, \tag{2.7}$$

which can be clearly seen in Figure 2. However, in a concrete computation, $z = \Delta T\lambda$ cannot be arbitrarily large, and therefore it is still possible to obtain an ideal parareal solver by choosing a properly chosen large mesh ratio J. More precisely, for some specified $z_{\max} := \Delta T\lambda_{\max}$, the convergence factor $\rho(J) := \max_{z\in[0,z_{\max}]} \mathcal{K}(z, J)$ tends to a constant around $\frac{1}{3}$, when J becomes larger and larger; this is illustrated in Figure 3 for the Parareal-TR algorithm (the plots for Parareal-Gauss4 look similar).

For the Parareal-DIRK3 algorithm, we have

$$\lim_{z\to+\infty} \mathcal{R}_{\mathrm{DIRK3}}(z) = (-0.73205)^J \Rightarrow \lim_{z\to+\infty} \frac{\left|\mathcal{R}^J_{\mathrm{DIRK3}}\left(\frac{z}{J}\right) - \frac{1}{1+z}\right|}{1-\frac{1}{1+z}} = 0.73205^J, \tag{2.8}$$

and this implies that $z \to +\infty$ will not lead to stagnation. We see in Figure 4 that

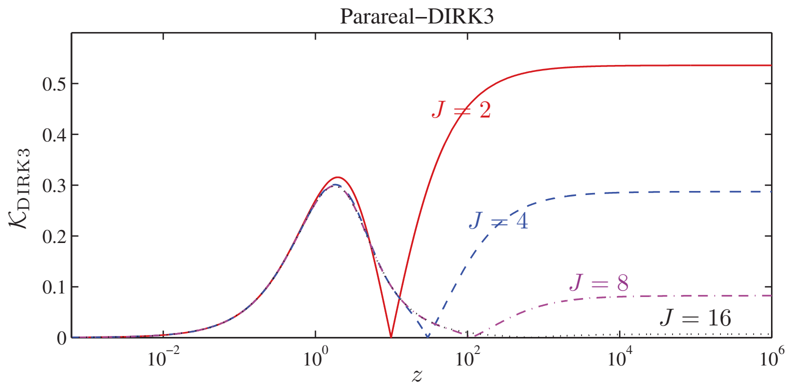

FIG. 4. *The contraction factor* $\mathcal{K}_{\mathrm{DIRK3}}(z, J)$ *as a function of* z.

for a relatively large mesh ratio J the global maximum of $\mathcal{K}_{\mathrm{DIRK3}}(z, J)$ is also around $\frac{1}{3}$, i.e., $\max_{z\in[0,+\infty)} \mathcal{K}_{\mathrm{DIRK3}}(z, J) \approx \frac{1}{3}$.

In summary, for the Parareal-TR, Parareal-DIRK3, and Parareal-Gauss4 algorithms, it seems that for given $z_{\max}$ there exists some critical J_{cri}, such that

$$\max_{z\in[0,z_{\max}]} \mathcal{K}(z, J) \le \frac{1}{3} \quad \text{if } J \ge J_{\mathrm{cri}}. \tag{2.9}$$

In what follows, we prove this and present explicit formulas for calculating J_{cri}. Our analysis is related to the dependence of $z_{\max}$ on J. First, if $z_{\max} = \Delta T\lambda_{\max}$ is independent of J, i.e., the coarse mesh size ΔT is fixed, the quantity J_{cri} is denoted by $J^*_{\min}$ (see Theorem 3.1). Second, if $z_{\max} = J\Delta t\lambda_{\max}$ depends on J, i.e., the fine mesh size Δt is fixed (in other words, the anticipated accuracy of the converged solution is fixed), J_{cri} is denoted by $\tilde{J}^*_{\min}$ (see Theorem 3.2).

3. Main results. In this section, we present the main results of this paper. The proofs are given in section 4.

THEOREM 3.1 (ΔT is fixed, independently of J). *Let $z_{\max} = \Delta T\lambda_{\max} > 0$ be a fixed quantity. Then, there exists some positive quantity $J^*_{\min}$ such that*

$$\max_{z\in[0,z_{\max}]} \mathcal{K}_{\mathrm{TR,DIRK3,Gauss4}}(z, J) \le \frac{1}{3} \ \textit{if } J \textit{ is even and } J \ge J^*_{\min}, \tag{3.1}$$

*where $\mathcal{K}_{\mathrm{TR,\ DIRK3,\ Gauss4}}$ are the contraction factors of the Parareal-TR, Parareal-DIRK*3*, and Parareal-Gauss*4 *algorithms, which are defined by* (2.4) *with stability functions $\mathcal{R}_{\mathrm{TR,\ DIRK3,\ Gauss4}}$ given in* (2.5).

The lower bound $J^*_{\min}$ for the three parareal algorithms is given by

$$\begin{aligned} &\text{Parareal-DIRK3}: \quad J^*_{\min} = 4 \ \forall z_{\max} \in \mathbb{R}^+; \\ &\text{Parareal-TR}: \quad J^*_{\min} = \begin{cases} 2 & \text{if } z_{\max} \le z^*_{\mathrm{TR}}, \\ J^*_{\mathrm{TR}}(z_{\max}) & \text{otherwise}; \end{cases} \\ &\text{Parareal-Gauss4}: \quad J^*_{\min} = \begin{cases} 2 & \text{if } z_{\max} \le z^*_{\mathrm{Gauss4}}, \\ J^*_{\mathrm{Gauss4}}(z_{\max}) & \text{otherwise}, \end{cases} \end{aligned} \tag{3.2}$$

where $J^*_{\mathrm{TR}}(z_{\max}) > 2$ and $J^*_{\mathrm{Gauss4}}(z_{\max}) > 2$ depend on $z_{\max}$ and are the unique positive roots of the two nonlinear equations

$$\underbrace{\left|\mathcal{R}_{\rm TR}\left(\frac{z_{\max}}{J}\right)\right|^J = \frac{3+z_{\max}}{3(1+z_{\max})}}_{\text{root}=J^*_{\rm TR}(z_{\max})}, \quad \underbrace{\left|\mathcal{R}_{\rm Gauss4}\left(\frac{z_{\max}}{J}\right)\right|^J = \frac{3+z_{\max}}{3(1+z_{\max})}}_{\text{root}=J^*_{\rm Gauss4}(z_{\max})}. \tag{3.3}$$

The quantities $z^*_{\rm TR}$ and $z^*_{\rm Gauss4}$ are the unique positive roots of $\mathcal{K}_{\rm TR}(z,2)=\frac{1}{3}$ and $\mathcal{K}_{\rm Gauss4}(z,2)=\frac{1}{3}$, i.e., $z^*_{\rm TR}=16.48528$ and $z^*_{\rm Gauss4}=45.43065$.[1]

Theorem 3.1 deals with the case that the number of coarse time intervals, i.e., $N=\frac{T}{\Delta T}$, is fixed. In this case, increasing J reduces the fine mesh size Δt, since $\Delta t=\frac{\Delta T}{J}$. We also need to consider the situation that the anticipated accuracy of the converged solution is fixed, i.e., Δt is fixed; in this case, the quantity $z_{\max}=\Delta T\lambda_{\max}=J\Delta t\lambda_{\max}$ increases as J increases. The next theorem shows that in this case there also exists some quantity $\tilde{J}^*_{\min}$ such that the convergence factors of the three parareal algorithms can be bounded from above by $\frac{1}{3}$, provided $J\geq\tilde{J}^*_{\min}$.

THEOREM 3.2 (Δt is fixed, independently of J). *Let $\tilde{z}_{\max}=\Delta t\lambda_{\max}$ be a fixed quantity. Then, there exists some positive quantity $\tilde{J}^*_{\min}$ such that*

$$\max_{z\in[0,J\tilde{z}_{\max}]}\mathcal{K}_{\rm TR,DIRK3,Gauss4}(z,J)\leq\frac{1}{3} \text{ if } J \text{ is even and } J\geq\tilde{J}^*_{\min}. \tag{3.4}$$

The definition of $\tilde{J}_{\min}$ is similar to that of $J^*_{\min}$ given by (3.2), but with $z^*_{\rm TR,\ Gauss4}$ and $J^*_{\rm TR,\ Gauss4}$ being replaced by $\tilde{z}^*_{\rm TR,\ Gauss4}$ and $\tilde{J}^*_{\rm TR,\ Gauss4}$, where the $\tilde{z}^*_{\rm TR,\ Gauss4}$ are the unique positive roots of $\mathcal{K}_{\rm TR,\ Gauss4}(2\tilde{z},2)=\frac{1}{3}$, i.e., $\tilde{z}^*_{\rm TR,\ Gauss4}=\frac{z^*_{\rm TR,\ Gauss4}}{2}$, and the $\tilde{J}^*_{\rm TR,\ Gauss4}$ are the unique positive roots of the equations

$$\underbrace{\left|\mathcal{R}_{\rm TR}\left(\tilde{z}_{\max}\right)\right|^J = \frac{3+J\tilde{z}_{\max}}{3(1+J\tilde{z}_{\max})}}_{\text{root}=\tilde{J}^*_{\rm TR}(\tilde{z}_{\max})}, \quad \underbrace{\left|\mathcal{R}_{\rm Gauss4}\left(\tilde{z}_{\max}\right)\right|^J = \frac{3+J\tilde{z}_{\max}}{3(1+J\tilde{z}_{\max})}}_{\text{root}=\tilde{J}^*_{\rm Gauss4}(\tilde{z}_{\max})}. \tag{3.5}$$

Remark 3.1 (about $J^*_{\min}$ and $\tilde{J}^*_{\min}$). The constraints $J\geq J^*_{\min}$ in (3.1) and $J\geq\tilde{J}^*_{\min}$ in (3.4) are equivalent to $\Delta T\geq\Delta tJ^*_{\min}$ and $\Delta T\geq\Delta t\tilde{J}^*_{\min}$. For prescribed fine mesh size, this equivalence is useful for choosing the coarse mesh size in a practical computation. For the Parareal-DIRK3 algorithm, Theorems 3.1 and 3.2 actually assert that for SPD problem (1.1) this algorithm always converges as fast as the Parareal-Euler algorithm, since $J\geq 4$ is a trivial restriction. For the other two parareal algorithms, Parareal-TR and Parareal-Gauss4, the influence of J on the convergence rates cannot be neglected, since for these two algorithms the lower bound $J^*_{\min}$ (resp., $\tilde{J}^*_{\min}$) increases as $z_{\max}=\Delta T\lambda_{\max}$ (resp., $\tilde{z}_{\max}=\Delta t\lambda_{\max}$) increases; see Figure 5. For given $z_{\max}$ (resp., $\tilde{z}_{\max}$), because of the uniqueness of the roots of the two nonlinear equations in (3.3) (resp., (3.5)), the quantity $J^*_{\min}$ (resp., $\tilde{J}^*_{\min}$) can be easily computed for example by the single-variable nonlinear solver *fzero in MATLAB.*

We now point out the relation between the work in this paper and previous studies. We show in Remark 3.2 that the results given by Theorems 3.1 and 3.2 can be easily generalized to time-periodic problems. In Remark 3.3, we show that the work in this

[1]The unique existence of $z^*_{\rm TR,\ Gauss4}$ can be proved easily and we omit the details.

