SIAM J. SCI. COMPUT.
Vol. 41, No. 2, pp. B375–B395

A NEW PARAREAL ALGORITHM FOR PROBLEMS WITH DISCONTINUOUS SOURCES*

MARTIN J. GANDER†, IRYNA KULCHYTSKA-RUCHKA‡, INNOCENT NIYONZIMA§, AND SEBASTIAN SCHÖPS‡

Abstract. The Parareal algorithm allows one to solve evolution problems exploiting parallelization in time. Its convergence and stability have been proved under the assumption of regular (smooth) inputs. We present and analyze here a new Parareal algorithm for ordinary differential equations which involve discontinuous right-hand sides. Such situations occur in various applications, e.g., when an electric device is supplied with a pulse-width-modulated signal. Our new Parareal algorithm uses a smooth input for the coarse problem with reduced dynamics. We derive error estimates that show how the input reduction influences the overall convergence rate of the algorithm. We support our theoretical results by numerical experiments, and also test our new Parareal algorithm in an eddy current simulation of an induction machine.

Key words. evolution problems, parallel-in-time solution, Parareal, ODEs with discontinuous inputs, convergence analysis

AMS subject classifications. 34A34, 34A36, 34A37, 65L20, 78M10

DOI. 10.1137/18M1175653

1. Introduction.

Due to the increasing computational power of modern computer systems, scientists are nowadays able to solve complex physical problems, and parallel computers allow one to reduce the time to obtain the solution further. The first and most natural approach to solve evolution problems in parallel is to perform parallel computations in space by domain decomposition; see [35, 41, 15, 6] and references therein. However, when space parallelization is exploited up to saturation, and more processors are still available, parallel-in-time methods are considered to be a complementary approach to achieve further numerical speed-up; see [16] for an overview of such techniques.

The Parareal algorithm was introduced by Lions, Maday, and Turinici in [26]. It has become a powerful tool, which allows one to solve time-dependent problems in a time-parallel fashion. The method has been applied to a wide range of problems [33], in particular, linear and nonlinear parabolic problems [39, 27, 43, 44], molecular dynamics [2], stochastic ordinary differential equations (ODEs) [3, 11], Navier–Stokes equations [42, 14], quantum control problems [30, 29] and optimal control problems [31], as well as low-frequency problems in electrical engineering [38].

*Submitted to the journal's Computational Methods in Science and Engineering section March 15, 2018; accepted for publication (in revised form) February 14, 2019; published electronically April 2, 2019.

http://www.siam.org/journals/sisc/41-2/M117565.html

Funding: This research was supported by the Excellence Initiative of the German Federal and State Governments and the Graduate School of Computational Engineering at TU Darmstadt, as well as by DFG grant SCHO1562/1-2 and BMBF grant 05M2018RDA (PASIROM).

†Section de Mathématiques, Université de Genève, CH-1211 Geneva, Switzerland (Martin.Gander@unige.ch).

‡Institut für Teilchenbeschleinigung und Elektromagnetische Felder, Technische Universität Darmstadt, D-64289 Darmstadt, Germany (kulchytska@gsc.tu-darmstadt.de, schoeps@temf.tu-darmstadt.de).

§Department of Civil Engineering and Engineering Mechanics, Columbia University, New York, NY 10027 (inno_niyo@yahoo.com).

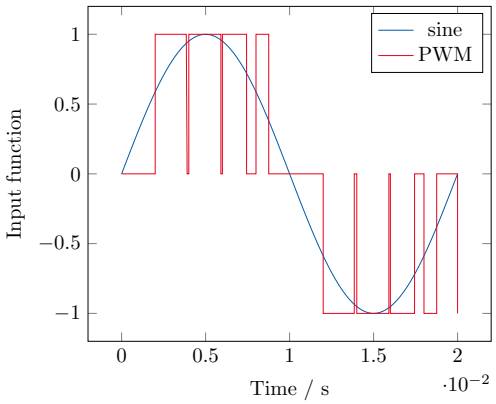

FIG. 1. *PWM signal with switching frequency of* $f_s = 500$ *Hz, generating a sine wave of* 50 *Hz.*

The Parareal algorithm is based on a decomposition of the time domain of interest into nonoverlapping time intervals (e.g., one time interval per processor) and the parallel solution of the governing equation on each time interval. Exchange of information at synchronization points is based on the action of fine and coarse propagators. Starting from a prescribed initial guess, both operators solve the underlying problem over each time interval and return the solution at the end of the time interval. The fine propagator is accurate and computationally expensive. It can be, for example, a classical time integrator, which uses a very fine time discretization. On the other hand, the coarse propagator is less accurate, but much less expensive than the fine propagator (e.g., via time stepping over a coarse partition). The Parareal algorithm corrects the approximate solution iteratively until convergence.

Several techniques for reducing the computational cost of Parareal are discussed in [28]. In particular, for the time domain solution of partial differential equations (PDEs), the use of a coarse mesh also in space for the coarse propagator is proposed. This approach can be used within a multiscale setting [1, 8] or with spatial averaging operators [4]. A second idea is to perform model order reduction (MOR) for the extraction of a coarse propagator from the fine problem. Further reduced order techniques, developed in [7, 23], involve spatial MOR also for the coarse problem.

These ideas help to reduce the cost of Parareal by simplifying the coarse model in space. In this paper we propose to use a simpler coarse problem with respect to the time variable, similar to [22], where Parareal was applied to PDEs which exhibit scale separation in time. Our method is specific for problems involving discontinuous or multirate excitations, e.g., pulse-width-modulated (PWM) signals, an example of which is shown in Figure 1 for multiharmonic signals. Its main idea is to supply the coarse propagator with a smooth input, which features reduced dynamics, e.g., a periodic waveform, which consists of the fundamental frequency only. For instance, in the case of the PWM signal containing 10 pulses on the time interval $[0, 0.02]$ s, one could choose a sine wave of 50 Hz to be the smooth input, also shown in Figure 1. This allows the coarse propagator to use larger time steps and a high order method.

Our paper is organized as follows. The problem setting is described in section 2. The original Parareal algorithm for a system of nonlinear ODEs, together with its error estimate from [17] are recalled in section 3. In section 4, we present our new Parareal algorithm for a subclass of Carathéodory equations—equations whose inputs may contain discontinuities with respect to the time variable, and we derive a

sharp convergence estimate using techniques developed in [17]. We then measure the convergence rate of our new Parareal algorithm numerically in section 5 for an RL-circuit model, and observe very good agreement with our theoretical estimates. In section 6, we test the new Parareal algorithm applied to an eddy current simulation of an induction machine. We finally present our conclusions in section 7.

2. Problem setting. We consider a nonlinear initial value problem (IVP) of nonautonomous ODEs of the form

$$\boldsymbol{u}'(t) = \boldsymbol{f}(t, \boldsymbol{u}(t)), \quad t \in \mathcal{I}, \tag{2.1}$$

$$\boldsymbol{u}(0) = \boldsymbol{u}_0 \tag{2.2}$$

with right-hand side (RHS) $\boldsymbol{f} : \mathcal{I} \times \mathbb{R}^n \to \mathbb{R}^n$ and solution $\boldsymbol{u} : \mathcal{I} \to \mathbb{R}^n$ on the time interval $\mathcal{I} := (0, T]$. We are interested in problems for which the nonsmooth (or even discontinuous) excitation can be separated from the smooth part of the RHS, i.e.,

$$\boldsymbol{f}(t, \boldsymbol{u}(t)) := \bar{\boldsymbol{f}}(t, \boldsymbol{u}(t)) + \tilde{\boldsymbol{f}}(t), \tag{2.3}$$

where $\bar{\boldsymbol{f}}(t, \boldsymbol{u}(t))$ and $\tilde{\boldsymbol{f}}(t)$ satisfy the following two assumptions.

Assumption 1. The function $\bar{\boldsymbol{f}}$ in (2.3) is bounded and sufficiently smooth in both arguments, and it is Lipschitz in the second argument with Lipschitz constant L.

Assumption 2. The function $\tilde{\boldsymbol{f}}$ in (2.3) belongs to $\boldsymbol{L}^p(\mathcal{I}, \mathbb{R}^n)$, $p \geq 1$, with its norm given by $C_p := \|\tilde{\boldsymbol{f}}\|_{L^p(\mathcal{I}, \mathbb{R}^n)}$. The notation $\boldsymbol{L}^p(\mathcal{I}, \mathbb{R}^n)$ describes the Lebesgue space of vector-valued functions (see, e.g., [45, Chapter 23.2]) $\tilde{\boldsymbol{f}} : \mathcal{I} \to \mathbb{R}^n$ with the norm given by

$$\|\tilde{\boldsymbol{f}}\|_{L^p(\mathcal{I}, \mathbb{R}^n)} = \left(\int_{\mathcal{I}} \|\tilde{\boldsymbol{f}}(t)\|_{\mathbb{R}^n}^p dt \right)^{1/p} < \infty.$$

For the case $p = \infty$, the norm is defined by

$$\|\tilde{\boldsymbol{f}}\|_{L^\infty(\mathcal{I}, \mathbb{R}^n)} := \operatorname*{ess\,sup}_{t \in \mathcal{I}} \|\tilde{\boldsymbol{f}}(t)\|_{\mathbb{R}^n} < \infty. \tag{2.4}$$

By $\|\cdot\|_{\mathbb{R}^n}$ we denote an arbitrary finite norm of an n-dimensional vector, e.g., the Euclidean norm.

Clearly, the total RHS $\boldsymbol{f}$ has no continuity or smoothness properties, and therefore the Lindelöf theory for existence and uniqueness of solutions cannot be applied to (2.1)–(2.2). However, one can use the solvability and uniqueness theory for Carathéodory equations, which can be found, e.g., in [13]. We recall that (2.1) is called a Carathéodory equation if its RHS $\boldsymbol{f}(t, \boldsymbol{u})$ satisfies the so-called Carathéodory conditions:

(a) $\boldsymbol{f}(t, \boldsymbol{u})$ is defined and continuous in $\boldsymbol{u}$ for almost all t;
(b) $\boldsymbol{f}(t, \boldsymbol{u})$ is measurable in t for each $\boldsymbol{u}$;
(c) $\|\boldsymbol{f}(t, \boldsymbol{u})\| \leq m(t)$ with m being a summable function on $\mathcal{I}$.

It was proved in [13] that there exists a solution to (2.1)–(2.2) if $\boldsymbol{f}(t, \boldsymbol{u})$ satisfies the Carathéodory conditions (a)–(c). Furthermore, if there exists a summable function $l(t)$ s.t. $\forall (t, \boldsymbol{v})$ and $\forall (t, \boldsymbol{u})$ with $t \in \mathcal{I}$

$$\|\boldsymbol{f}(t, \boldsymbol{u}) - \boldsymbol{f}(t, \boldsymbol{v})\| \leq l(t)\|\boldsymbol{u} - \boldsymbol{v}\|, \tag{2.5}$$

then the solution is unique. Note that Assumptions 1 and 2 imply that the Carathéodory conditions and (2.5) are satisfied and, hence, there exists a unique solution to (2.1)–(2.2).

3. Original Parareal algorithm and convergence for smooth RHSs. We now recall the original Parareal algorithm from [26] in the form described in [18] for solving (2.1)–(2.2). The initial step of the algorithm consists in partitioning the time domain $(0,T]$ into nonoverlapping time intervals $(T_{n-1},T_n]$, $n=1,\dots,N$, with $0=T_0<T_1<T_2<\cdots<T_N=T$. One can then define an evolution problem on each time interval,

$$\boldsymbol{u}_n'(t)=\boldsymbol{f}(t,\boldsymbol{u}_n(t)),\quad t\in(T_{n-1},T_n], \tag{3.1}$$

$$\boldsymbol{u}_n(T_{n-1})=\boldsymbol{U}_{n-1} \tag{3.2}$$

for $n=1,\dots,N$. The initial values $\boldsymbol{U}_{n-1}$, $n=1,\dots,N$, need to be determined such that the solutions on each time interval $(T_{n-1},T_n]$ coincide with the restriction of the solution of (2.1)–(2.2) to that time interval. The Parareal algorithm computes by iteration better and better approximations of these initial conditions: for a given initial guess $\boldsymbol{U}_n^{(0)}$, $n=0,\dots,N$, it solves for $k=0,1,\dots,K$

$$\boldsymbol{U}_0^{(k+1)}=\boldsymbol{u}_0, \tag{3.3}$$

$$\begin{aligned}\boldsymbol{U}_n^{(k+1)}&=\mathcal{F}\big(T_n,T_{n-1},\boldsymbol{U}_{n-1}^{(k)}\big)+\mathcal{G}\big(T_n,T_{n-1},\boldsymbol{U}_{n-1}^{(k+1)}\big)\\&\quad-\mathcal{G}\big(T_n,T_{n-1},\boldsymbol{U}_{n-1}^{(k)}\big),\quad n=1,\dots,N.\end{aligned} \tag{3.4}$$

In (3.4) we denote by $\mathcal{F}(t,T_{n-1},\boldsymbol{U}_{n-1})$ and $\mathcal{G}(t,T_{n-1},\boldsymbol{U}_{n-1})$ the numerical solution propagators of the IVP (3.1)–(3.2). Both of them propagate the initial value $\boldsymbol{U}_{n-1}$ in time on $(T_{n-1},T_n]$, but they differ in accuracy: the fine propagator $\mathcal{F}$ gives a very accurate, but expensive approximate solution to the IVP, whereas the coarse propagator $\mathcal{G}$ gives an inexpensive, but less accurate solution. The first term of the RHS in (3.4) involves quantities which are already known at the iteration $k+1$ and, therefore, can be computed in parallel. The last one is known as well, since it has been already computed at the previous iteration. The term $\mathcal{G}(T_n,T_{n-1},\boldsymbol{U}_{n-1}^{(k+1)})$ involves the approximation $\boldsymbol{U}_{n-1}^{(k+1)}$, $n=1,\dots,N$, which has not yet been obtained at the beginning of the iteration $k+1$. Therefore, its calculation cannot be parallelized and the coarse but inexpensive propagator $\mathcal{G}$ is applied sequentially.

We now state the convergence result for problems with smooth RHS $\boldsymbol{f}$, which was proved in [17] under the assumption that each time interval has the same length $\Delta T=T/N$.

THEOREM 3.1. *Assume that the RHS $\boldsymbol{f}$ is smooth enough and $\mathcal{F}(T_n,T_{n-1},\boldsymbol{U}_{n-1}^{(k)})$ is the exact solution to* (3.1)–(3.2) *at T_n with initial value $\boldsymbol{U}_{n-1}^{(k)}$. Furthermore,*

- *let $\mathcal{G}(T_n,T_{n-1},\mathbf{U}_{n-1}^{(k)})$ be an approximate solution with local truncation error bounded by $C_3\Delta T^{l+1}$, which can be expanded for ΔT small as*
  $$\mathcal{F}(T_n,T_{n-1},\boldsymbol{U})-\mathcal{G}(T_n,T_{n-1},\boldsymbol{U})=c_{l+1}(\boldsymbol{U})\Delta T^{l+1}+c_{l+2}(\boldsymbol{U})\Delta T^{l+2}+\cdots \tag{3.5}$$
  with an initial value $\boldsymbol{U}$ and continuously differentiable functions c_i, $i=l+1,l+2,\dots$;
- *assume that $\mathcal{G}$ satisfies the Lipschitz condition*
  $$\|\mathcal{G}\big(t+\Delta T,t,\boldsymbol{U}\big)-\mathcal{G}\big(t+\Delta T,t,\boldsymbol{V}\big)\|\le(1+C_2\Delta T)\|\boldsymbol{U}-\boldsymbol{V}\| \tag{3.6}$$
  for $t\in\mathcal{I}$ and for all $\boldsymbol{U},\boldsymbol{V}$, with constant C_2.

Then at iteration k of the Parareal algorithm (3.3)–(3.4) *we have the error bound*

$$\|\boldsymbol{u}(T_n) - \boldsymbol{U}_n^{(k)}\| \leq \frac{C_3}{C_1}\frac{(C_1\Delta T^{l+1})^{k+1}}{(k+1)!}(1+C_2\Delta T)^{n-k-1}\prod_{j=0}^{k}(n-j), \tag{3.7}$$

where the constant C_1 comes from the expansion (3.5) *and the Lipschitz continuity of c_i, $i = l+1, l+2, \ldots$; see the proof in* [17].

4. A new Parareal algorithm for nonsmooth sources. We now omit the assumption of smoothness on the RHS and allow discontinuities in the time-dependent input $\tilde{\boldsymbol{f}}$, considering the IVP (2.1)–(2.2) with $\boldsymbol{f}$ as in (2.3) such that only Assumptions 1 and 2 are satisfied.

When one deals with a highly oscillatory or discontinuous source, the coarse propagator $\mathcal{G}$ might not capture its dynamics if low accuracy, i.e., big time steps, is used. This may lead to solving a coarse problem, which does not contain enough information about the original input, and it is not clear how this influences the overall convergence of the Parareal algorithm. For this reason, we propose to define a smooth input, which is appropriate for coarse discretization. Therefore, in our new Parareal algorithm, the coarse propagator solves the modified problem with reduced dynamics

$$\bar{\boldsymbol{u}}'(t) = \bar{\boldsymbol{f}}(t, \bar{\boldsymbol{u}}(t)), \quad t \in \mathcal{I}, \tag{4.1}$$

$$\bar{\boldsymbol{u}}(0) = \boldsymbol{u}_0, \tag{4.2}$$

while the fine propagator $\mathcal{F}$ is still applied to the original problem (2.1)–(2.2). In particular, the coarse propagator $\bar{\mathcal{G}}$ on the time interval $(T_{n-1}, T_n]$ for $n = 1, \ldots, N$ solves

$$\bar{\boldsymbol{u}}_n'(t) = \bar{\boldsymbol{f}}(t, \bar{\boldsymbol{u}}_n(t)), \quad t \in (T_{n-1}, T_n], \tag{4.3}$$

$$\bar{\boldsymbol{u}}_n(T_{n-1}) = \boldsymbol{U}_{n-1}. \tag{4.4}$$

Our new Parareal algorithm then computes for $k = 0, 1, \ldots, K$ and $n = 1, \ldots, N$

$$\boldsymbol{U}_0^{(k+1)} = \boldsymbol{u}_0, \tag{4.5}$$

$$\boldsymbol{U}_n^{(k+1)} = \mathcal{F}\big(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(k)}\big) + \bar{\mathcal{G}}\big(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(k+1)}\big) - \bar{\mathcal{G}}\big(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(k)}\big). \tag{4.6}$$

The initial approximation can be calculated using the coarse propagator,

$$\boldsymbol{U}_0^{(0)} := \boldsymbol{u}_0, \tag{4.7}$$

$$\boldsymbol{U}_n^{(0)} := \bar{\mathcal{G}}\big(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(0)}\big), \quad n = 1, \ldots, N. \tag{4.8}$$

For a given initial value $\boldsymbol{U}$, we define the difference between the exact solution of (3.1) and the numerical solution of the reduced coarse problem (4.3) as

$$\boldsymbol{\epsilon}_n(T_n, \boldsymbol{U}) := \mathcal{F}(T_n, T_{n-1}, \boldsymbol{U}) - \bar{\mathcal{G}}(T_n, T_{n-1}, \boldsymbol{U}). \tag{4.9}$$

Let $\boldsymbol{u}_n$ and $\bar{\boldsymbol{u}}_n$ solve (3.1)–(3.2) and (4.3)–(4.4), respectively. We define the error $\boldsymbol{\epsilon}_{f,n}$ as the difference

$$\boldsymbol{\epsilon}_{f,n} := \boldsymbol{u}_n - \bar{\boldsymbol{u}}_n \text{ on } [T_{n-1}, T_n] \tag{4.10}$$

and show that it does not depend on the initial value $\boldsymbol{U}_{n-1}$ in the following proposition.