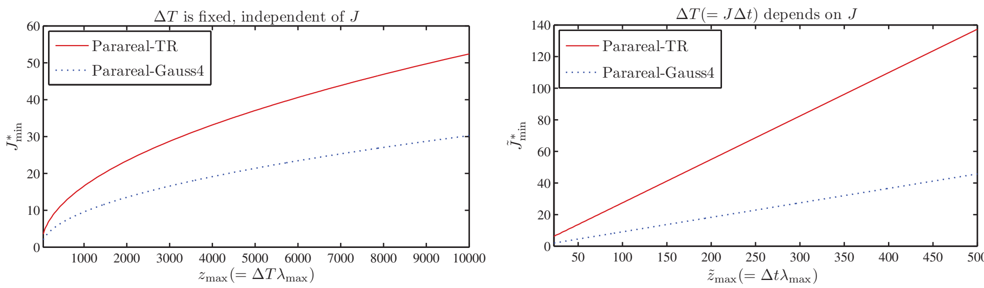

FIG. 5. *For the Parareal-TR and Parareal-Gauss4 algorithms, the quantities $J^*_{\min}$ (left panel) and $\tilde{J}^*_{\min}$ (right panel) are given as a function of $z_{\max}(=\Delta T\lambda_{\max})$ and $\tilde{z}_{\max}(=\Delta t\lambda_{\max})$.*

paper is closely related to previous studies with $\mathcal{F}=e^{-A\Delta T}$, i.e., the fine propagator is chosen as the exact time integrator.

Remark 3.2 (application to time-periodic SPD problems). Theorems 3.1 and 3.2 can be easily applied to time-periodic SPD problems of the form

$$\begin{cases}\mathbf{u}'(t)+A\mathbf{u}=g,\ t\in[0,T],\\ \mathbf{u}(0)=\mathbf{u}(T).\end{cases} \tag{3.6}$$

In [13], the so-called PP-PC (periodic parareal algorithm with periodic coarse problem) and PP-IC (periodic parareal algorithm with initial value coarse problem) algorithms are proposed for (3.6). Here, we are interested in PP-PC, because it converges faster than PP-IC. Similarly to (2.2), the compact form of the PP-PC algorithm reads

$$\begin{aligned} u_0^{k+1}&=\mathcal{G}\left(T_{N-1},u_{N-1}^{k+1},\Delta T\right)+\mathcal{F}^J\left(T_{N-1},u_{N-1}^{k},\Delta t\right)-\mathcal{G}\left(T_{N-1},u_{N-1}^{k},\Delta T\right),\\ u_{n+1}^{k+1}&=\mathcal{G}\left(T_n,u_n^{k+1},\Delta T\right)+\mathcal{F}^J\left(T_n,u_n^{k},\Delta t\right)-\mathcal{G}\left(T_n,u_n^{k},\Delta T\right),\end{aligned} \tag{3.7}$$

where $n=0,1,\ldots,N-2$. In (3.7), the fine propagator is applied to nonperiodic problems and the time periodicity on the coarse time grids will be recovered upon convergence. Define $\{e_n^k:=\mathbf{u}_n^k-\mathbf{u}_n\}_{n=0}^N$. Then, if we use the backward-Euler method as the $\mathcal{G}$-propagator and a concrete time integrator as the $\mathcal{F}$-propagator, following the analysis in [13, section 3], the errors $\{e_n^k\}$ and $\{e_n^{k+1}\}$ also satisfy (2.3)–(2.4). Hence, Theorems 3.1 and 3.2 are also applicable to the PP-PC algorithm (3.7).

Remark 3.3 (relevance to $\mathcal{F}=e^{-A\Delta T}$). In the analysis of the parareal algorithm, one often uses the exact propagator $\mathcal{F}=e^{-A\Delta T}$ and this significantly simplifies the convergence analysis; see for example [1, 4, 12, 13, 19]. If we use the backward-Euler method as the $\mathcal{G}$-propagator, this simplification yields

$$\max_{z\in[0,z_{\max}]}\mathcal{K}(z,J)=\max_{z\in[0,z_{\max}]}\frac{\left|e^{-z}-\frac{1}{1+z}\right|}{1-\frac{1}{1+z}}\le\frac{1}{3}\ (\textit{independent of } J). \tag{3.8}$$

In a practical numerical computation, the fine propagator $\mathcal{F}$ is defined by a concrete time integrator, instead of the exact propagator $\mathcal{F}=e^{-A\Delta T}$. Therefore, (3.8) does not always predict the real convergence rate very well; in fact, (3.8) coincides with the practical situation when we use for $\mathcal{F}$ the backward-Euler method [24], the second-order DIRK and TR/BDF2 methods [30], while for $\mathcal{F}$ using the trapezoidal rule and the fourth-order Gauss RK method, we now know that this is not always the case (see

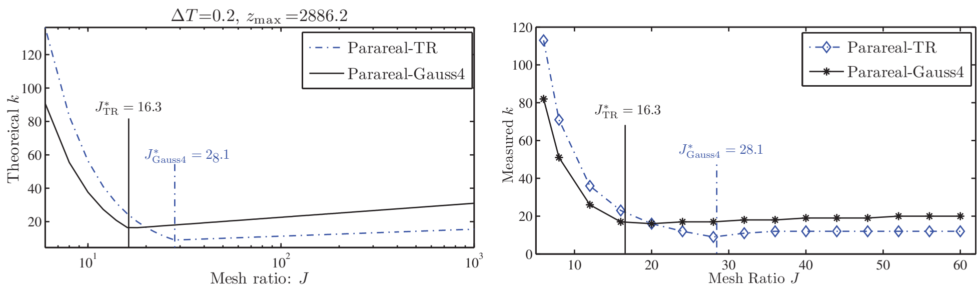

FIG. 6. *With the discretization/problem parameters shown on the top of Figure* 8*, we show the number of iterations required to reach the accuracy of the $\mathcal{F}$-propagator, i.e., $(\Delta T/J)^p$ for the Parareal-TR ($p = 2$) and Parareal-Gauss4 ($p = 4$) algorithms. Left: number of iterations by theoretical predication; Right: number of iterations measured in a numerical experiment.*

Figure 2). For these two algorithms Parareal-TR and Parareal-Gauss4, for specified $z_{\max}(= \Delta T\lambda_{\max}$ or $J\Delta t\lambda_{\max})$, (3.8) still coincides with the practical situation if we impose a restriction on the mesh ratio J: $J \geq J^*_{\min}$ (or $J \geq \tilde{J}^*_{\min}$).

At the end of this section, we address the following question: *for ΔT fixed, how does the number of iterations required to match the accuracy of $\mathcal{F}$ with the fine mesh size Δt, i.e., $\mathcal{O}(\Delta t^p)$, behave as J increases?* Here, p denotes the order of accuracy of $\mathcal{F}$. For Parareal-Euler, Parareal-TR/BDF2, Parareal-DIRK2, and Parareal-DIRK3, it is easy to answer this question, because the number of iterations is

$$k = \mathcal{O}\left(\frac{\log{(\Delta T/J)^p}}{\log \rho}\right) = \mathcal{O}(\log J) \ \left(\text{because } \rho \approx \frac{1}{3} \forall J, \Delta T, z_{\max}\right), \tag{3.9}$$

and therefore k increases as J increases. For the Parareal-TR and Parareal-Gauss4 algorithms, for given $z_{\max}$ and ΔT, Theorem 3.1 implies that $\rho \approx \frac{1}{3}$ for $J \geq J^*_{\min}$ and therefore the required number of iterations also behaves like $k = \mathcal{O}(\log J)$. For $J < J^*_{\min}$, the proof of Theorem 3.1 given in section 4 shows that $\rho = \mathcal{K}_{\text{TR, Gauss4}}(z_{\max}, J)$ and ρ decreases as J increases, and thus it is hard to say whether $k = \mathcal{O}(\log J/\log(1/\rho))$ is still an increasing function of J or not. We now study this problem numerically. For the Riesz fractional PDEs used in subsection 5.1 with the discretization/problem parameters shown on the top of Figure 8 (i.e., $\Delta T = 0.2$ and $z_{\max} = 2886.2$), we plot in Figure 6 on the left the quantity $\frac{\log(\Delta T/J)^p}{\log \mathcal{K}_{\text{TR, Gauss4}}(z_{\max}, J)}$ as a function of J, where p=2 and $p = 4$ for $\mathcal{F}$ using the trapezoidal rule and the fourth-order Gauss RK method, respectively. We see that for $J < J^*_{\min}$ the quantity k decreases as J increases. This conclusion is further confirmed by the results shown in Figure 6 on the right, where we plot the measured number of iterations required to reach the tolerances $(\frac{\Delta T}{J})^2$ and $(\frac{\Delta T}{J})^4$ for the two parareal algorithms. We have also tested many other choices for the discretization/problem parameters and this conclusion always seems to hold.

4. The proofs. For the algorithms Parareal-Euler [24], Parareal-TR/BDF2, and Parareal-DIRK2 [30], it holds that $\max_{z\in[0,\infty)} \mathcal{K}(z, J) \leq \frac{1}{3}$(for all $J \geq 2$ and J even). The proof of this result depends on the L-stability of the corresponding $\mathcal{F}$-propagator (i.e., $\lim_{z\to+\infty} \mathcal{R}_f(z) = 0$) and it is relatively simple compared to the proof of Theorems 3.1 and 3.2. The techniques proposed in [24, 30] cannot be simply generalized to analyze the Parareal-TR, Parareal-DIRK3, and Parareal-Gauss4 algorithms. For each algorithm, the proof is different and is shown in three subsections. For Parareal-TR and Parareal-Gauss4, the proof of Theorem 3.2 is closely connected to that of Theorem 3.1 and we only present the details for the Parareal-TR algorithm in subsection 4.1. For the Parareal-DIRK3 algorithm, we present a uniform proof for Theorems 3.1 and 3.2. Since Theorems 3.1 and 3.2 only concern even mesh ratios J, it is convenient to rewrite the contraction factors $\mathcal{K}_{\text{TR, DIRK3, Gauss4}}(z, J)$ given by (2.4) as

$$\mathcal{K}_{\text{TR, DIRK3, Gauss4}}(z, J) = \frac{\left|\left|\mathcal{R}_{\text{TR, DIRK3, Gauss4}}\left(\frac{z}{J}\right)\right|^J - \frac{1}{1+z}\right|}{1 - \left|\frac{1}{1+z}\right|}, \tag{4.1}$$

where $\mathcal{R}_{\text{TR, DIRK3, Gauss4}}$ are the stability functions of the trapezoidal rule, the third-order DIRK, and fourth-order Gauss RK methods given by (2.5). By rewriting the contraction factors $\mathcal{K}_{\text{TR, DIRK3, Gauss4}}$ in this form, the quantity J does not have to be an integer and throughout this section we assume that $J \geq 2$ is a real number.

4.1. The Parareal-TR algorithm. To prove Theorems 3.1 and 3.2 for the Parareal-TR algorithm, the following lemma is useful.