PROPOSITION 4.1. *If Assumptions* 1 *and* 2 *hold, then the error* $\boldsymbol{\epsilon}_{f,n}$ *from* (4.10) *solves the IVP*

$$\begin{aligned} \boldsymbol{\epsilon}'_{f,n}(t) &= \boldsymbol{J}(t, \boldsymbol{\epsilon}_{f,n}(t))\boldsymbol{\epsilon}_{f,n}(t) + \tilde{\boldsymbol{f}}(t), \quad t \in (T_{n-1}, T_n], \\ \boldsymbol{\epsilon}_{f,n}(T_{n-1}) &= 0, \end{aligned} \tag{4.11}$$

where $\boldsymbol{J}(t, \boldsymbol{\epsilon}_{f,n}(t))$ *is defined in* [9] *as the neighborhood average of the Jacobian, given by*

$$\boldsymbol{J}(t, \boldsymbol{\epsilon}_{f,n}(t)) = \int_0^1 \frac{\partial \bar{\boldsymbol{f}}}{\partial \boldsymbol{u}} \left(t, \bar{\boldsymbol{u}}(t) + \theta \boldsymbol{\epsilon}_{f,n}(t)\right) \mathrm{d}\theta. \tag{4.12}$$

Proof. Subtracting (4.3) from (3.1) and initial condition (4.4) from (3.2) gives

$$\begin{aligned} \boldsymbol{\epsilon}'_{f,n}(t) &= \bar{\boldsymbol{f}}\left(t, \bar{\boldsymbol{u}}_n(t) + \boldsymbol{\epsilon}_{f,n}(t)\right) - \bar{\boldsymbol{f}}\left(t, \bar{\boldsymbol{u}}_n(t)\right) + \tilde{\boldsymbol{f}}(t), \quad t \in (T_{n-1}, T_n], \\ \boldsymbol{\epsilon}_{f,n}(T_{n-1}) &= 0. \end{aligned} \tag{4.13}$$

Using the fundamental theorem of calculus, we get

$$\begin{aligned} &\bar{\boldsymbol{f}}\left(t, \bar{\boldsymbol{u}}_n(t) + \boldsymbol{\epsilon}_{f,n}(t)\right) - \bar{\boldsymbol{f}}\left(t, \bar{\boldsymbol{u}}_n(t)\right) \\ &\quad = \int_0^1 \frac{\partial \bar{\boldsymbol{f}}}{\partial \theta} \left(t, \bar{\boldsymbol{u}}_n(t) + \theta \boldsymbol{\epsilon}_{f,n}(t)\right) \mathrm{d}\theta \end{aligned} \tag{4.14}$$

$$\begin{aligned} &\quad = \int_0^1 \frac{\partial \bar{\boldsymbol{f}}}{\partial \boldsymbol{u}} \left(t, \bar{\boldsymbol{u}}_n(t) + \theta \boldsymbol{\epsilon}_{f,n}(t)\right) \boldsymbol{\epsilon}_{f,n}(t) \mathrm{d}\theta \\ &\quad =: \boldsymbol{J}(t, \boldsymbol{\epsilon}_{f,n}(t))\boldsymbol{\epsilon}_{f,n}(t), \end{aligned} \tag{4.15}$$

which leads to (4.11). ∎

Remark 4.2. We note that the IVP (4.11) is again well-defined in the sense of Carathéodory theory.

In the following lemma we derive a bound for the error $\boldsymbol{\epsilon}_{f,n}(T_n)$, the solution to (4.11) at T_n.

LEMMA 4.3. *Let Assumptions* 1 *and* 2 *hold, and let the time interval length* $\Delta T = T/N$ *be small. Then there exists* $C_4 > 0$ *s.t. the solution to* (4.11) *can be bounded at* T_n *by*

$$\|\boldsymbol{\epsilon}_{f,n}(T_n)\| \leq C_4 C_p \Delta T^{1/q}, \tag{4.16}$$

where the integer $q \geq 1$ *is defined by the relation* $1/p + 1/q = 1$, *and* $C_p := \|\tilde{\boldsymbol{f}}\|_{L^p(\mathcal{I}, \mathbb{R}^n)}$ *is from Assumption* 2.

Proof. Let us denote an arbitrary spatial norm of $\tilde{\boldsymbol{f}}(t)$ in $\mathbb{R}^n$ by $\varepsilon(t) := \|\tilde{\boldsymbol{f}}(t)\|$. Then, based on [21, Theorem 10.2] for $t \geq T_{n-1}$, one can bound the error $\boldsymbol{\epsilon}_{f,n}$ by

$$\|\boldsymbol{\epsilon}_{f,n}(t)\| \leq \mathrm{e}^{L(t-T_{n-1})} \int_{T_{n-1}}^{t} \mathrm{e}^{-L(s-T_{n-1})} \varepsilon(s) \mathrm{d}s, \tag{4.17}$$

since initially at T_{n-1} the error $\boldsymbol{\epsilon}_{f,n}$ equals zero and is thus bounded, the norm $\|\tilde{\boldsymbol{f}}(t)\|$ is bounded by $\varepsilon(t)$, and the function $\bar{\boldsymbol{f}}$ is Lipschitz continuous with Lipschitz constant

L, as stated in Assumption 1. Taking $t = T_n$ in (4.17) and using Hölder's inequality together with a Taylor expansion for ΔT small, we obtain

$$\begin{aligned}\|\boldsymbol{\epsilon}_{f,n}(T_n)\| &\le \mathrm{e}^{L\Delta T}\int_{T_{n-1}}^{T_n}\left|\mathrm{e}^{-L(s-T_{n-1})}\varepsilon(s)\right|\mathrm{d}s\\ &\le \mathrm{e}^{L\Delta T}\left(\int_{T_{n-1}}^{T_n}\left|\mathrm{e}^{-L(s-T_{n-1})}\right|^q\mathrm{d}s\right)^{1/q}\left(\int_{T_{n-1}}^{T_n}|\varepsilon(s)|^p\mathrm{d}s\right)^{1/p}\\ &= \left(1+L\Delta T+\mathcal{O}\left(\Delta T^2\right)\right)\left[\Delta T+\mathcal{O}\left(\Delta T^2\right)\right]^{1/q}\|\varepsilon\|_{L^p(T_{n-1},T_n)}\\ &\le C_p\Delta T^{1/q}+\mathcal{O}\left(\Delta T^{2/q}\right)\le C_4C_p\Delta T^{1/q}\end{aligned}$$

with $q \ge 1$ satisfying $1/p + 1/q = 1$ and the constant C_4 coming from the definition of the Landau symbol "big $\mathcal{O}$." ∎

We can now prove a convergence result for our new Parareal algorithm for nonsmooth input (4.5)–(4.6) for problem (2.1)–(2.2), which is similar to that of Theorem 3.1, derived for the case of smooth RHS. Like in Theorem 3.1, we also assume that the time intervals have equal length, $\Delta T = T/N$. For analysis purposes, we introduce an additional propagator $\bar{\mathcal{F}}$, which, as $\bar{\mathcal{G}}$, solves (4.1)–(4.2), but is exact. Note that assuming that both fine propagators $\mathcal{F}(T_n, T_{n-1}, \boldsymbol{U})$ and $\bar{\mathcal{F}}(T_n, T_{n-1}, \boldsymbol{U})$ give exact solutions to (3.1) and (4.3), respectively, with an initial value $\boldsymbol{U}$ at T_{n-1}, one can write the error $\boldsymbol{\epsilon}_{f,n}$ defined in (4.10) at T_n as $\boldsymbol{\epsilon}_{f,n}(T_n) = \mathcal{F}(T_n, T_{n-1}, \boldsymbol{U}) - \bar{\mathcal{F}}(T_n, T_{n-1}, \boldsymbol{U})$. With this setting we derive the error estimate of the new Parareal algorithm with reduced coarse problem in our main theorem below.

THEOREM 4.4. *Let Assumptions* 1 *and* 2 *be satisfied, and assume that* $\mathcal{F}(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(k)})$ *is the exact solution to* (3.1)–(3.2) *at* T_n *with initial value* $\boldsymbol{U}_{n-1}^{(k)}$. *Furthermore,*

- *let* $\bar{\mathcal{G}}(T_n, T_{n-1}, \mathbf{U}_{n-1}^{(k)})$ *be an approximate solution to* (4.3)–(4.4) *with local truncation error bounded by* $\bar{C}_3\Delta T^{l+1}$, *which can be expanded for* ΔT *small as*
  (4.18)
  $$\bar{\mathcal{F}}(T_n, T_{n-1}, \boldsymbol{U}) - \bar{\mathcal{G}}(T_n, T_{n-1}, \boldsymbol{U}) = \bar{c}_{l+1}(\boldsymbol{U})\Delta T^{l+1} + \bar{c}_{l+2}(\boldsymbol{U})\Delta T^{l+2} + \cdots$$
  with continuously differentiable functions $\bar{c}_i$, $i = l+1, l+2, \ldots$, *and where* $\bar{\mathcal{F}}(T_n, T_{n-1}, \boldsymbol{U})$ *denotes the exact solution to* (4.3) *at* T_n, *starting from the initial value* $\boldsymbol{U}$;
- *assume* $\bar{\mathcal{G}}$ *satisfies the Lipschitz condition*
  (4.19)
  $$\|\bar{\mathcal{G}}\big(t+\Delta T, t, \boldsymbol{U}\big) - \bar{\mathcal{G}}\big(t+\Delta T, t, \boldsymbol{V}\big)\| \le (1 + C_2\Delta T)\|\boldsymbol{U} - \boldsymbol{V}\|$$
  for $t \in \mathcal{I}$ *and for all* $\boldsymbol{U}$, $\boldsymbol{V}$.

Then at iteration k, *the new Parareal algorithm* (4.5)–(4.6) *satisfies the error bound*

$$\begin{aligned}&\|\boldsymbol{u}(T_n) - \boldsymbol{U}_n^{(k)}\|\\ &\le \bar{C}_1^k\left[C_4C_p\Delta T^{(l+1)k+1/q} + \bar{C}_3\left(\Delta T^{l+1}\right)^{k+1}\right]\frac{(1+C_2\Delta T)^{n-k-1}}{(k+1)!}\prod_{j=0}^{k}(n-j)\end{aligned} \tag{4.20}$$

with the integer $q \ge 1$ *defined by the relation* $1/p + 1/q = 1$, *constants* C_p *and* C_4 *from Lemma* 4.3, *and* $\bar{C}_1 > 0$ *determined by the Lipschitz constant of* $\bar{c}_{l+1}$ *and the expansion* (4.18).

Proof. By adding and subtracting the same terms, we obtain from the new Parareal update formula for the error of (4.6)

$$
\begin{aligned}
&\boldsymbol{u}(T_n) - \boldsymbol{U}_n^{(k+1)} \\
&\quad = \mathcal{F}\left(T_n, T_{n-1}, \boldsymbol{u}(T_{n-1})\right) - \mathcal{F}\left(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(k)}\right) \\
&\qquad + \bar{\mathcal{G}}\left(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(k)}\right) - \bar{\mathcal{G}}\left(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(k+1)}\right) \\
&\qquad \pm \bar{\mathcal{F}}\left(T_n, T_{n-1}, \boldsymbol{u}(T_{n-1})\right) \pm \bar{\mathcal{G}}\left(T_n, T_{n-1}, \boldsymbol{u}(T_{n-1})\right) \pm \bar{\mathcal{F}}\left(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(k)}\right) \\
&\quad = \underbrace{\mathcal{F}\left(T_n, T_{n-1}, \boldsymbol{u}(T_{n-1})\right) - \bar{\mathcal{F}}\left(T_n, T_{n-1}, \boldsymbol{u}(T_{n-1})\right)}_{=\boldsymbol{\epsilon}_{f,n}(T_n)} \\
&\qquad + \underbrace{\bar{\mathcal{F}}\left(T_n, T_{n-1}, \boldsymbol{u}(T_{n-1})\right) - \bar{\mathcal{G}}\left(T_n, T_{n-1}, \boldsymbol{u}(T_{n-1})\right)}_{=\bar{c}_{l+1}(\boldsymbol{u}(T_{n-1}))\Delta T^{l+1}+\cdots} \\
&\qquad - \underbrace{\left(\mathcal{F}\left(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(k)}\right) - \bar{\mathcal{F}}\left(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(k)}\right)\right)}_{=\boldsymbol{\epsilon}_{f,n}(T_n)} \\
&\qquad - \underbrace{\left(\bar{\mathcal{F}}\left(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(k)}\right) - \bar{\mathcal{G}}\left(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(k)}\right)\right)}_{=\bar{c}_{l+1}\left(\boldsymbol{U}_{n-1}^{(k)}\right)\Delta T^{l+1}+\cdots} \\
&\qquad + \bar{\mathcal{G}}\left(T_n, T_{n-1}, \boldsymbol{u}(T_{n-1})\right) - \bar{\mathcal{G}}\left(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(k+1)}\right). \qquad (4.21)
\end{aligned}
$$

Using the Lipschitz continuity of $\bar{c}_{l+1}$ and the Lipschitz condition (4.19), we obtain the bound

$$
\|\boldsymbol{u}(T_n) - \boldsymbol{U}_n^{(k+1)}\| \le \bar{C}_1 \Delta T^{l+1} \|\boldsymbol{u}(T_n) - \boldsymbol{U}_{n-1}^{(k)}\| + (1 + C_2 \Delta T) \|\boldsymbol{u}(T_{n-1}) - \boldsymbol{U}_{n-1}^{(k+1)}\| \qquad (4.22)
$$

with a positive constant $\bar{C}_1$. In order to obtain a bound on the error, we now consider the corresponding recurrence relation

$$
e_n^{k+1} = \alpha e_{n-1}^{k} + \beta e_{n-1}^{k+1} \qquad (4.23)
$$

with $\alpha = \bar{C}_1 \Delta T^{l+1}$ and $\beta = 1 + C_2 \Delta T$. Due to the initial guess from the coarse propagator (4.8), the initial error can be estimated for $n = 1, \ldots, N$ by

$$
\begin{aligned}
\|\boldsymbol{u}(T_n) - \boldsymbol{U}_n^{(0)}\| &= \left\| \mathcal{F}\left(T_n, T_{n-1}, \boldsymbol{u}(T_{n-1})\right) - \bar{\mathcal{G}}\left(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(0)}\right) \right\| \\
&\le \| \mathcal{F}\left(T_n, T_{n-1}, \boldsymbol{u}(T_{n-1})\right) - \bar{\mathcal{G}}\left(T_n, T_{n-1}, \boldsymbol{u}(T_{n-1})\right) \| \\
&\quad + \left\| \bar{\mathcal{G}}\left(T_n, T_{n-1}, \boldsymbol{u}(T_{n-1})\right) - \bar{\mathcal{G}}\left(T_n, T_{n-1}, \boldsymbol{U}_{n-1}^{(0)}\right) \right\| \\
&\le \| \mathcal{F}\left(T_n, T_{n-1}, \boldsymbol{u}(T_{n-1})\right) - \bar{\mathcal{F}}\left(T_n, T_{n-1}, \boldsymbol{u}(T_{n-1})\right) \| \\
&\quad + \bar{C}_3 \Delta T^{l+1} + (1 + C_2 \Delta T) \|\boldsymbol{u}(T_n) - \boldsymbol{U}_{n-1}^{(0)}\|.
\end{aligned}
$$

Now Lemma 4.3 gives us a bound for the first term on the RHS above, and we thus obtain for the bounding initial recurrence relation

$$
e_n^0 = \gamma + \beta e_{n-1}^0, \qquad (4.24)
$$

where $\gamma := C_4 C_p \Delta T^{1/q} + \bar{C}_3 \Delta T^{l+1}$. We can now follow the same reasoning as in [17] to obtain the estimate (4.20). Multiplying (4.23) and (4.24) by ζ^n and summing over $n \geq 1$, we obtain that the generating function $\rho_k(\zeta) := \sum_{n\geq 1} e_n^k \zeta^n$ satisfies the recurrence relation

$$\rho_{k+1}(\zeta) = \alpha\zeta\rho_k(\zeta) + \beta\zeta\rho_{k+1}(\zeta), \quad \rho_0(\zeta) = \gamma\frac{\zeta}{1-\zeta} + \beta\zeta\rho_0(\zeta).$$

Solving for $\rho_k(\zeta)$, we get by induction

$$\rho_k(\zeta) = \gamma\alpha^k \frac{\zeta^{k+1}}{(1-\zeta)} \frac{1}{(1-\beta\zeta)^{k+1}}.$$

Coefficients in the power series of $\rho_k(\zeta)$ will increase when the factor $1-\zeta$ in the denominator is replaced by $1-\beta\zeta$. Then using the binomial series expansion

$$\frac{1}{(1-\beta\zeta)^{k+2}} = \sum_{j\geq 0} \binom{k+1+j}{j} \beta^j \zeta^j,$$

one can derive from $\rho_k(\zeta)$ the following bound for e_n^k,

$$e_n^k \leq \gamma\alpha^k \beta^{n-k-1} \binom{n}{k+1},$$

which concludes the proof. ∎

Remark 4.5. We note that depending on the selected L^p-space the order of $\Delta T^{(l+1)k+1/q}$ attains a value from $[(l+1)k, (l+1)k+1]$. When $p=\infty$ the highest order $(l+1)k+1$ can be obtained, and the corollary below gives the error estimate for our new Parareal algorithm (4.5)–(4.6) in that case.

COROLLARY 4.6. *Let the assumptions of Theorem* 4.4 *be satisfied. If* $\tilde{\boldsymbol{f}} \in \boldsymbol{L}^\infty(\mathcal{I}, \mathbb{R}^n)$ *in Assumption* 2, *then the estimate* (4.20) *becomes*

(4.25)
$$\|\boldsymbol{u}(T_n) - \boldsymbol{U}_n^{(k)}\| \leq \bar{C}_1^k \left[C_4 C_\infty \Delta T^{(l+1)k+1} + \bar{C}_3 \left(\Delta T^{l+1}\right)^{k+1} \right] \frac{(1+C_2\Delta T)^{n-k-1}}{(k+1)!} \prod_{j=0}^{k} (n-j).$$

Proof. Using [21, Theorem 10.2] and boundedness of the vector norm $\|\tilde{\boldsymbol{f}}(t)\| \leq C_\infty$ on $\mathcal{I}$, we obtain the bound

$$\|\boldsymbol{\epsilon}_{f,n}(t)\| \leq \frac{C_\infty}{L}\left(\mathrm{e}^{L(t-T_{n-1})} - 1\right), \quad t \leq T_{n-1}. \tag{4.26}$$

For small ΔT this implies that there exists $C_4 > 0$ s.t.

$$\|\boldsymbol{\epsilon}_{f,n}(T_n)\| \leq C_4 C_\infty \Delta T, \tag{4.27}$$

and following the proof of Theorem 4.4, we obtain the estimate (4.25). ∎

Remark 4.7. From the convergence estimate (4.20), we see that if the norm $\|\tilde{\boldsymbol{f}}\|_{L^p(\mathcal{I},\mathbb{R}^n)}$ in Assumption 2 is small enough, then the second term in the estimate (4.20) will dominate initially, and the convergence rate will be as for the original Parareal algorithm, where coarse and fine propagators both solve the same problem. This explains the key innovation in our new Parareal algorithm, namely, to use a suitable smooth input $\bar{\boldsymbol{f}}$ for our new coarse propagator $\bar{\mathcal{G}}$, in order to avoid a considerable reduction of the Parareal convergence order.

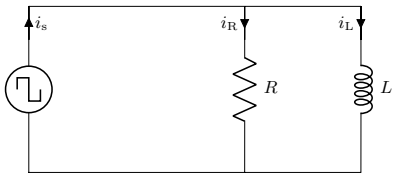

FIG. 2. *RL-circuit model.*

Remark 4.8. We would like to note that the key difference between the convergence results derived in Theorem 4.4 and the standard Parareal error estimate from Theorem 3.1 lies in the introduction of the additional operator $\bar{\mathcal{F}}$. In fact, the bound for the error at iteration $k+1$ given in (4.22) is the same for both methods. The additional term of order $\Delta T^{1/q}$ in the error estimate (4.20) compared to (3.7) comes only from the initial error e_n^0 in (4.24). Therefore, Theorem 4.4 can be considered to be a natural extension of the standard result from Theorem 3.1. Indeed, applying Theorem 4.4 to a problem with a smooth RHS would lead directly to the original error estimate (3.7) via assigning $\tilde{\boldsymbol{f}}(t) := \mathbf{0}$ in (2.3).