LEMMA 4.1. *For any $J \in [2, \infty)$, the following equation has a unique root $s_0 \geq 2$:*

$$F_{\text{TR}}(s, J) := \underbrace{J \log \left|\frac{s-2}{s+2}\right| + \log(1 + Js) = 0}_{\text{root}=s_0 \geq 2}. \tag{4.2}$$

Proof. For $J \geq 2$ and $s \in (0, 2)$, we have

$$\underbrace{\log\left(\frac{2-s}{2+s}\right) + s < 0}_{\text{can be verified easily for } s \in (0,2)} \quad \Rightarrow \log\left(\frac{2-s}{2+s}\right) + \frac{s}{1+Js} < 0 \Rightarrow \partial_J F_{\text{TR}} < 0,$$

and this implies $F_{\text{TR}}(s, J) < \log(\frac{2-s}{2+s}) + \log(1+s) < 0$. Hence, for any $J \geq 2$, it holds that $F_{\text{TR}}(s, J) < 0$ for all $s \in (0, 2)$. For $s > 2$, F_{TR} is an increasing function of z; this, together with $\lim_{s \to 2^+} F_{\text{TR}}(s, J) = -\infty$ and $\lim_{s \to +\infty} F_{\text{TR}}(s, J) = +\infty$, implies that F_{TR} has a unique root $s_0 > 2$. ∎

Based on Lemma 4.1, the proof of Theorem 3.1 for the Parareal-TR algorithm consists of the following three lemmas.

LEMMA 4.2. *For any $J \geq 2$, it holds that $\partial_z \mathcal{K}_{\text{TR}}(z, J) > 0$ for $z > Js_0$.*

Proof. To prove this claim, we first show that the contraction factor $\mathcal{K}_{\text{TR}}(z, J)$ given by (4.1) can be rewritten as

$$\mathcal{K}_{\text{TR}}(z, J) = \begin{cases} \frac{\frac{1}{1+z} - \left|\mathcal{R}_{\text{TR}}\left(\frac{z}{J}\right)\right|^J}{1 - \frac{1}{1+z}} & \text{if } z \in (0, Js_0], \\ \frac{\left|\mathcal{R}_{\text{TR}}\left(\frac{z}{J}\right)\right|^J - \frac{1}{1+z}}{1 - \frac{1}{1+z}} & \text{if } z > Js_0. \end{cases} \tag{4.3}$$

Let $s = \frac{z}{J}$. Then, it holds that $|\mathcal{R}_{\text{TR}}(\frac{z}{J})|^J - \frac{1}{1+z} = |\mathcal{R}_{\text{TR}}(s)|^J - \frac{1}{1+Js}$. Clearly,

$$\text{sign}\left(\left|\mathcal{R}_{\text{TR}}\left(\frac{z}{J}\right)\right|^J - \frac{1}{1+z}\right) = \text{sign}\left(F_{\text{TR}}(s, J)\right),$$

where F_{TR} is defined by (4.2). For $J \in [2, \infty)$, in the proof of Lemma 4.1 we have proved that $F_{\text{TR}}(s, J) \leq 0$ if $s \leq s_0$ and $F_{\text{TR}}(s, J) > 0$ if $s > s_0$. This gives (4.3).

Since $Js_0 > 2J$ and for $z > 2J$ it holds that $\mathcal{R}_{\rm TR}(z/J) < 0$, from (4.3) we have $\mathcal{K}_{\rm TR}(z,J) = \frac{(-\mathcal{R}_{\rm TR}(\frac{z}{J}))^J - \frac{1}{1+z}}{1-\frac{1}{1+z}}$ for $z > Js_0$. Hence, for $z > Js_0$ we have

$$z^2\partial_z\mathcal{K}_{\rm TR}(z,J) = \left(-\mathcal{R}_{\rm TR}\left(\frac{z}{J}\right)\right)^{J-1}\frac{4z(1+z)}{\left(2+\frac{z}{J}\right)^2} + \left[1-\left(-\mathcal{R}_{\rm TR}\left(\frac{z}{J}\right)\right)^J\right] > 0. \tag{4.4}$$

Hence, it holds that $\partial_z\mathcal{K}_{\rm TR}(z,J) > 0$ for $z > Js_0$. ∎

LEMMA 4.3. *For any $J \geq 2$, it holds that* $\max_{z\in[0,Js_0]}\mathcal{K}_{\rm TR}(z,J) \leq \frac{1}{3}$.

Proof. To prove this claim, we first set an ansatz $\max_{z\in[0,Js_0]}\mathcal{K}_{\rm TR}(z,J) \leq \theta$ with $\theta \geq \frac{1}{4}$ and then we prove that θ can reach $\frac{1}{3}$. Clearly,

$$\mathcal{K}_{\rm TR}(z,J) \leq \theta \Leftrightarrow \frac{1-\theta z}{1+z} \leq \left|\mathcal{R}_{\rm TR}\left(\frac{z}{J}\right)\right|^J \leq \frac{1+\theta z}{1+z}\ \forall z\in[0,Js_0]. \tag{4.5}$$

We have already proved $\frac{1}{1+z} \geq |\mathcal{R}_{\rm TR}(\frac{z}{J})|^J$ for $z\in[0,Js_0]$; this, together with $\frac{1+\theta z}{1+z} \geq \frac{1}{1+z}$, implies that $|\mathcal{R}_{\rm TR}(\frac{z}{J})|^J \leq \frac{1+\theta z}{1+z}$ holds for all $\theta > 0$. Therefore,

$$\mathcal{K}_{\rm TR}(z,J) \leq \theta \Leftrightarrow \frac{1-\theta z}{1+z} \leq \left|\mathcal{R}_{\rm TR}\left(\frac{z}{J}\right)\right|^J\ \ \forall z\in[0,Js_0]. \tag{4.6}$$

We next prove that the minimal value of θ, which guarantees this inequality, is $\theta_{\min} = \frac{1}{3}$. By the assumption $\theta \geq \frac{1}{4}$, it suffices to prove (4.6) for $z\in[0,4]$. For $z\in[0,4]$ and $J \geq 2$, using the second result in Lemma 4.5 given below yields $|\mathcal{R}_{\rm TR}(\frac{z}{J})|^J \geq \mathcal{R}^2_{\rm TR}(\frac{z}{2})$. Hence, it suffices to require $\mathcal{R}^2_{\rm TR}(\frac{z}{2}) \geq \frac{1-\theta z}{1+z}$ to guarantee (4.6). We have

$$\mathcal{R}^2_{\rm TR}\left(\frac{z}{2}\right) \geq \frac{1-\theta z}{1+z} \Leftrightarrow \theta \geq \max_{z\in[0,4]}\frac{1-(1+z)\mathcal{R}^2_{\rm TR}\left(\frac{z}{2}\right)}{z} = \frac{1}{3}, \tag{4.7}$$

and this gives $\theta_{\min} = \frac{1}{3}$ and we finally arrive at $\max_{z\in[0,Js_0]}\mathcal{K}_{\rm TR}(z,J) \leq \frac{1}{3}$. ∎

LEMMA 4.4. *For $J\in[2,\infty)$, $\mathcal{K}_{\rm TR}(z,J) = \frac{1}{3}$ has a unique root $z_0(J) > Js_0$ and $z_0(J)$ is an increasing function of J.*

Proof. We have $\mathcal{K}_{\rm TR}(Js_0,J) = 0$ and from Lemma 4.2 we also have $\partial_z\mathcal{K}_{\rm TR}(z,J) > 0$ for $z > Js_0$. These, together with $\lim_{z\to+\infty}\mathcal{K}_{\rm TR}(z,J) = 1$, imply the unique existence of the root of $\mathcal{K}_{\rm TR}(z,J) = \frac{1}{3}$ for $z\in[Js_0,+\infty)$. It remains to prove that $z_0(J)$ is an increasing function of J. Let $J_2 > J_1 \geq 2$ be two real numbers and the roots of $\mathcal{K}_{\rm TR}(z,J_1) = \frac{1}{3}$ and $\mathcal{K}_{\rm TR}(z,J_2) = \frac{1}{3}$ be $z_0(J_1)(> J_1s_0)$ and $z_0(J_2)(> J_2s_0)$, respectively. We suppose $z_0(J_2) < z_0(J_1)$. Then, from Lemma 4.2, we have $\mathcal{K}_{\rm TR}(z_0(J_2),J_2) < \mathcal{K}_{\rm TR}(z_0(J_1),J_2)$, i.e.,

$$\frac{1}{3} = \frac{\left|\mathcal{R}_{\rm TR}\left(\frac{z_0(J_2)}{J_2}\right)\right|^{J_2} - \frac{1}{1+z_0(J_2)}}{1-\frac{1}{1+z_0(J_2)}} < \frac{\left|\mathcal{R}_{\rm TR}\left(\frac{z_0(J_1)}{J_2}\right)\right|^{J_2} - \frac{1}{1+z_0(J_1)}}{1-\frac{1}{1+z_0(J_1)}}. \tag{4.8}$$

Since $z_0(J_1) > z_0(J_2) > J_2s_0 > 2J_2$, by using the first result in Lemma 4.5 we have

$$\left|\mathcal{R}_{\rm TR}\left(\frac{z_0(J_1)}{J_2}\right)\right|^{J_2} < \left|\mathcal{R}_{\rm TR}\left(\frac{z_0(J_1)}{J_1}\right)\right|^{J_1}.$$

This, together with (4.8), gives

$$\frac{1}{3} < \frac{\left|\mathcal{R}_{\rm TR}\left(\frac{z_0(J_1)}{J_1}\right)\right|^{J_1} - \frac{1}{1+z_0(J_1)}}{1-\frac{1}{1+z_0(J_1)}} = \mathcal{K}_{\rm TR}(z_0(J_1),J_1). \tag{4.9}$$

Clearly, this contradicts the assumption $\mathcal{K}_{\mathrm{TR}}(z_0(J_1), J_1) = \frac{1}{3}$. The case $z_0(J_2) = z_0(J_1)$ can also be excluded, since $J_1 \neq J_2$ and $\mathcal{K}_{\mathrm{TR}}(z, J)$ is an increasing function of z, when $z \geq Js_0$. Hence, there is only one possibility left: $z_0(J_2) > z_0(J_1)$. ◻

Now, from Lemmas 4.2 and 4.3 we have

$$\max_{z\in[0,z_{\max}]} \mathcal{K}_{\mathrm{TR}}(z, J) \leq \max\left\{\frac{1}{3}, \mathcal{K}_{\mathrm{TR}}(z_{\max}, J)\right\}. \tag{4.10}$$

Then, for the Parareal-TR algorithm, the conclusion stated in Theorem 3.1 follows from Lemma 4.4, plus a simple logic deduction.

The following lemma, which reveals a special property of the stability function of the trapezoidal rule, plays a central role in the aforementioned analysis. For the other two parareal algorithms, Parareal-DIRK3 and Parareal-Gauss4, similar lemmas will be presented and proved in subsections 4.2 and 4.3.