5. Numerical experiments for a model problem. We now compare the performance of our new Parareal algorithm to the one of the original Parareal algorithm, and test the accuracy of our error estimates on the model of the RL-circuit shown in Figure 2. The equations for this circuit are

$$\begin{aligned} \frac{1}{R}\phi'(t) + \frac{1}{L}\phi(t) &= f_m(t), \quad t \in (0,T], \\ \phi(0) &= 0, \end{aligned} \tag{5.1}$$

where $R = 0.01\ \Omega$ is the resistance, $L = 0.001$ H denotes the inductance, $T = 0.02$ s is the period, and f_m is the supplied PWM current source (in A) with m denoting the number of pulses, i.e.,

$$f_m(t) = \begin{cases} \operatorname{sign}\left[\sin\left(\dfrac{2\pi}{T}t\right)\right], & s_m(t) - \left|\sin\left(\dfrac{2\pi}{T}t\right)\right| < 0, \\ 0, & \text{otherwise}, \end{cases} \tag{5.2}$$

where $s_m(t) = \dfrac{m}{T}t - \lfloor \dfrac{m}{T}t \rfloor$, $t \in [0,T]$ is the common sawtooth pattern. In Figure 1 we showed already the PWM of switching frequency $f_s = m/T = 500$ Hz, which consists of $m = 10$ pulses. Note that the values, which the depicted PWM signal attains, are only $-1, 0, 1$. Our numerical tests deal with the base frequency of 50 Hz and a modulation of 20 kHz ($m = 400$), which is practically relevant in many applications in electrical engineering, e.g., for a DC-AC converter described in [34]. Details about PWM signals and their frequencies used for power semiconductor devices can be found in [5, 32]

5.1. Performance of the original Parareal algorithm. In the original Parareal algorithm, both the fine and the coarse problem use the PWM signal (5.2). For a small number of processors, $N \ll m$, the coarse propagator will not resolve the dynamics of the excitation and, therefore, the original convergence arguments are not applicable: Theorem 3.1 is valid only for N large enough, when the coarse propagator resolves all the pulses and the function is locally smooth, and only in this regime one can expect that the high convergence rate of the original Parareal algorithm is

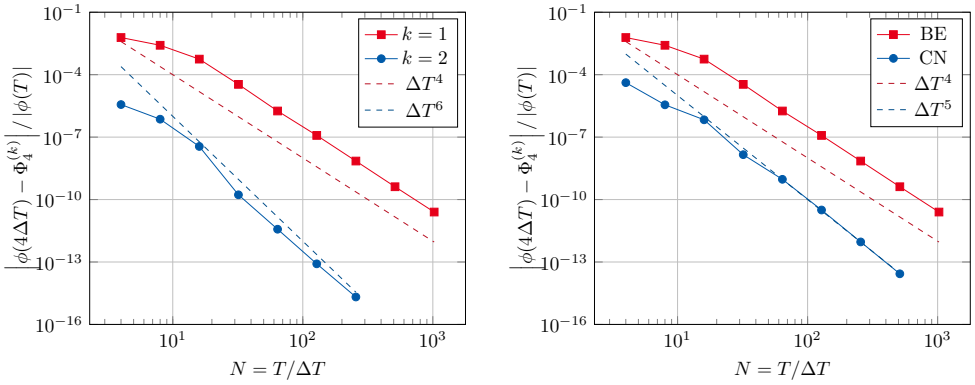

FIG. 3. *Dependence on N of the convergence of the original Parareal algorithm. Left: for $k = 1$ using BE, where we clearly see order reduction for $N < 20$, and the asymptotic convergence order is only reached for larger N. For $k = 2$, however, order reduction can be observed also for bigger N. Right: for $k = 1$ using CN, where the order is reduced for all N, in contrast to BE for $k = 1$. Note, BE for $k = 1$ is shown in both plots for reference.*

maintained. This is illustrated in Figure 3 for backward Euler (BE) on the left, where we see that for large N we obtain fourth order convergence for $k = 1$ which matches the prediction $(l + 1)(k + 1)$ in (3.7) for the BE method of order $l = 1$. However, for small N (less than 20), the convergence order is much lower. Convergence slowdown is even more apparent for $k = 2$, since in that case order reduction remains even for larger N. On the right in Figure 3, we show the corresponding results for the Crank–Nicolson (CN) scheme, which is of order $l = 2$, and we iterate only once, $k = 1$. Here we observe that the asymptotic order is reduced to 5 instead of the predicted order 6 by the original Parareal estimate in (3.7) for the case of smooth input

5.2. Performance of the new Parareal algorithm. We now test our new Parareal algorithm using two choices of input for the coarse propagator with reduced dynamics. On the one hand, one could make the naive choice of a step function

$$\bar{f}_{\text{step}}(t) = \begin{cases} 1, & t \in [0, T/2], \\ -1, & t \in (T/2, T] \end{cases} \tag{5.3}$$

on $[0, T]$. This is not globally smooth but piecewise, which suffices, since we consider in the following experiments only single step time stepping methods that restart at $T/2$.

On the other hand, in power engineering, the PWM is commonly used as a cheap surrogate for sinusoidal excitation. Therefore, its first and dominant harmonic, i.e., the sine wave

$$\bar{f}_{\text{sin}}(t) = \sin\left(\frac{2\pi}{T}t\right), \quad t \in [0, T], \tag{5.4}$$

is a more reasonable choice for the coarse problem. The IVP with reduced dynamics for our model problem is defined by

$$\begin{aligned} \frac{1}{R}\phi'(t) + \frac{1}{L}\phi(t) &= \bar{f}(t), \quad t \in (0, T], \\ \phi(0) &= 0 \end{aligned} \tag{5.5}$$

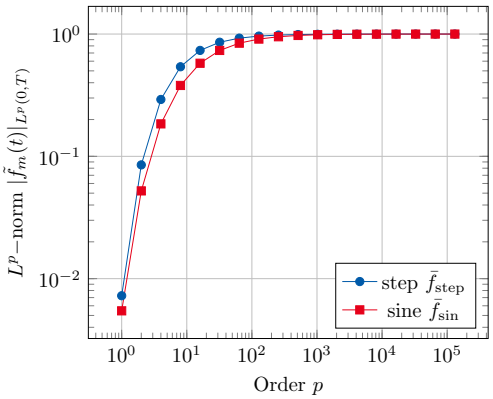

FIG. 4. *L^p-norm $|\tilde{f}_m(t)|_{L^p(0,T)} := |f_m(t) - \bar{f}(t)|_{L^p(0,T)}$ for the sine and the step coarse inputs.*

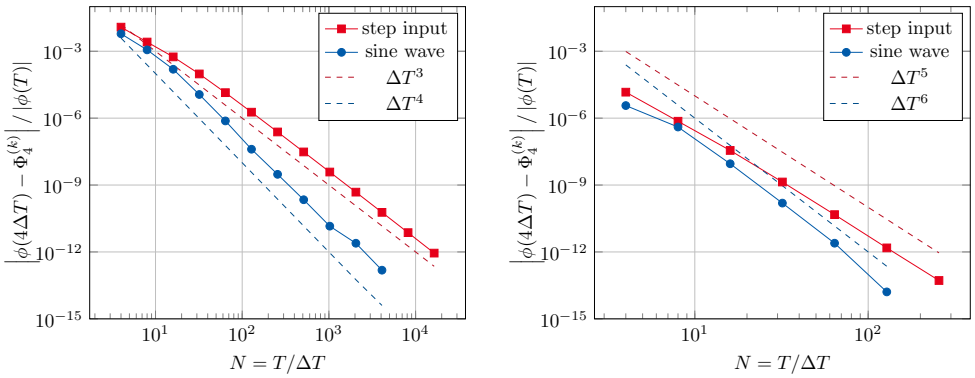

FIG. 5. *Dependence on N of the convergence of the new Parareal algorithm using BE and the coarse propagators with reduced dynamics* (5.3) *and* (5.4). *Left: for $k = 1$. Right: for $k = 2$.*

with $\bar{f}$ being one of the functions in (5.3) or (5.4). The coarse propagator $\bar{\mathcal{G}}$ will solve the problem (5.5), while the fine propagator $\mathcal{F}$ will solve the original problem (5.1). The non-smooth part of the input is then given by

$$\tilde{f}_m(t) := f_m(t) - \bar{f}(t). \tag{5.6}$$

Clearly, $|\tilde{f}_m(t)| \in L^\infty(0,T)$ for both sinusoidal and step coarse input. This leads to the freedom in the choice of $p \geq 1$ in the norm $|\tilde{f}_m(t)|_{L^p(0,T)}$, which in turn influences the convergence rate due to the estimate (4.20) and Remark 4.5. In fact, it could be seen that $|\tilde{f}_m(t)|_{L^\infty(0,T)} = 1$ for both sinusoidal and step coarse inputs. However, based on the calculated values of the norms presented in Figure 4 we could conclude that for any $1 \leq p < \infty$ the L^p-norm of $\tilde{f}_m$ is smaller in the case of sine than in the case of the step. These data imply that the sinusoidal waveform is a superior candidate for the coarse input than the step function and, therefore, one could expect a higher convergence rate in that case.

We show in Figure 5 a comparison of the convergence behavior of the new Parareal algorithm using BE for $k = 1$ and $k = 2$ iterations using the two different choices of reduced input dynamics. We see that in both cases when the reduced dynamics of the step function $\bar{f}_{\text{step}}$ in (5.3) is used for the coarse propagator, one obtains an order

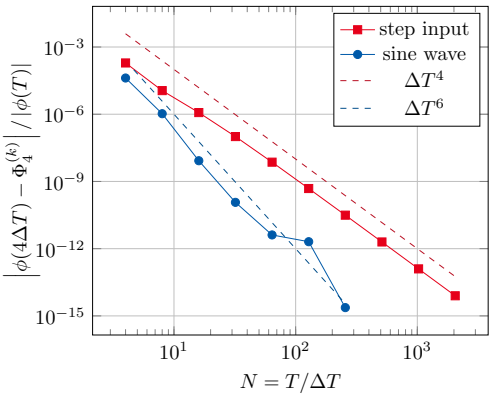

FIG. 6. *Dependence on N of the convergence of the new Parareal algorithm using CN and the coarse propagators with reduced dynamics* (5.3) *and* (5.4) *for $k = 1$.*

reduction: for $k = 1$ we get third order, and for $k = 2$ we get fifth order, which matches the theoretical predictions because the lower order term in (4.20) could have the maximal order $(l+1)k+1 = 3$ for $k = 1$ and $(l+1)k+1 = 5$ for $k = 2$. On the other hand, convergence of order $(l+1)(k+1) = 4$ for $k = 1$ (left) and $(l+1)(k+1) = 6$ for $k = 2$ (right) is observed for the coarse sine input $\bar{f}_{\sin}$, given in (5.4), which means that indeed the second term $\bar{C}_3(\Delta T^{l+1})^{k+1}$ in our estimate (4.20) is dominant over the first one. Hence, the sinusoidal function appears to be a well-chosen reduced dynamics for the coarse problem, which does not slow down the convergence of the Parareal algorithm, as the bound in (3.7) gives the same rate.

We next test CN with our new Parareal algorithm. For one iteration, $k = 1$, we show in Figure 6 how in this case the step input function $\bar{f}_{\text{step}}$ also gives order reduction, we only observe fourth order convergence, which is in good agreement with our convergence estimate since the first term in (4.20) is in the best case of order $(l+1)k+1 = 4$, whereas with the sine input function $\bar{f}_{\sin}$ we get as expected the full sixth order convergence.

Remark 5.1. In the current section we have observed that the choice of the coarse input plays an important role in the convergence rate of our new Parareal approach. A suitable smooth function can be selected based on, e.g., Fourier analysis. Indeed, application of the Fourier transform to a given excitation gives the information about the magnitudes of its harmonic coefficients. The norm $\|\tilde{\boldsymbol{f}}\|_{L^p(\mathcal{I},\mathbb{R}^n)}$ could be small enough when there are harmonics with dominating magnitutes compared to the rest of the Fourier coefficients. However, we note that the proposed Parareal method converges for any choice of the smooth coarse input.

5.3. A nonlinear example. Now we would like to investigate the performance of our Parareal method in the nonlinear case. For that reason we consider a nonlinear model problem, similar to the one described previously. In particular, we solve the IVP

$$\begin{aligned} R^{-1}\phi'(t) + L^{-1}(|\phi|)\phi(t) &= f_m(t), \quad t \in (0,T], \\ \phi(0) &= 0 \end{aligned} \tag{5.7}$$

with the resistance $R = 0.01\ \Omega$ and a flux-dependent inverse of inductance $L^{-1}(|\phi|) := k_{\text{L}}|\phi(t)|$, where $k_{\text{L}} = 10^3\ \text{A/Wb}^2$. The PWM current source f_m is defined in (5.2) and

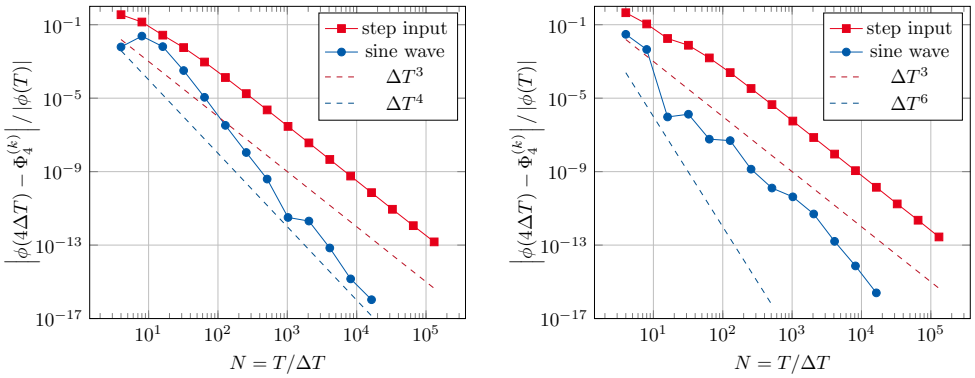

FIG. 7. *Dependence on N of the convergence of the new Parareal algorithm using BE (left) and CN (right) and the coarse propagators with reduced dynamics* (5.3) *and* (5.4) *for $k = 1$.*

has the switching frequency of 20 kHz. As in (5.1) we consider the period $T = 0.02$ s. Convergence rates obtained after one iteration of Parareal with the sinusoidal (5.4) and step (5.3) coarse inputs are shown in Figure 7.

When applying the BE method as the time integrator (of order $l = 1$) we observe a faster convergence rate when using the sinusoidal coarse input, compared to that obtained with the step function. In particular, Parareal equipped with the sine has order close to 5, which is higher than $(l+1)(k+1) = 4$ for almost all the tested values of ΔT, whereas the step coarse input leads to order reduction by giving the order $(l+1)k+1 = 3$. A similar situation is also observed in the case of the CN scheme. The sinusoidal waveform on the coarse level allows Parareal to converge with order close to $(l+1)(k+1) = 6$. In contrast to this, the step excitation causes convergence of order $(l+1)k = 3$, which is the lowest convergence rate of our Parareal method due to the estimate (4.20). These results support our hypothesis that the sine (5.4) is a more beneficial choice of the coarse excitation than (5.3), as we have also concluded for the linear example.

6. Application to an induction machine. Due to the low-frequency operating regime of electrical machines, their simulation is usually performed assuming that the displacement current density is negligible with respect to the other current densities [36], and one derives a parabolic-elliptic initial boundary value problem from Maxwell's equations [24]. This is called the eddy current problem and it reads in terms of the magnetic vector potential $\vec{A} : \Omega \times \mathcal{I} \to \mathbb{R}^3$,

$$\sigma \partial_t \vec{A}(\vec{r},t) + \nabla \times \left(\nu \nabla \times \vec{A}(\vec{r},t)\right) = \vec{J}_{\text{src}}(\vec{r},t) \qquad \text{in } \Omega \times \mathcal{I}, \tag{6.1}$$

$$\vec{n} \times \vec{A}|_\Gamma = 0 \qquad \text{on } \Gamma \times \mathcal{I}, \tag{6.2}$$

$$\vec{A}(\vec{r},t_0) = \vec{A}_0(\vec{r}), \qquad \vec{r} \in \Omega, \tag{6.3}$$

where Ω represents the spatial domain of the machine, consisting of a rotor, a stator, and the air gap in-between, $\Gamma = \partial\Omega$ denotes its boundaries, and $\mathcal{I} := (t_0, t_{\text{end}}]$ is the time interval. The geometry is encoded in the scalar-valued electric conductivity $\sigma = \sigma(\vec{r}) \geq 0$ and the magnetic reluctivity $\nu = \nu(\vec{r}, \|\nabla \times \vec{A}\|) > 0$. The source current density

$$\vec{J}_{\text{src}} = \sum_{s=1}^{n_{\text{src}}} \vec{\chi}_s i_s$$

impresses lumped currents due to an attached electric network in terms of the winding functions $\vec{\chi}_s : \Omega \to \mathbb{R}^3$ which homogeneously distribute the currents $i_s : \mathcal{I} \to \mathbb{R}$ among $n_{\mathrm{src}} = 3$ stranded conductors [37], since we deal with a three-phase excitation within this application. The electric circuit establishes a relation between the current i_s and the voltage

$$v_s(t) = R_s i_s(t) + \int_\Omega \vec{\chi}_s(\vec{r}) \cdot \partial_t \vec{A}(\vec{r}, t) \, \mathrm{d}\Omega \tag{6.4}$$

with $s = 1, 2, 3$ and R_s denoting the DC resistance of the sth stranded conductor.

Furthermore, in order to include the rotation of the motor, the equation of motion is additionally considered: the movement is represented in the mesh by the moving band approach [12]. The angular velocity of the rotor,

$$\omega(t) = \mathrm{d}_t \theta(t), \quad t \in \mathcal{I}, \tag{6.5}$$

with a given initial rotor angle $\theta(t_0) = \theta_0$ can be determined via

$$I \mathrm{d}_t \omega + C\omega = T_{\mathrm{mag}}(\vec{A}) \quad \text{in } \mathcal{I}, \tag{6.6}$$

$$\omega(t_0) = \omega_0, \tag{6.7}$$

where I is the moment of inertia, and C is the friction coefficient. System (6.6)–(6.7) is excited with the torque T_{mag}, which is defined on the boundary of the air gap.

We consider in the following a two-dimensional (2D) computational domain $\Omega_{\mathrm{2D}} \subset \mathbb{R}^2$, which represents the cross section of the electrical machine. The reduction to the 2D setting and discretization of (6.1)–(6.3) using finite elements with n_{a} degrees of freedom gives together with (6.4), (6.5), and (6.6) an IVP for a coupled system of differential-algebraic equations (DAEs) of the form

$$\mathbf{M} \mathrm{d}_t \mathbf{u}(t) + \mathbf{K}\left(\mathbf{u}(t)\right) \mathbf{u}(t) = \mathbf{f}(t), \quad t \in \mathcal{I}, \tag{6.8}$$

$$\mathbf{u}(t_0) = \mathbf{u}_0 \tag{6.9}$$

with unknown $\mathbf{u} : \mathcal{I} \to \mathbb{R}^n$ and the initial condition $\mathbf{u}_0 \in \mathbb{R}^n$. The function $\mathbf{u}$ is a joint composition of several quantities such that $\mathbf{u}(t)^\top = [\mathbf{a}(t)^\top, \mathbf{i}(t)^\top, \theta(t), \omega(t)]$. Here the superscript $\top$ denotes the transpose of a vector. At each point t in time, $\mathbf{a}(t) \in \mathbb{R}^{n_{\mathrm{a}}}$ is the vector of (line-integrated) magnetic vector potentials, $\mathbf{i}(t) \in \mathbb{R}^3$ represents the currents of the three phases, $\theta(t) \in \mathbb{R}$ denotes the rotor angle, and $\omega(t) \in \mathbb{R}$ is the rotor's angular velocity ($n = n_{\mathrm{a}} + 5$). The differential-algebraic nature of the system (6.8) originates from the fact that $\mathbf{M}$ inherits the singularity from the finite element conductivity matrix due to the presence of nonconducting materials in the domain, i.e., where $\sigma = 0$. The RHS $\mathbf{f}(t)$ consists of given voltages $\mathbf{v}(t) \in \mathbb{R}^3$ and the mechanical excitation. We refer to [20] for details. Finally, the time-dependent problem (6.8)–(6.9) has to be solved via application of a time integrator.

Remark 6.1. We note that the DAE (6.8) is not immediately covered by our analysis for ODEs, but we do not see difficulties for the case of index-1 problems due the reasoning in [38].