LEMMA 4.5. *Let $\mathcal{R}_{\mathrm{TR}}(z) = \frac{1-z/2}{1+z/2}$ be the stability function of the trapezoidal rule and $J^* \geq 2$. Then, for any real numbers $J_1, J_2 \geq 2$ it holds that*

1. *for $z > 2J^*$, $|\mathcal{R}_{\mathrm{TR}}(\frac{z}{J_2})|^{J_2} < |\mathcal{R}_{\mathrm{TR}}(\frac{z}{J_1})|^{J_1}$ $\forall J_1 < J_2 \leq J^*$;*
2. *for $z \in (0, 2J^*]$, $|\mathcal{R}_{\mathrm{TR}}(\frac{z}{J_2})|^{J_2} > |\mathcal{R}_{\mathrm{TR}}(\frac{z}{J_1})|^{J_1}$ $\forall J_2 > J_1 \geq J^*$.*

Proof. For the first claim, since $J^* \geq J_2 > J_1$ and $z > 2J^*$, we have

$$\left|\mathcal{R}_{\mathrm{TR}}\left(\frac{z}{J_2}\right)\right| < \left|\mathcal{R}_{\mathrm{TR}}\left(\frac{z}{J_1}\right)\right| < 1 \Rightarrow \left|\mathcal{R}_{\mathrm{TR}}\left(\frac{z}{J_2}\right)\right|^{J_1} < \left|\mathcal{R}_{\mathrm{TR}}\left(\frac{z}{J_1}\right)\right|^{J_1} < 1,$$

which gives $|\mathcal{R}_{\mathrm{TR}}(\frac{z}{J_2})|^{J_2} < |\mathcal{R}_{\mathrm{TR}}(\frac{z}{J_1})|^{J_1}$.

To prove the second claim, we note that $|\mathcal{R}_{\mathrm{TR}}(\frac{z}{J})| = \mathcal{R}_{\mathrm{TR}}(\frac{z}{J})$ for all $z \in (0, 2J^*]$ and $J \geq J^*$. Moreover, a partial derivative of $\mathcal{R}^J_{\mathrm{TR}}(\frac{z}{J})$ with respect to J leads to

$$\partial_J\left[\mathcal{R}^J_{\mathrm{TR}}\left(\frac{z}{J}\right)\right] = \mathcal{R}^J_{\mathrm{TR}}\left(\frac{z}{J}\right)\left[\log \mathcal{R}_{\mathrm{TR}}\left(\frac{z}{J}\right) - \frac{z}{J}\frac{\partial_z \mathcal{R}_{\mathrm{TR}}\left(\frac{z}{J}\right)}{\mathcal{R}_{\mathrm{TR}}\left(\frac{z}{J}\right)}\right]. \tag{4.11}$$

Let $s = \frac{z}{J}$, $r(s) = \mathcal{R}_{\mathrm{TR}}(s)$, and $R(s) = \log r(s) - s\frac{r'(s)}{r(s)}$. Clearly, $\partial_J[\mathcal{R}^J_{\mathrm{TR}}(\frac{z}{J})] = r^J(s)R(s)$. Since $z \in (0, 2J^*]$ and $J > J^*$, we have $s \in (0, 2)$ and then we know $r(s) = \mathcal{R}_{\mathrm{TR}}(\frac{z}{J}) > 0$. Hence, proving the second claim is equivalent to proving $R(s) > 0$ for $s \in (0, 2)$. To this end, by noticing $R(s) = \log r(s) - s\frac{d \log r(s)}{ds}$ we have

$$R'(s) = \frac{d\log r(s)}{ds} - \frac{d\log r(s)}{ds} - s\frac{d^2\log r(s)}{ds^2} = -s\frac{d\left(\frac{r'(s)}{r(s)}\right)}{ds} = \frac{8s^2}{4-s^2} > 0. \tag{4.12}$$

This, together with $R(0) = 0$, gives $R(s) > 0$ for $s \in (0, 2)$. ◻

It remains to prove Theorem 3.2. For fixed $\tilde{z}_{\max} := \Delta t\lambda_{\max}$, the quantity $z_{\max} = J\tilde{z}_{\max}$ increases as J increases. For given J, by using Lemmas 4.2 and 4.3, we have

$$\mathcal{K}_{\mathrm{TR}}(z, J) \leq \begin{cases} \frac{1}{3}, & z \in [0, Js_0], \\ \mathcal{K}_{\mathrm{TR}}(z_{\max}, J) = \frac{(-\mathcal{R}_{\mathrm{TR}}(\tilde{z}_{\max}))^J - \frac{1}{1+J\tilde{z}_{\max}}}{1-\frac{1}{1+J\tilde{z}_{\max}}}, & z \in [Js_0, J\tilde{z}_{\max}]. \end{cases} \tag{4.13}$$

For $\tilde{z}_{\max} > 2$ and any two real numbers J_1 and J_2, we claim that

$$\frac{\left(-\mathcal{R}_{\mathrm{TR}}\left(\tilde{z}_{\max}\right)\right)^{J_2} - \frac{1}{1+J_2\tilde{z}_{\max}}}{1 - \frac{1}{1+J_2\tilde{z}_{\max}}} < \frac{\left(-\mathcal{R}_{\mathrm{TR}}\left(\tilde{z}_{\max}\right)\right)^{J_1} - \frac{1}{1+J_1\tilde{z}_{\max}}}{1 - \frac{1}{1+J_1\tilde{z}_{\max}}}, \tag{4.14}$$

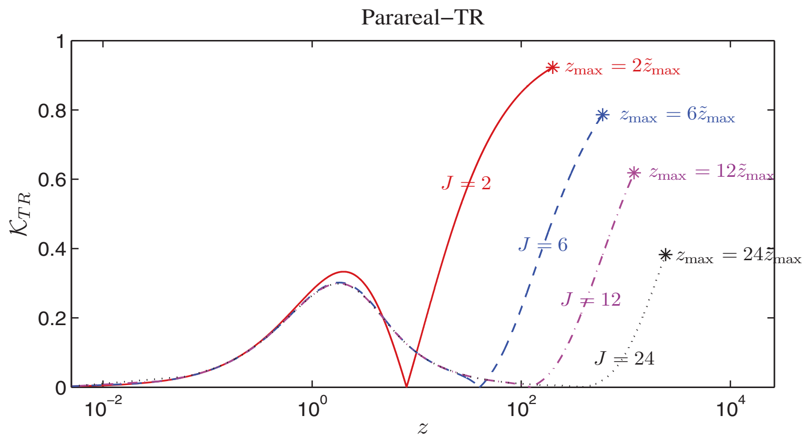

FIG. 7. *For $\tilde{z}_{\max} = \Delta t\lambda_{\max}$ being a fixed quantity, $\mathcal{K}_{\rm TR}(z_{\max}, J)$ is a decreasing function of J.*

provided $J_2 > J_1$ and $(1 + J_{1,2}\tilde{z}_{\max})[-\mathcal{R}_{\rm TR}(\tilde{z}_{\max})]^{J_{1,2}} \ge 1$. Let $H(J) = \frac{(1+J\tilde{z})r^J-1}{J\tilde{z}}$ with $r \in (0,1)$ and $\tilde{z} > 0$. Then, for J satisfying $(1 + J\tilde{z})r^J \ge 1$ we have

$$\frac{dH}{dJ} = \frac{r^J(1 + J\tilde{z})\log(r^J) - r^J + 1}{J^2\tilde{z}} \le \frac{\log(r^J) - r^J + 1}{J^2\tilde{z}}, \tag{4.15}$$

where we have used $r \in (0,1)$ ($\Rightarrow \log(r^J) < 0$). Let $\tilde{r} = r^J \in (0,1)$ and $H_1(\tilde{r}) = \log(\tilde{r}) - \tilde{r} + 1$. Then, it is easy to get $\frac{dH_1}{d\tilde{r}} > 0$; this, together with $H_1(1) = 0$, implies $H_1(\tilde{r}) = \log(r^J) - r^J + 1 < 0$. Hence, $\frac{dH}{dJ} < 0$, which gives (4.14). (An illustration of the monotonic property (4.14) is shown in Figure 7.) Now, the monotonic property (4.14), together with (4.10) and a simple logic deduction, proves Theorem 3.2.

Clearly, the proof of Theorem 3.2 given above holds, in fact, for the Parareal-Gauss4 algorithm, because we have not yet used specific properties of the stability function $\mathcal{R}_f(z)$ of the $\mathcal{F}$-propagator. (The proof only requires $|\mathcal{R}_f(z)| < 1(\forall z > 0)$, which is true for these two $\mathcal{F}$-propagators, since they are A-stable time integrators.)

4.2. The Parareal-DIRK3 algorithm. To prove Theorems 3.1 and 3.2 for the Parareal-DIRK3 algorithm, we need the following lemma, which reveals a special property of the stability function of the third-order DIRK method.

LEMMA 4.6. *Let $\mathcal{R}_{\rm DIRK3}(z) = \frac{\sqrt{3}+z-\gamma z^2}{\sqrt{3}(z\gamma+1)^2}$ and $\gamma = \frac{3+\sqrt{3}}{6}$ and $z^* = \frac{1+\sqrt{1+4\sqrt{3}\gamma}}{2\gamma} =$ 2.24583. Then, for $J^* \ge 2$ and $J_2 > J_1 \ge J^*$, it holds that*

$$\left|\mathcal{R}_{\rm DIRK3}\left(\frac{z}{J_2}\right)\right|^{J_2} > \left|\mathcal{R}_{\rm DIRK3}\left(\frac{z}{J_1}\right)\right|^{J_1} \quad \forall z \in (0, z^*J^*]. \tag{4.16}$$

Proof. For $z \in [0, z^*J^*]$ and $J \ge J^*$, a routine calculation yields $|\mathcal{R}_{\rm DIRK3}(\frac{z}{J})| = \mathcal{R}_{\rm DIRK3}(\frac{z}{J})$. Let $s = \frac{z}{J}$, $r(s) = \mathcal{R}_{\rm DIRK3}(s)$, and $R(s) = \log r(s) - s\frac{r'(s)}{r(s)}$. Then, for $z \in (0, z^*J^*]$ and $J \ge J^*$, it holds that $s \in (0, 2.24583)$ and $\text{sign}(\partial_J \mathcal{R}^J_{\rm DIRK3}(\frac{z}{J})) = \text{sign}(R(s))$. Similarly to (4.12), a routine calculation yields

$$R'(s) = \frac{3\gamma^2 s^3(2\gamma s + 2\sqrt{3}\gamma - 1)}{[(1+\gamma s)(\sqrt{3} + s - \gamma s^2)]^2} > 0 \ \forall s \in (0, z^*).$$

This, together with $R(0) = 0$, gives $\partial_J \mathcal{R}^J_{\rm DIRK3}(\frac{z}{J}) > 0$ for $z \in (0, z^*J^*]$. $\square$

We now prove Theorems 3.1 and 3.2 for the Parareal-DIRK3 algorithm.

Proof. Similarly to (4.3), the quantity $\mathcal{K}_{\mathrm{DIRK3}}$ given by (4.1) can be written as

$$\mathcal{K}_{\mathrm{DIRK3}}(z,J)=\begin{cases}\frac{\frac{1}{1+z}-\left|\mathcal{R}_{\mathrm{DIRK3}}\left(\frac{z}{J}\right)\right|^J}{1-\frac{1}{1+z}}, & \text{if } z\in(0,Js_0],\\ \frac{\left|\mathcal{R}_{\mathrm{DIRK3}}\left(\frac{z}{J}\right)\right|^J-\frac{1}{1+z}}{1-\frac{1}{1+z}} & \text{if } z>Js_0,\end{cases} \tag{4.17}$$

where $s_0 > 2.24583$ is the unique positive root of the equation[2]

$$F_{\mathrm{DIRK3}}(s,J) := \underbrace{J\log\left|\frac{\sqrt{3}+s-\gamma s^2}{\sqrt{3}(s\gamma+1)^2}\right| + \log(1+Js) = 0}_{\text{root}=s_0>2.24583}. \tag{4.18}$$

For $J \geq 2$ and $z > Js_0$, we have $|\mathcal{R}_{\mathrm{DIRK3}}(\frac{z}{J})| = -\mathcal{R}_{\mathrm{DIRK3}}(\frac{z}{J})$ and therefore

$$\begin{aligned} z^2\partial_z\mathcal{K}_{\mathrm{DIRK3}}(z,J) &= \left[-\mathcal{R}_{\mathrm{DIRK3}}\left(\frac{z}{J}\right)\right]^{J-1}\frac{z(1+z)\left(2\sqrt{3}+3\gamma\frac{z}{J}-1\right)}{\sqrt{3}\left(1+\frac{z}{J}\gamma\right)^3}\\ &\quad+\left[1-\left(-\mathcal{R}_{\mathrm{DIRK3}}\left(\frac{z}{J}\right)\right)^J\right]>0, \end{aligned}$$

which gives

$$\partial_z\mathcal{K}_{\mathrm{DIRK3}}(z,J) > 0 \ \forall z > Js_0. \tag{4.19}$$

For $z \in [0, Js_0]$, we claim

$$\max_{z\in[0,Js_0]}\mathcal{K}_{\mathrm{DIRK3}}(z,J) \leq \frac{1}{3}. \tag{4.20}$$

Using Lemma 4.6, the proof of (4.20) is the same as that of Lemma 4.3 and the only difference lies in replacing $\mathcal{R}^2_{\mathrm{TR}}(z/2)$ by $\mathcal{R}^2_{\mathrm{DIRK3}}(z/2)$ in (4.7), which yields

$$\mathcal{R}^2_{\mathrm{DIRK3}}\left(\frac{z}{2}\right) \geq \frac{1-\theta z}{1+z} \Leftrightarrow \theta \geq \max_{z\in[0,4]}\frac{1-(1+z)\mathcal{R}^2_{\mathrm{DIRK3}}\left(\frac{z}{2}\right)}{z} = 0.31554.$$

Combining (4.19) and (4.20), we have

$$\begin{aligned}\max_{z\in[0,z_{\max}]}\mathcal{K}_{\mathrm{DIRK3}}(z,J) &\leq \max\left\{\frac{1}{3},\mathcal{K}_{\mathrm{DIRK3}}(z_{\max},J)\right\}\\ &\leq \max\left\{\frac{1}{3},\lim_{z_{\max}\to+\infty}\mathcal{K}_{\mathrm{DIRK3}}(z_{\max},J)\right\}.\end{aligned}$$

Then, by using (2.8), we have

$$\max_{z\in[0,z_{\max}]}\mathcal{K}_{\mathrm{DIRK3}}(z,J) \leq \max\left\{\frac{1}{3}, 0.73205^J\right\} \ \forall z_{\max} > 0. \tag{4.21}$$

Hence, for $J \geq 4$ it holds that $\max_{z\in[0,z_{\max}]}\mathcal{K}_{\mathrm{DIRK3}}(z,J) \leq \frac{1}{3}\ \forall z_{\max} > 0$. ∎

[2]The proof of the unique existence of the root is similar to that of Lemma 4.1.

4.3. The Parareal-Gauss4 algorithm. For the fourth-order Gauss RK method, the stability function $\mathcal{R}_{\rm Gauss4}(z) = \frac{1-\frac{z}{2}+\frac{z^2}{12}}{1+\frac{z}{2}+\frac{z^2}{12}}$ satisfies $\mathcal{R}_{\rm Gauss4}(z) > 0$ $\forall z \geq 0$. Hence, we have $|\mathcal{R}_{\rm Gauss4}(z)| = \mathcal{R}_{\rm Gauss4}(z)$ for all $z \geq 0$.

LEMMA 4.7. *Let* $\mathcal{R}_{\rm Gauss4}(z) = \frac{1-z/2+z^2/12}{1+z/2+z^2/12}$. *Then, for* $J_2 > J_1 \geq 2$*, it holds that*

$$\mathcal{R}^{J_2}_{\rm Gauss4}\left(\frac{z}{J_2}\right) < \mathcal{R}^{J_1}_{\rm Gauss4}\left(\frac{z}{J_1}\right) \ \forall z > 0. \tag{4.22}$$

Proof. Let $s = \frac{z}{J}$, $r(s) = \mathcal{R}_{\rm Gauss4}(s)$, and $R(s) = \log r(s) - s\frac{r'(s)}{r(s)}$. Then, for $z > 0$ and $J \geq 2$, we have

$$\operatorname{sign}\left(\partial_J \mathcal{R}^J_{\rm Gauss4}\left(\frac{z}{J}\right)\right) = \operatorname{sign}\left(R(s)\right), \ R'(s) = \frac{24s^4(s^2-24)}{\left[\left(1-6z+z^2\right)\left(1+6z+z^2\right)\right]^2}.$$

The second equality implies that $\max_{s\geq 0} R(s) = \max\{R(0), \lim_{s\to+\infty} R(s)\} = 0$, since $r(0) = \lim_{s\to+\infty} r(s) = 1$, $r'(0) = -1$, and $\lim_{s\to+\infty} r'(s) = 0$. Hence, it holds that $\partial_J \mathcal{R}^J_{\rm Gauss4} < 0$ for any $J \geq 2$; this proves (4.22). ∎

Similarly to the trapezoidal rule and the third-order DIRK method, this lemma reveals a special property of the stability function of the fourth-order Gauss RK method. We now start the proof of Theorem 3.1 for the Parareal-Gauss4 algorithm. The proof of Theorem 3.2 is similar to that of Parareal-TR given in subsection 4.1.

Proof. Similarly to (4.3), the contraction factor $\mathcal{K}_{\rm Gauss4}$ can be rewritten as

$$\mathcal{K}_{\rm Gauss4}(z, J) = \begin{cases} \frac{\frac{1}{1+z} - \mathcal{R}^J_{\rm DIRK3}\left(\frac{z}{J}\right)}{1-\frac{1}{1+z}} & \text{if } z \in (0, Js_0], \\ \frac{\mathcal{R}^J_{\rm Gauss4}\left(\frac{z}{J}\right) - \frac{1}{1+z}}{1-\frac{1}{1+z}} & \text{if } z > Js_0, \end{cases} \tag{4.23}$$

where $s_0 > 4$ is the unique positive root of the equation

$$F_{\rm Gauss4}(s, J) := \underbrace{J \log\left(\frac{1-s/2+s^2/12}{1+s/2+s^2/12}\right) + \log(1+Js) = 0}_{\text{root}=s_0>4}. \tag{4.24}$$

The unique existence of such a positive root can be shown like this: first, a routine calculation yields $\operatorname{sign}(\partial_s F_{\rm Gauss4}(s, J)) = \operatorname{sign}(s^3 + 12Js^2 - 144)$; then, as s varies from 0 to $+\infty$, we first meet $\partial_s F_{\rm Gauss4} < 0$, then $\partial_s F_{\rm Gauss4} = 0$, and then $\partial_s F_{\rm Gauss4} > 0$; and finally, the sign of $\partial_s F_{\rm Gauss4}(s, J)$, together with $F_{\rm Gauss4}(0, J) = 0$ and $F_{\rm Gauss4}(+\infty, J) = +\infty$, implies that $F_{\rm Gauss4}(s, J)$ has a unique positive root. The claim $s_0 > 4$ holds, because $\mathcal{K}_{\rm Gauss4}(4, J) = J \log\frac{1}{13} + \log(1 + 4J) < 0$ $(\forall J \geq 2)$.

Similarly to (4.4), since $s_0 > 4 (\Rightarrow J^2 s_0^2 > 12J^2)$, it holds for all $z \geq Js_0$ that

$$z^2 \partial_z \mathcal{K}_{\rm Gauss4}(z, J) = \mathcal{R}^{J-1}_{\rm Gauss4}\left(\frac{z}{J}\right) \frac{z(1+z)\left(\frac{z^2}{12J^2} - 1\right)}{\left(1 + \frac{z}{2J} + \frac{z^2}{12J^2}\right)^2} + \left[1 - \mathcal{R}^J_{\rm Gauss4}\left(\frac{z}{J}\right)\right] > 0.$$

Hence, we have

$$\partial_z \mathcal{K}_{\rm Gauss4}(z, J) > 0 \ \forall z > Js_0. \tag{4.25}$$

We next claim that

$$\max_{z\in[0,Js_0]} \mathcal{K}_{\text{Gauss4}}(z,J) \le \frac{1}{3}. \tag{4.26}$$

The proof of (4.26) is slightly different from that of (4.20) and we now give the details. Assuming that $\max_{z\in[0,Js_0]} \mathcal{K}_{\text{Gauss4}}(z,J) \le \theta$ with $\theta \ge \frac{1}{4}$, we have

$$\mathcal{K}_{\text{Gauss4}}(z,J) \le \theta \Leftrightarrow \frac{1-\theta z}{1+z} \le \mathcal{R}^J_{\text{Gauss4}}\left(\frac{z}{J}\right) \ \forall z \in [0, Js_0]. \tag{4.27}$$

By the assumption $\theta \ge \frac{1}{4}$, it is sufficient to prove the second inequality in (4.27) for $z \in [0,4]$. By using Lemma 4.7, we have

$$\mathcal{R}^J_{\text{Gauss4}}\left(\frac{z}{J}\right) \ge \lim_{J\to+\infty} \mathcal{R}^J_{\text{Gauss}}\left(\frac{z}{J}\right) = \lim_{J\to+\infty}\left(1-\frac{z}{J}\mathcal{I}(J)\right)^J, \tag{4.28}$$

where $\mathcal{I}(J) = [1+z/(2J)+z^2/(12J^2)]^{-1}$. Since $\lim_{J\to+\infty}\mathcal{I}(J)=1$, it holds that

$$\mathcal{R}^J_{\text{Gauss4}}\left(\frac{z}{J}\right) \ge \lim_{J\to+\infty} \mathcal{R}^J_{\text{Gauss}}\left(\frac{z}{J}\right) = e^{-z}. \tag{4.29}$$