6.1. Numerical model. We will now illustrate the performance of our new Parareal algorithm (4.5)–(4.6) for the semidiscrete eddy current problem (6.8)–(6.9), supplied with a three-phase PWM voltage source. As a concrete example we consider a four-pole squirrel-cage induction motor, illustrated in Figure 8, and carry out the

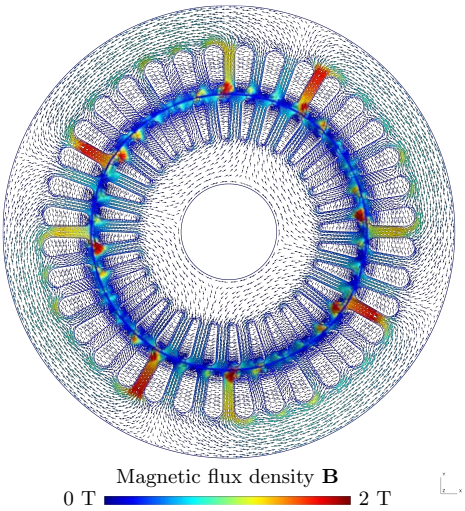

FIG. 8. *Magnetic field of the four-pole induction machine model "im_3kw"* [20] *at time instant* $t = 0.02$ *s if excited by a sinusoidal voltage excitation. The numerical simulation with GetDP* [19] *considers only a quarter of the machine geometry with periodic boundary conditions.*

computations under a no-load operation condition. The simulation of the 2D machine model was performed using the GetDP library [19] using $n = 4400$ degrees of freedom. The machine is supplied with a three-phase PWM voltage source of 20 kHz, which corresponds to $m = 400$ pulses on the time interval $[0,\ 0.02]$ s, and is practically relevant for numerous applications in electrical engineering. For $t \in \mathcal{I}$ and $s = 1, 2, 3$, the excitation (in V) with m pulses is given by

$$v_s^m(t) = \operatorname{sign}\left[\sin\left(\frac{2\pi}{T}t + \varphi_s\right) - b_m(t)\right], \tag{6.10}$$

where φ_s denotes one of the three phases $\varphi_1 = 0$, $\varphi_2 = -2/3\pi$, $\varphi_3 = -4/3\pi$, and

$$b_m(t) = 2\left(\frac{m}{T}t - \left\lfloor\frac{m}{T}t\right\rfloor\right) - 1 \tag{6.11}$$

is determined by the bipolar trailing-edge modulation using a sawtooth carrier [40].

We consider $T = 0.02$ s to be the electric period, which corresponds to a frequency of 50 Hz. As an example of the voltage source, a PWM signal v_1^{100} of 5 kHz (corresponding to $m = 100$ pulses on $[0, 0.02]$ s) is shown in Figure 9. An initial ramp up of the applied voltage was used for reducing the transient behavior of the motor, as was proposed by the original authors of the model [20]. Phase 1 of the three-phase sinusoidal voltage source of 50 Hz is shown in Figure 9. This waveform will be used as an input for the reduced coarse problem within our new Parareal method.

The current waveforms, obtained by solving the DAE (6.8) excited by the PWM signal of 20 kHz, are shown in Figure 10. The waveform has multiharmonic characteristics; one observes three different time scales: the underlying sinusoidal excitation (50 Hz), an additional sinusoidal behavior due to "slotting" of the machine, and finally the small high-frequency oscillations ("ripples") in the current waveforms due to the PWM excitations (20 kHz). The consideration of all those frequencies may

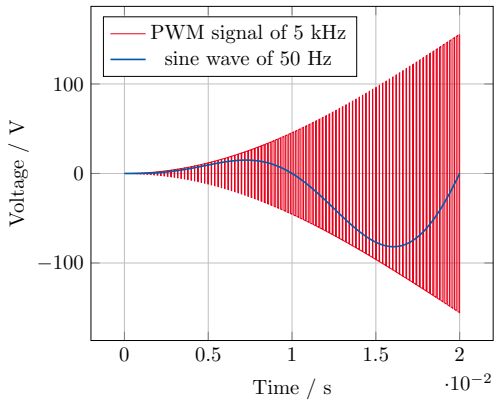

FIG. 9. *PWM voltage source of* 5 *kHz with a ramp up and phase* 1 *of the corresponding sinusoidal waveform of* 50 *Hz.*

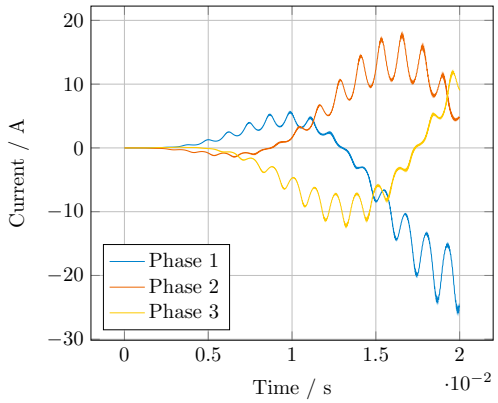

FIG. 10. *Stator current waveforms for the three-phase PWM voltage source of* 20 *kHz.*

be, for example, important if an engineer is concerned with the acoustic design of the machine.

6.2. Application of the new Parareal algorithm. The new Parareal algorithm (4.5)–(4.6) was implemented in GNU Octave [10] and uses OpenMP parallelized calls of GetDP. The code is available in [25]. It was executed on an Intel Xeon cluster with 80×2.00 GHz cores and 1 TB DDR3 memory.

The reduced coarse propagator $\bar{\mathcal{G}}$ solves (6.8)–(6.9) with the input three-phase voltage

$$\bar{v}_s(t) = \sin\left(\frac{2\pi}{T}t + \varphi_s\right), \quad s = 1, 2, 3, \tag{6.12}$$

for $t \in \mathcal{I} = [0,\ 0.02]$, shown in Figure 11. The fine solver $\mathcal{F}$ uses the original PWM input v_s^{400}, $s = 1, 2, 3$, from (6.10). Both propagators solve the IVP using the BE method with the time step sizes $\Delta t = 10^{-3}$ s and $\delta t = 10^{-6}$ s for the coarse and the fine problems, respectively. We have used $N = 20$ cores within the new Parareal simulation, and for comparison we also simulated the machine with the original Parareal algorithm (3.3)–(3.4). In the original Parareal algorithm both coarse

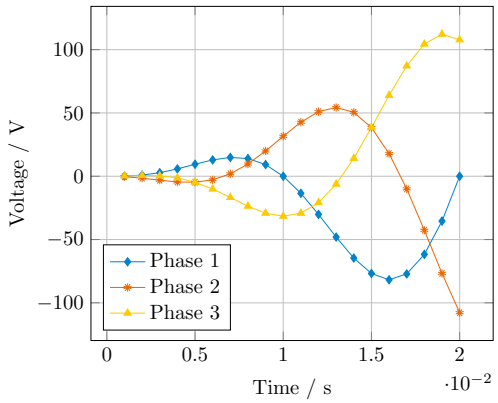

FIG. 11. *Three-phase sinusoidal voltage source of 50 Hz, used as an input for the coarse propagator with reduced dynamics in our new Parareal algorithm. The coarse discretization is obtained with the time step* $\Delta t = 10^{-3}$.

and fine problems have the same PWM voltage input v_s^{400}, $s = 1, 2, 3$, and we use the same time step sizes Δt and δt defined above.

In order to evaluate the convergence of the original and the new Parareal algorithms, we used the error norm from [21, Chapter II.4], i.e., the vector $\mathbf{u}$ is considered close to $\mathbf{v} \in \mathbb{R}^n$ if

$$\text{err} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \frac{|u_i - v_i|^2}{\left(\text{atol} + \text{rtol}\,|v_i|\right)^2}} < 1, \tag{6.13}$$

where $\text{atol} = \text{rtol} = 1.5 \cdot 10^{-5}$ are prescribed absolute and relative tolerances. The error norm (6.13) is applied to each jump at the $N - 1$ synchronization points. The Parareal iteration is terminated if the mismatch of the biggest jump, measured by (6.13), is below 1.

The numerical results in Figure 12 show that the new Parareal algorithm with well-chosen reduced coarse input works very well in practice: at each iteration, it is about one order of magnitude more accurate, and also needs, in this example, 25% fewer iterations than the original Parareal algorithm to converge to the desired tolerance, capturing well all relevant frequencies of the multiharmonic solution. Furthermore, we have estimated computational costs of both Parareal algorithms in terms of the required number of linear solutions. In particular, the standard Parareal converged in 8 iterations, effectively requiring 156 309 sequential solutions of linearized systems of algebraic equations. In contrast to this, the new Parareal reached the prescribed tolerance after 6 iterations, thereby needing 121 264 linear system solves. For completeness of the comparison we also calculated that within the sequential simulation on $[0, 0.02]$ s 244 104 linear systems were solved. These data illustrate that the Parareal approach, equipped with the smoothed coarse excitation, converges faster than the Parareal with the PWM source on both fine and coarse levels, and it also gives a speed-up of factor 2 compared to the standard time integration.

7. Conclusions. In this paper we proposed a new Parareal algorithm for problems which involve discontinuous inputs. Our new Parareal algorithm uses a smooth part of the source representing reduced dynamics for the coarse propagator, which is suitable for coarse discretization in time. We analyzed the new Parareal algorithm

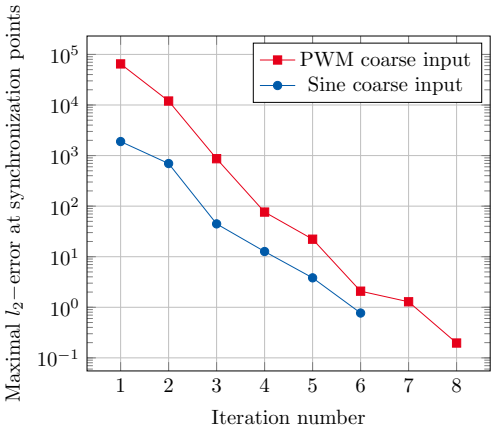

FIG. 12. *Comparison of the convergence of the original and the new Parareal algorithms.*

and derived precise error estimates, which show that order reduction is possible if the coarse model is not good enough. In particular, if the chosen nonsmooth part $\tilde{\boldsymbol{f}}$ of the input belongs to $L^p(\mathcal{I}, \mathbb{R}^n)$ with $p \geq 1$, and a time integrator of order l is used as the coarse propagator for the reduced problem, we proved that the order reduction can be at most $l + 1/p$. However, if the corresponding input with reduced dynamics is a good approximation to the original (discontinuous) input, i.e., they are close in the sense of the L^p-norm, then the new Parareal algorithm with a coarse propagator using large time steps reaches the same order as the original Parareal algorithm that would need a coarse propagator with very small time steps. We illustrated the accuracy of our estimates with numerical experiments on an RL-circuit model as well as for a nonlinear model problem with PWM signal as input. We also tested the new Parareal algorithm on an eddy current problem, describing the operation of an induction machine. The new Parareal algorithm is about an order of magnitude more accurate in each iteration than the original one, and reaches a prescribed tolerance in 25% fewer iterations in the eddy current example.