So, it suffices to require $e^{-z} \ge \frac{1-\theta z}{1+z}$, i.e., $\theta \ge \max_{z\in[0,4]} \frac{1-(1+z)e^{-z}}{z} = 0.2984$, to guarantee the second inequality in (4.27). This gives (4.26). Now, from (4.25)–(4.26) we have

$$\max_{z\in[0,z_{\max}]} \mathcal{K}_{\text{Gauss4}}(z,J) \le \max\left\{\frac{1}{3}, \mathcal{K}_{\text{Gauss4}}(z_{\max},J)\right\}. \tag{4.30}$$

Since $\lim_{z\to+\infty}\mathcal{K}_{\text{Gauss4}}(z,J)=1$ and $\mathcal{K}_{\text{Gauss4}}(Js_0,J)=0$, by using (4.25) we know that $\mathcal{K}_{\text{Gauss4}}(z,J)=\frac{1}{3}$ has a unique root $z_0(J) > Js_0$. Because of the monotonicity of $\mathcal{R}^J_{\text{Gauss4}}(z/J)$ with respect to J (see Lemma 4.7), it can be shown that $z_0(J)$ is an increasing function of J. (The proof of this claim is the same as that of Lemma 4.4.) Hence, by using (4.30) and a simple logic deduction, the proof of Theorem 3.1 concerning the Parareal-Gauss4 algorithm is completed. ∎

5. Numerical results. In this section, we provide numerical results to validate Theorems 3.1 and 3.2. We test the performance of four parareal algorithms, Parareal-Euler, Parareal-TR, Parareal-DIRK3, and Parareal-Gauss4, for two problems: Riesz fractional PDEs and Galerkin approximations of the stochastic diffusion equations. The Parareal-Euler algorithm is used as a reference and we compare the other three algorithms to it. In each experiment, the initial iterate is chosen randomly and the iteration stops when the error between the iterate and the target solution satisfies

$$\max_n \left\|\mathbf{u}_n^k - \mathbf{u}_n\right\|_\infty \le 10^{-12}. \tag{5.1}$$

5.1. Riesz fractional PDEs. We first consider the Riesz fractional PDE

$$\begin{cases} \frac{\partial u(x,t)}{\partial t} = D\frac{\partial^\alpha u(x,t)}{\partial |x|^\alpha} + \sin(xt), & (x,t)\in(0,1)\times(0,T),\\ u(x,0)=0, & x\in(0,1),\\ u(0,t)=u(1,t)=0, & t\in(0,T), \end{cases} \tag{5.2}$$

where $\alpha \in (1,2]$, $D>0$ is a constant number, and $\frac{\partial^\alpha}{\partial|x|^\alpha}$ is the Riesz fractional derivative operator defined by

$$\frac{\partial^\alpha v(x)}{\partial|x|^\alpha} = -\frac{1}{2\cos(\alpha\pi/2)\Gamma(2-\alpha)}\frac{d^2}{dx^2}\int_{-\infty}^{\infty}|x-\tau|^{1-\alpha}v(\tau)d\tau. \tag{5.3}$$

Let Δx be the spatial mesh size, $m = \frac{1}{\Delta x} - 1$ be the number of spatial grid points located in the interior of $[0, 1]$, and $x_j = j\Delta x$, $j = 1, 2, \ldots, m$. Then, applying the fractional centered difference formula to (5.2) yields the system of ODEs [25, 3]

$$\mathbf{u}'(t) + A\mathbf{u}(t) = \begin{bmatrix} \sin(x_1 t) \\ \sin(x_2 t) \\ \vdots \\ \sin(x_m t) \end{bmatrix}, \; A = \frac{1}{2\Delta x^\alpha} \begin{bmatrix} \omega_0 & \omega_1 & \omega_2 & \ldots & \omega_{m-1} \\ \omega_1 & \omega_0 & \omega_1 & \ddots & \vdots \\ \omega_2 & \omega_1 & \omega_0 & \ddots & \omega_2 \\ \vdots & \ddots & \ddots & \ddots & \omega_1 \\ \omega_{m-1} & \ldots & \omega_2 & \omega_1 & \omega_0 \end{bmatrix}, \tag{5.4}$$

where

$$\omega_0 = \frac{\Gamma(1+\alpha)}{\Gamma^2(1+\alpha/2)}, \quad \omega_{j+1} = \left(1 - \frac{\alpha+1}{\alpha/2 + j + 1}\right)\omega_j, j = 1, 2, \ldots \tag{5.5}$$

Let λ be an eigenvalue of A. Then, by Gerschgorin's disk theorem, we have

$$0 < \lambda < \omega_0 \Delta x^{-\alpha} \quad \text{(see Theorem 4.1 in [3]).} \tag{5.6}$$

With the upper bound given in (5.6), the critical quantities $J^*_{\min}$ and $\tilde{J}^*_{\min}$, which guarantee rapid convergence of the Parareal-TR and Parareal-Gauss4 algorithms, can be calculated according to Theorems 3.1 and 3.2. In Figure 8, we show the measured number of iterations required to satisfy the stopping criterion (5.1) for different choices of the mesh ratio J, in two situations: $\Delta T = 0.2$ independent of J (top), and ΔT connected to J (bottom)—$\Delta T = J\Delta t$ with Δt being fixed to 0.01. (For the bottom panel, the number of coarse intervals, i.e., $N = \frac{T}{\Delta T}$, is fixed at $N = 250$.) We see that for both the Parareal-TR and Parareal-Gauss4 algorithms, if J is larger than the corresponding $J^*_{\min}$ and $\tilde{J}^*_{\min}$, the convergence rates are the same as that of the Parareal-Euler algorithm. These numerical results coincide with theoretical predictions implied by Theorems 3.1 and 3.2.

For the Parareal-DIRK3 algorithm, both Theorems 3.1 and 3.2 imply that if $J \geq 4$ the convergence rate is the same as that of the Parareal-Euler algorithm. If $J = 2$, from Figure 4 we know that Parareal-Euler converges faster than Parareal-DIRK3. These theoretical predications are validated in Figure 9. (In this figure, we consider the case that ΔT is fixed to 0.1, independently of J; for the case $\Delta T = J\Delta t$ with Δt being fixed, the corresponding convergence plot looks similar.) For the Parareal-DIRK3 algorithm applied to problem (5.4)–(5.5) with the three mesh ratios J used in Figure 9, the convergence factor $\rho := \max_{z\in[0,z_{\max}]} \mathcal{K}_{\text{DIRK3}}(z, J)$ (with $z_{\max} = \Delta T \lambda_{\max}$) reads $\rho = 0.535$, $\rho = 0.3$, and $\rho = 0.298$. The results shown in Figure 10 show that the convergence factor predicts very well the measured convergence rate, where we plot the measured error and the error predicted by ρ, as $\rho^k \max_n \|\mathbf{u}_n^0 - \mathbf{u}_n\|_\infty$.

5.2. Galerkin approach for random diffusion problems. In this example, we consider a time-periodic diffusion problem with random coefficient,

$$\begin{cases} \frac{\partial u(x,\omega,t)}{\partial t} = \nabla \cdot (\kappa(x,\omega)\nabla u(x,\omega,t)) + f(x,\omega,t), & (x,\omega,t) \in \Omega \times \Lambda \times (0,T), \\ u(x,\omega,0) = u(x,\omega,T), & (x,\omega) \in \Omega \times \Lambda, \\ u(x,\omega,t)|_{x\in\partial\Omega} = 0, & t \in (0,T), \end{cases} \tag{5.7}$$

where $\Omega = [0, 1]$, $\Lambda = [-1, 1]$, x is the spatial coordinate, and ω is a random parameter. We assume that the random diffusion function κ is of the form

$$\kappa(x,\omega) = \kappa_0(x) + \kappa_1(x)\omega, \tag{5.8}$$

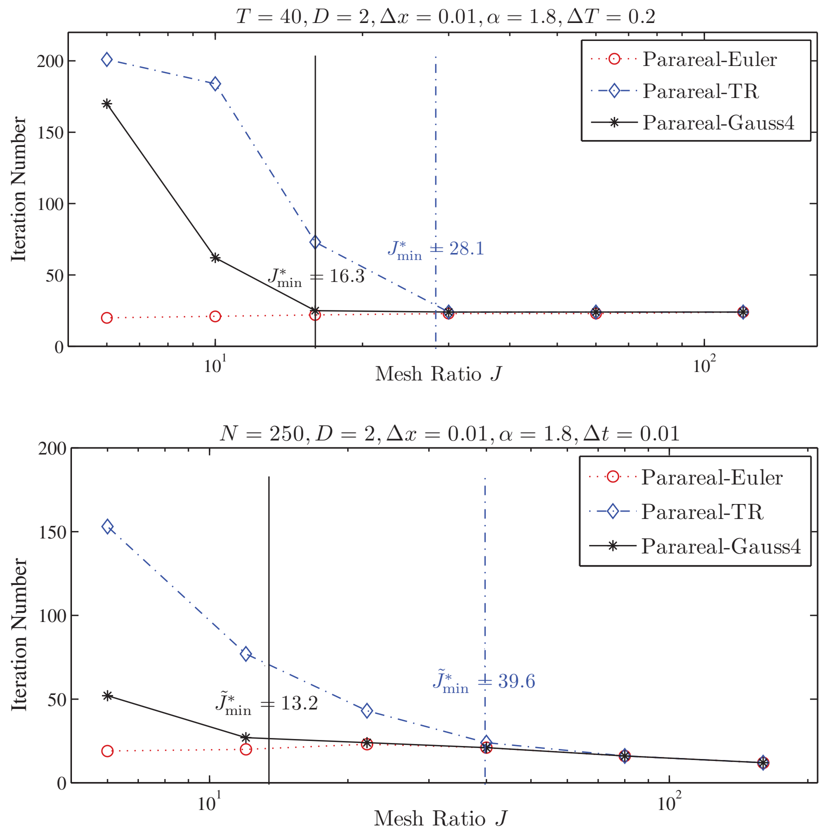

FIG. 8. *For Parareal-Euler, Parareal-TR, and Parareal-Gauss*4*, the iteration number required to satisfy the stopping criterion* (5.1) *for several different mesh ratios J. Top: ΔT is fixed to* 0.2*, independent of J; Bottom: ΔT is connected to J, $\Delta T = J\Delta t$ with $\Delta t = 0.01$ being fixed.*

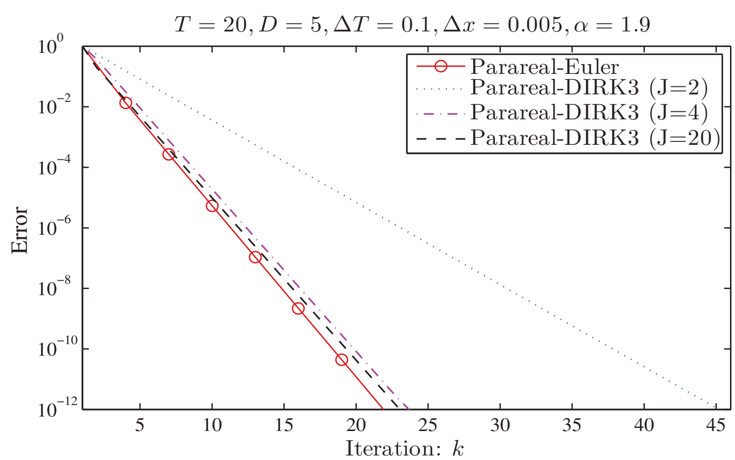

FIG. 9. *Measured convergence rates of the Parareal-Euler and Parareal-DIRK*3 *algorithms for $J = 2, 4$, and* 20*. (The convergence rate of the Parareal-Euler algorithm with the three mesh ratios looks similar and thus we only show the case $J = 20$.)*