REFERENCES

[1] M. ASTORINO, F. CHOULY, AND A. QUARTERONI, *Multiscale Coupling of Finite Element and Lattice Boltzmann Methods for Time Dependent Problems*, preprint, https://hal.archives-ouvertes.fr/hal-00746942v1 (2012).

[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, https://doi.org/10.1103/PhysRevE.66.057701.

[3] G. BAL, *Parallelization in Time of (Stochastic) Ordinary Differential Equations*, manuscript.

[4] G. I. BARENBLATT AND A. J. CHORIN, *New perspectives in turbulence: Scaling laws, asymptotics, and intermittency*, SIAM Rev., 40 (1998), pp. 265–291.

[5] B. K. BOSE, *Power Electronics And Motor Drives*, Academic Press, Burlington, MA, 2006.

[6] Y. BOUBENDIR, X. ANTOINE, AND C. GEUZAINE, *A quasi-optimal non-overlapping domain decomposition algorithm for the Helmholtz equation*, J. Comput. Phys., 231 (2012), pp. 262–280.

[7] F. CHEN, J. S. HESTHAVEN, AND X. ZHU, *On the use of reduced basis methods to accelerate and stabilize the Parareal method*, in Reduced Order Methods for Modeling and Computational Reduction, Springer, Cham, Switzerland, 2014, pp. 187–214, https://doi.org/10.1007/978-3-319-02090-7_7.

[8] F. Chouly and A. Lozinski, *Parareal multi-model numerical zoom for parabolic multiscale problems*, C. R. Math., 352 (2014), pp. 535–540, https://doi.org/10.1016/j.crma.2014.03.018.
[9] G. Dahlquist and Å. Björck, *Numerical Methods in Scientific Computing, Volume* I, SIAM, Philadelphia, 2008.
[10] J. W. Eaton, D. Bateman, S. Hauberg, and R. Wehbring, *GNU Octave (version* 5.1.0*)*, http://www.gnu.org/software/octave/doc/interpreter.
[11] S. Engblom, *Parallel in time simulation of multiscale stochastic chemical kinetics*, Multiscale Model. Simul., 8 (2009), pp. 46–68, https://doi.org/10.1137/080733723.
[12] M. V. Ferreira da Luz, P. Dular, N. Sadowski, C. Geuzaine, and J. P. A. Bastos, *Analysis of a permanent magnet generator with dual formulations using periodicity conditions and moving band*, IEEE Trans. Magn., 38 (2002), pp. 961–964, https://doi.org/10.1109/20.996247.
[13] A. F. Filippov, *Differential Equations with Discontinuous Righthand Sides: Control Systems*, Vol. 18, Springer, Berlin, 1988.
[14] P. F. Fischer, F. Hecht, and Y. Maday, *A Parareal in time semi-implicit approximation of the Navier-Stokes equations*, in Domain Decomposition Methods in Science and Engineering, R. Kornhuber et al., ed., Lect. Notes Comput. Sci. Eng. 40, Springer, Berlin, 2005, pp. 433–440, https://doi.org/10.1007/3-540-26825-1_44.
[15] M. J. Gander, *Optimized Schwarz methods*, SIAM J. Numer. Anal., 44 (2006), pp. 699–731.
[16] M. J. Gander, 50 *years of time parallel time integration*, in Multiple Shooting and Time Domain Decomposition Methods, T. Carraro, M. Geiger, S. Körkel, and R. Rannacher, eds., Contrib. Math. Comput. Sci. 9, Springer, Berlin, 2015, pp. 69–113, https://doi.org/10.1007/978-3-319-23321-5_3.
[17] M. J. Gander and E. Hairer, *Nonlinear convergence analysis for the Parareal algorithm*, in Domain Decomposition Methods in Science and Engineering XVII, Springer, Berlin, 2008, pp. 45–56, https://doi.org/10.1007/978-3-540-75199-1_4.
[18] M. J. Gander and S. Vandewalle, *Analysis of the Parareal time-parallel time-integration method*, SIAM J. Sci. Comput., 29 (2007), pp. 556–578.
[19] C. Geuzaine, *GetDP: A general finite-element solver for the de Rham complex*, PAMM Proc. Appl. Math, Mech. 7, Wiley (2007), pp. 1010603–1010604, https://doi.org/10.1002/pamm.200700750.
[20] J. Gyselinck, L. Vandevelde, and J. Melkebeek, *Multi-slice FE modeling of electrical machines with skewed slots-the skew discretization error*, IEEE Trans. Magn., 37 (2001), pp. 3233–3237, https://doi.org/10.1109/20.952584.
[21] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations* I: *Nonstiff Problems*, 2nd ed., Springer Ser. Comput. Math., Springer, Berlin, 1993.
[22] T. Haut and B. Wingate, *An asymptotic parallel-in-time method for highly oscillatory PDEs*, SIAM J. Sci. Comput., 36 (2014), pp. A693–A713, https://doi.org/10.1137/130914577.
[23] L. He, *The reduced basis technique as a coarse solver for Parareal in time simulations*, J. Comput. Math., 28 (2010), pp. 676–692, https://doi.org/10.4208/jcm.1003-m2980.
[24] J. D. Jackson, *Classical Electrodynamics*, 3rd ed., Wiley, New York, 1998.
[25] I. Kulchytska-Ruchka, S. Schöps, I. Niyonzima, and M. J. Gander, *A Parareal Implemenation in Matlab/Octave for Problems with Discontinuous Sources*, http://www.github.com/temf/parareal (2019).
[26] J.-L. Lions, Y. Maday, and G. Turinici, *A Parareal in time discretization of PDEs*, C. R. Acad. Sci. Ser. I Math., 332 (2001), pp. 661–668, https://doi.org/10.1016/S0764-4442(00)01793-6.
[27] J. Liu and Y.-L. Jiang, *A Parareal waveform relaxation algorithm for semi-linear parabolic partial differential equations*, J. Comput. Appl. Math., 236 (2012), pp. 4245–4263, https://doi.org/10.1016/j.cam.2012.05.014.
[28] Y. Maday, *The Parareal in time algorithm*, in Substructuring Techniques and Domain Decomposition Methods, Saxe-Coburg Publications, Kippen, Stirling, Scotland, 2010, pp. 19–44.
[29] Y. Maday, J. Salomon, and G. Turinici, *Monotonic Parareal control for quantum systems*, SIAM J. Numer. Anal., 45 (2007), pp. 2468–2482, https://doi.org/10.1137/050647086.
[30] Y. Maday and G. Turinici, *Parallel in time algorithms for quantum control: Parareal time discretization scheme*, Int. J. Quantum. Chem., 93 (2003), pp. 223–228, https://doi.org/10.1002/qua.10554.
[31] 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.
[32] N. Mohan, T. M. Undeland, and W. P. Robbins, *Power Electronics: Converters, Applications and Design*, 3 ed., Wiley, Hoboken, NJ, 2003.

[33] A. S. Nielsen, *Feasibility Study of the Parareal Algorithm*, Master's thesis, Technical University of Denmark, Kongens Lyngby, Denmark. 2012.
[34] K. Niyomsatian, P. Vanassche, R. V. Sabariego, and J. Gyselinck, *Systematic control design for half-bridge converters with LCL output filters through virtual circuit similarity transformations*, in 2017 IEEE Energy Conversion Congress and Exposition (ECCE), IEEE, Piscataway, NJ, 2017, pp. 2895–2902, https://doi.org/10.1109/ECCE.2017.8096535.
[35] A. Quarteroni and A. Valli, *Domain Decomposition Methods for Partial Differential Equations*, Numer. Math. Sci. Comput., Oxford University Press, Oxford, 1999.
[36] K. Schmidt, O. Sterz, and R. Hiptmair, *Estimating the eddy-current modeling error*, IEEE Trans. Magn., 44 (2008), pp. 686–689, https://doi.org/10.1109/TMAG.2008.915834.
[37] S. Schöps, H. De Gersem, and T. Weiland, *Winding functions in transient magnetoquasistatic field-circuit coupled simulations*, COMPEL, 32 (2013), pp. 2063–2083, https://doi.org/10.1108/COMPEL-01-2013-0004.
[38] S. Schöps, I. Niyonzima, and M. Clemens, *Parallel-in-time simulation of eddy current problems using Parareal*, IEEE Trans. Magn., 54 (2018), 7200604, https://doi.org/10.1109/TMAG.2017.2763090.
[39] G. Staff, *Convergence and Stability of the Parareal Algorithm*, Master's thesis, Norwegian University of Science and Technology, Trondheim, Norway, 2003.
[40] J. Sun, *Pulse-width modulation*, in Dynamics and Control of Switched Electronic Systems, Springer, London, 2012, pp. 25–61.
[41] A. Toselli and O. B. Widlund, *Domain Decomposition Methods—Algorithms and Theory*, Springer Ser. Comput. Math. 34, Springer, Berlin, 2005, https://doi.org/10.1007/b137868.
[42] J. M. F. Trindade and J. C. F. Pereira, *Parallel-in-time simulation of the unsteady Navier-Stokes equations for incompressible flow*, Internat. J. Numer. Methods Fluids, 45 (2004), pp. 1123–1136, https://doi.org/10.1002/fld.732.
[43] S.-L. Wu, *A second-order Parareal algorithm for fractional PDEs*, J. Comput. Phys., 307 (2016), pp. 280–290.
[44] S.-L. Wu, B. Shi, and C. Huang, *Parareal-Richardson algorithm for solving nonlinear ODEs and PDEs*, Commun. Comput. Phys., 6 (2009), pp. 883–902.
[45] E. Zeidler, *Nonlinear Functional Analysis and its Applications* II/A. *Linear Monotone Operators*, Springer, New York, 1990.