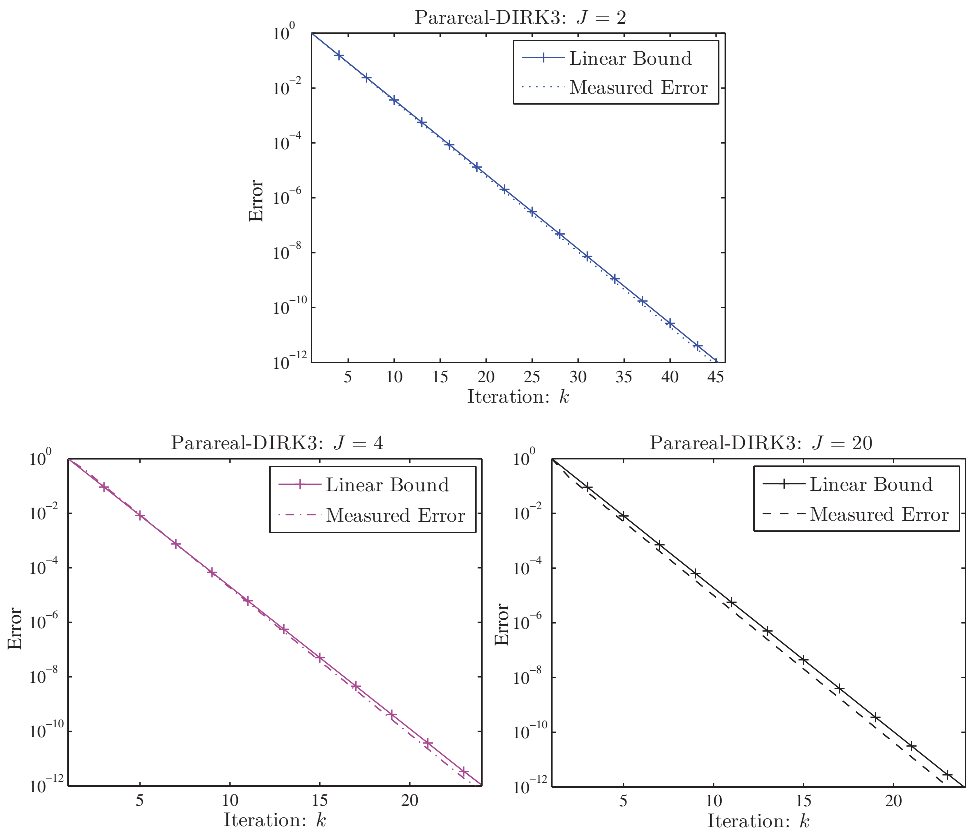

FIG. 10. *For the Parareal-DIRK*3 *algorithm, we show the measured error and the error predicted by the convergence factor* $\rho := \max_{z\in[0,z_{\max}]} \mathcal{K}_{\mathrm{DIRK3}}(z, J)$ *with the three mesh ratios* J *used in Figure* 9.

where $\kappa_0(x)$ and $\kappa_1(x)$ are known functions and $\kappa(x,\omega) > \delta > 0$ $\forall(x,\omega) \in \Omega \times \Lambda$. Here, we assume that the parameter ω is uniformly distributed in Λ.

A popular approach for solving the above random PDE is to construct the so-called generalized polynomial chaos (gPC) expansions [32, 33]:

$$u(x,\omega,t) = \sum_{j=0}^{N} \widehat{u}_j(x,t)L_j(\omega), \tag{5.9}$$

where $L_j(\omega)$ is the jth-order Legendre polynomial, and the $\{\widehat{u}_j(x,t)\}_{j=0}^N$ are the unknown functions to be computed. Substituting the approximation (5.9) into the governing equation (5.7) and then projecting the resulting equation onto the subspace spanned by the first $(N+1)$ gPC basis polynomials, gives the deterministic gPC system

$$\begin{cases} \frac{\partial \mathbf{u}}{\partial t} = \nabla\cdot[\mathbf{A}(x)\nabla\mathbf{u}] + \mathbf{f}(x,t), & (x,t)\in\Omega\times(0,T),\\ \mathbf{u}(x,0) = \mathbf{u}(x,T), & x\in\Omega,\\ \mathbf{u}(x,t)|_{x\in\partial\Omega} = 0, & t\in(0,T), \end{cases} \tag{5.10}$$

where $\mathbf{u} = (\widehat{u}_0(x,t),\ldots,\widehat{u}_N(x,t))^\top$ and $\mathbf{f}(x,t) = (\widehat{f}_0(x,t),\ldots,\widehat{f}_N(x,t))^\top$ with

$$\widehat{f}_j(x,t) = \int_\Lambda f(x,\omega,t)L_j(\omega)d\omega,\ j = 0,1,\ldots,N. \tag{5.11}$$

The components of the coefficient matrix $\mathbf{A}(x) = [a_{i,j}(x)]$ are defined as

$$a_{i,j}(x) = \int_\Lambda \kappa(x,\omega)L_i(\omega)L_j(\omega)d\omega. \tag{5.12}$$

In general, an M-point Gauss quadrature rule can be used to compute $a_{i,j}(x)$,

$$a_{i,j}(x) = \int_\Lambda \kappa(x,\omega)L_i(\omega)L_j(\omega)d\omega \approx \sum_{k=1}^{M} \omega_k \kappa(x,\omega_k)L_i(\omega_k)L_j(\omega_k), \tag{5.13}$$

where $\{y_k, \omega_k\}$ are the Gauss–Legendre points and the corresponding weights.

By solving the unknown variable $\mathbf{u}$ in (5.10) one gets the so-called gPC solution (5.9) of the original problem (5.7). With (5.8), the matrix $\mathbf{A}(x)$ can be expressed as

$$\mathbf{A}(x) = \kappa_0(x)\mathbf{A}_0 + \kappa_1(x)\mathbf{A}_1, \tag{5.14}$$

where $\mathbf{A}_0$ and $\mathbf{A}_1$ are constant matrices given explicitly by

$$\begin{aligned}
&\mathbf{A}_0 := \left[\int_{-1}^{1} L_i(\omega)L_j(\omega)d\omega\right] = \begin{pmatrix} \frac{2}{3} & & & \\ & \frac{2}{5} & & \\ & & \ddots & \\ & & & \frac{2}{2N+1} \end{pmatrix}_{N\times N}, \\
&\mathbf{A}_1 := \left[\int_{-1}^{1} \omega L_i(\omega)L_j(\omega)d\omega\right] = \begin{pmatrix} 0 & a_1 & & & \\ a_1 & 0 & a_2 & & \\ & \ddots & \ddots & \ddots & \\ & & a_{N-2} & 0 & a_{N-1} \\ & & & a_{N-1} & 0 \end{pmatrix}_{N\times N}, \\
&a_j = \frac{2(j+1)}{4(j+1)^2-1}, j = 1, 2, \ldots, N-1.
\end{aligned} \tag{5.15}$$

Clearly, $\mathbf{A}(x) = \mathbf{A}^\top(x)$ (for all $x \in \Omega$). Moreover, it can be shown that this matrix is strictly diagonal dominant if the perturbation term $\kappa_1(x)$ is small [33, 34]. Hence, $\mathbf{A}(x)$ is an SPD matrix for all $x \in \Omega$. In our numerical experiments we choose

$$\kappa_0 = \frac{1}{5},\ \kappa_1(x) = \frac{1}{20}(1+\cos(2\pi x)),\ \hat{f}_j(x,t) = \frac{t}{1+j}\cos(jxt\pi),\ j = 1, \ldots, N. \tag{5.16}$$

Applying the centered finite difference scheme to (5.10) with mesh size Δx gives

$$\begin{cases} \frac{d\mathbf{U}}{dt} + \frac{1}{\Delta x^2}\mathbf{Q}\mathbf{U} = \mathbf{F}(t),\ \mathbf{U}(0) = \mathbf{U}(T), \\ \mathbf{Q} = \begin{pmatrix} 2\tilde{\mathbf{A}}_1 & -\mathbf{A}_{1+\frac{1}{2}} & & & \\ -\mathbf{A}_{2-\frac{1}{2}} & 2\tilde{\mathbf{A}}_2 & -\mathbf{A}_{2+\frac{1}{2}} & & \\ & \ddots & \ddots & \ddots & \\ & & -\mathbf{A}_{m-1-\frac{1}{2}} & 2\tilde{\mathbf{A}}_{m-1} & -\mathbf{A}_{m-1+\frac{1}{2}} \\ & & & -\mathbf{A}_{m-\frac{1}{2}} & 2\tilde{\mathbf{A}}_m \end{pmatrix}_{mN\times mN}, \end{cases} \tag{5.17}$$

where $m = \frac{1}{\Delta x} - 1$, $\tilde{\mathbf{A}}_j = \frac{\mathbf{A}_{j-\frac{1}{2}} + \mathbf{A}_{j+\frac{1}{2}}}{2} = \frac{\mathbf{A}(x_{j-\frac{1}{2}}) + \mathbf{A}(x_{j+\frac{1}{2}})}{2}$, $x_{j\pm\frac{1}{2}} = x_j \pm \frac{\Delta x}{2}$, $x_j = j\Delta x$, and $j = 1, 2, \ldots, m$. Since $\mathbf{A}(x) = \mathbf{A}^\top(x)$ (for all $\in (0,1)$), we have $\mathbf{Q} = \mathbf{Q}^\top$. We

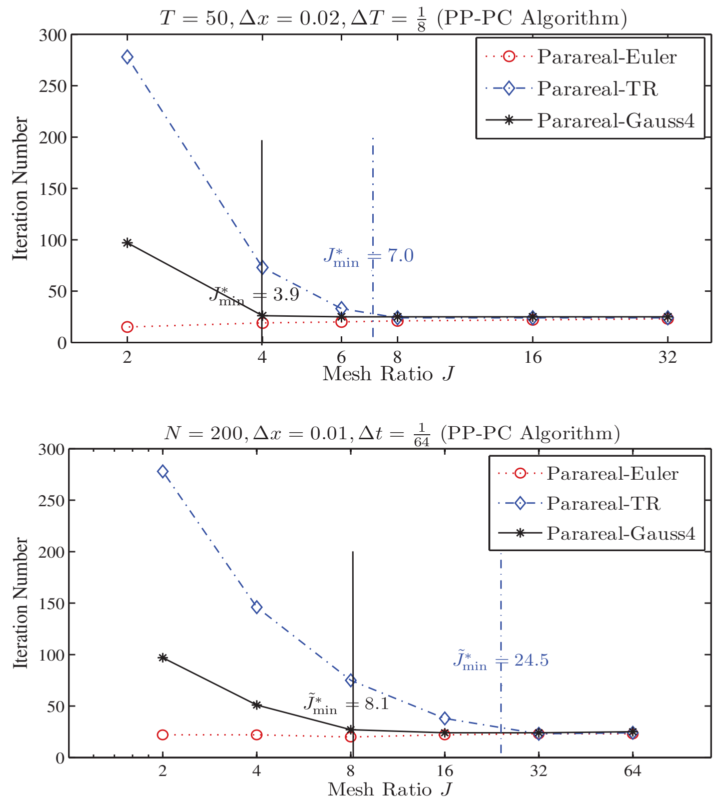

FIG. 11. *For the time-periodic problem* (5.17)*, the measured iteration number (required to satisfy the stopping criterion* (5.1)*) is shown for the Parareal-Euler, Parareal-TR, and Parareal-Gauss*4 *algorithms, for several different mesh ratios J. Top: ΔT is fixed to* $1/8$*, independent of J; Bottom: ΔT is connected to J, $\Delta T = J\Delta t$ with $\Delta t = 1/64$ being fixed.*

choose $\Delta x = \frac{1}{50}$ (or $\frac{1}{100}$) and $N = 6$. Computing the eigenvalues of $\mathbf{Q}$ numerically, we find that all of the eigenvalues are positive numbers. So, $\mathbf{Q}$ is an SPD matrix.

Equation (5.17) is a time-periodic problem and the PP-PC algorithm (3.7), proposed recently by Gander et al. [13], is suitable to solve it. To be precise, we use the backward-Euler method as the $\mathcal{G}$-propagator and several implicit time integrators, i.e., the backward-Euler method, the trapezoidal rule, the third-order DIRK method, and the fourth-order Gauss RK method, as the $\mathcal{F}$-propagator. We call the corresponding PP-PC algorithm Parareal-Euler, Parareal-TR, Parareal-DIRK3, and Parareal-Gauss4, too. In Figure 11, for several choices of the mesh ratio J, we plot the number of iterations (required to satisfy the stopping criterion (5.1)) for the Parareal-Euler, Parareal-TR, and Parareal-Gauss4 algorithms, in two different situations: (1) the length of the time interval and the coarse mesh size ΔT is fixed (top); (2) the number of coarse time intervals and the fine mesh size Δt is fixed (bottom). In both situations, we see that (a) the Parareal-Euler algorithm needs minimal iterations and is insensitive to the change of J; (b) the Parareal-TR and Parareal-Gauss4 algorithms converge as fast as Parareal-Euler, when $J \geq J^*_{\min}$ (or $J \geq \tilde{J}^*_{\min}$). Figure 11 shows that Theorems 3.1 and 3.2 also seem to apply to the PP-PC algorithm.

We also compared the performance of the Parareal-DIRK3 and Parareal-Euler algorithms; the result is very similar to Figure 9, so we omit the presentation.

6. Conclusions. For the fully discretized parareal algorithm using for $\mathcal{G}$ the backward-Euler method and for $\mathcal{F}$ some usual Runge–Kutta methods with order 1 to 4, the results obtained in this paper, together with the ones given recently in [24, Lemma 4.3] and [30], show the following convergence properties:

- for the algorithms Parareal-Euler [24], Parareal-TR/BDF2 [30], Parareal-DIRK2 [30], and Parareal-DIRK3 (this paper), the convergence factor can be a satisfactory quantity, say $\rho \approx \frac{1}{3}$—the best result we can expect so far, independent of the discretization/problem parameters (for the Parareal-DIRK3 algorithm we need $J \geq 4$, but this is a trivial restriction);
- for the algorithms Parareal-TR and Parareal-Gauss4 (this paper), they can also possess a convergence factor around $\frac{1}{3}$, but the mesh ratio J needs to be relatively large. The lower bound of J, which guarantees such a small convergence factor, is given in this paper.

The efficiency of the parareal algorithm is a different issue, but closely connected with the convergence analysis presented in this paper. Our results on efficiency will appear elsewhere [31, section 4].

Acknowledgment. The authors are very grateful to the anonymous referees for the careful reading of a preliminary version of the manuscript and their valuable suggestions and comments, which greatly improved the quality of this paper.

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, Lecture Notes Comput. Sci. Eng. 40, 2005, pp. 425–432.

[2] L. Baffico, S. Bernard, Y. Maday, G. Turinici, and G. Zérah, *Parallel-in-time molecular-dynamics simulations*, Phys. Rev. E (3), 66 (2002), 057701.

[3] C. Çelik and M. Duman, *Crank-Nicolson method for the fractional diffusion equation with the Riesz fractional derivative*, J. Comput. Phys., 231 (2012), pp. 1743–1750.

[4] E. Celledoni and T. Kvamsdal, *Parallelization in time for thermo-viscoplastic problems in extrusion of aluminium*, Internat. J. Numer. Methods Engrg., 79 (2009), pp. 576–598.

[5] J. Cortial and C. Farhat, *A time-parallel implicit method for accelerating the solution of nonlinear structural dynamics problems*, Internat. J. Numer. Methods Engrg., 77 (2008), pp. 451–470.

[6] X. H. Du, M. Sarkis, C. F. Schaerer, and D. B. Szyld, *Inexact and truncated parareal-in-time Krylov subspace methods for parabolic optimal control problems*, Electron. Trans. Numer. Anal., 40 (2013), pp. 36–57.

[7] X. Dai and Y. Maday, *Stable parareal in time method for first- and second-order hyperbolic systems*, SIAM J. Sci. Comput., 35 (2013), pp. A52–A78.

[8] X. Dai, C. Bris, F. Legoll, and Y. Maday, *Symmetric parareal algorithms for Hamiltonian systems*, M2AN Math. Model Numer. Anal., 47 (2012), pp. 717–742.

[9] M. Emmett and M. L. Minion, *Toward an efficient parallel in time method for partial differential equations*, Comm. Appl. Math. Comput. Sci., 7 (2012), pp. 105–132.

[10] C. Farhat and M. Chandesris, *Time-decomposed parallel time-integrators: Theory and feasibility studies for fluid, structure, and fluid-structure applications*, Internat. J. Numer. Methods Engng., 58 (2003), pp. 1397–1434.

[11] M. J. Gander and E. Hairer, *Nonlinear convergence analysis for the parareal algorithm*, in Domain Decomposition Methods in Science and Engineering, XVII, Lecture Notes Comput. Sci. Eng. 60, 2008, pp. 45–56.

[12] M. J. Gander and S. Vandewalle, *Analysis of the parareal time-parallel time-integration method*, SIAM J. Sci. Comput., 29 (2007), pp. 556–578.

[13] M. J. Gander, Y.-L. Jiang, B. Song, and H. Zhang, *Analysis of two parareal algorithms for time-periodic problems*, SIAM J. Sci. Comput., 35 (2013), pp. A2393–A2415.

[14] M. J. Gander and E. Hairer, *Analysis for parareal algorithms applied to Hamiltonian differential equations*, J. Comput. Appl. Math., 259 (2014), pp. 2–13.

[15] L. He and M. He, *Parareal in time simulation of morphological transformation in cubic alloys with spatially dependent composition*, Commun. Comput. Phys., 11 (2012), pp. 1697–1717.
[16] M. J. Gander, 50 *years of time parallel time integration*, in Multiple Shooting and Time Domain Decomposition, T. Cararro, M. Geiger, S. Körkel, and R. Rannacher, eds., Springer-Verlag, Heidelburg, 2015.
[17] X. J. Li, T. Tang, and C. J. Xu, *Parallel in time algorithm with spectral-subdomain enhancement for Volterra integral equations*, SIAM J. Numer. Anal., 51 (2013), pp. 1735–1756.
[18] J.-L. Lions, Y. Maday, and G. Turinici, *A "parareal" in time discretization of PDE's*, C. R. Acad. Sci. Paris Sér. I Math., 332 (2001), pp. 661–668.
[19] F. Legoll, T. Lelièvre, and G. Samaey, *A micro-macro parareal algorithm: Application to singularly perturbed ordinary differential equations*, SIAM J. Sci. Comput., 35 (2013), pp. A1951–A1986.
[20] Y. Maday, J. Salomon, and G. Turinici, *Monotonic parareal control for quantum systems*, SIAM J. Numer. Anal., 45 (2007), pp. 2468–2482.
[21] M. L. Minion, *A hybrid parareal spectral deferred corrections method*, Appl. Math. Comput. Sci., 5 (2010), pp. 265–301.
[22] M. J. Gander, M. Petcu, *Analysis of a Krylov subspace enhanced parareal algorithm for linear problems*, ESAIM Proceedings, 25 (2008), pp. 114–129.
[23] Y. Maday, M.-K. Riahi, and J. Salomon, *Parareal in time intermediate targets methods for optimal control problems*, in Control and Optimization with PDE Constraints, Internat. Ser. Numer. Math. 164, 2013, pp. 79–92.
[24] T. P. Mathew, M. Sarkis, and C. E. Schaerer, *Analysis of block parareal preconditioners for parabolic optimal control problems*, SIAM J. Sci. Comput., 32 (2010), pp. 1180–1200.
[25] I. Podlubny, A. Chechkin, I. Skovranek, Y. Chen, and B. M. Vinagre Jara, *Matrix approach to discrete fractional calculus* II: *Partial fractional differential equations*, J. Comput. Phys., 228 (2009), pp. 3137–3153.
[26] J. M. Reynolds-Barredo, D. E. Newman, R. Sanchez, D. Samaddar, L. A. Berry, and W. R. Elwasif, *Mechanisms for the convergence of time-parallelized, parareal turbulent plasma simulations*, J. Comput. Phys., 231 (2012), pp. 7851–7867.
[27] J. M. Reynolds-Barredo, D. E. Newman, and R. Sanchez, *An analytic model for the convergence of turbulent simulations time-parallelized via the parareal algorithm*, J. Comput. Phys., 255 (2013), pp. 293–315.
[28] G. A. Staff and E. M. Rønquist, *Stability of the parareal algorithm*, in Domain Decomposition Methods in Science and Engineering, Lecture Notes Comput. Sci. Eng. 40, 2005, pp. 449–456.
[29] S. L. Wu, B. C. Shi, and C. M. Huang, *Parareal-Richardson algorithm for solving nonlinear ODEs and PDEs*, Commun. Comput. Phys., 6 (2009), pp. 883–902.
[30] S. L. Wu, *Convergence analysis of some second-order parareal algorithms*, IMA J. Numer. Anal., to appear.
[31] S. L. Wu, *Rapidly convergent parareal solvers with order* 2 *and* 4 *for SPD ODEs*, IMA J. Numer. Anal., submitted.
[32] D. Xiu and G. E. Karniadakis, *The Wiener–Askey polynomial chaos for stochastic differential equations*, SIAM J. Sci. Comput., 24 (2002), pp. 619–644.
[33] D. Xiu and J. Shen, *Efficient stochastic Galerkin methods for random diffusion equations*, J. Comput. Phys., 228 (2009), pp. 266–281.
[34] T. Zhou and T. Tang, *Note on coefficient matrices from stochastic Galerkin methods for random diffusion equations*, J. Comput. Phys., 229 (2010), pp. 8225–8230.