SIAM J. SCI. COMPUT.
Vol. 38, No. 6, pp. A3540–A3564

PARAREAL MULTISCALE METHODS FOR HIGHLY OSCILLATORY DYNAMICAL SYSTEMS*

GIL ARIEL†, SEONG JUN KIM‡, AND RICHARD TSAI§

Abstract. We introduce a new strategy for coupling the parallel in time (parareal) iterative methodology with multiscale integrators. Following the parareal framework, the algorithm computes a low-cost approximation of all slow variables in the system using an appropriate multiscale integrator, which is refined using parallel fine scale integrations. Convergence is obtained using an alignment algorithm for fast phase-like variables. The method may be used either to enhance the accuracy and range of applicability of the multiscale method in approximating only the slow variables, or to resolve all the state variables. The numerical scheme does not require that the system is split into slow and fast coordinates. Moreover, the dynamics may involve hidden slow variables, for example, due to resonances. We propose an alignment algorithm for almost-periodic solutions, in which case convergence of the parareal iterations is proved. The applicability of the method is demonstrated in numerical examples.

Key words. multiscale computation, parallel algorithms, highly oscillatory problems

AMS subject classifications. 34K28, 37N30

DOI. 10.1137/15M1011044

1. Introduction.

The parallel in time, also known as the "parareal" method, introduced by Lions, Maday and Turinici [52] is a simple yet effective scheme for the parallelization of numerical solutions for a large class of time dependent problems [53]. It consists of a fixed point iteration involving a coarse-but-cheap and a fine-but-expensive integrators. Computational time is reduced by parallelization of the fine integrations. For problems with separated multiple scales, it is tempting to apply a multiscale solver as a coarse integrator. So far, such types of parallel methods are limited to a few special multiscale cases such as chemical kinetics [16, 24, 34], dissipative ordinary differential equations (ODEs) [50], and highly oscillatory (HiOsc) problems in which the oscillatory behavior is relatively simple [20, 33]. One difficulty stems from a fundamental difference between the parareal and the multiscale philosophies—while the former requires pointwise convergence of the numerical solvers (in the state variable), most multiscale schemes gain efficiency by only approximating a reduced set of slowly varying coarse/slow/macroscopic variables [10, 11, 23, 30, 43, 46, 58].

In this paper, we develop a general strategy that couples multiscale integrators and fully resolved fine scale integration for parallel in time computation of HiOsc solutions of a class of ODEs. There are several advantages in such coupling strategies. First, some multiscale methods (such as the Poincaré-map technique [2]) only approximate the slow constituents or slow variables of the dynamics. Proper coupling of

*Submitted to the journal's Methods and Algorithms for Scientific Computing section March 4, 2015; accepted for publication (in revised form) August 31, 2016; published electronically November 8, 2016.

http://www.siam.org/journals/sisc/38-6/M101104.html

Funding: The work of the third author was supported by a Simons Foundation Fellowship, NSF grants DMS-1217203, and DMS-1318975.

†Department of Mathematics, Bar-Ilan University, Ramat-Gan 5290002, Israel (arielg@math.biu.ac.il).

‡Department of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332 (skim396@math.gatech.edu).

§Department of Mathematics and Institute for Computational Engineering and Sciences (ICES), The University of Texas at Austin, Austin, TX 78712 and KTH Royal Institute of Technology, Sweden (ytsai@ices.utexas.edu).

multiscale and fine scale solvers via a parareal-like framework can be efficient (by parallelization) in computing full detailed solutions, including the fast phase in the HiOsc dynamics. A good class of multiscale methods may be used as a coarse integrator as long as fast phase-like variables are appropriately aligned with sufficient efficiency and accuracy, as described below. Convergence in the slow constituents or in all of the numerical solutions can be shown, provided that the alignment strategy satisfies the conditions prescribed in "Local alignment" or "Forward alignment of step size H" in section 3. Second, the parareal iterations enhance the stability and accuracy of the multiscale scheme, in particular, when the scale separation in the system is not significant and the corresponding sampling/averaging errors are nonnegligible. Finally, parareal multiscale coupling schemes can deal with more challenging situations, for example, (a) *the effective equation is valid almost everywhere macroscopically, but is not an adequate description of the system at small but a priori "unpredictable" locations in the phase space (as these regions may depend on the solutions)*; and (b) *the influence of microscopic solutions in these regions on the macroscopic solution elsewhere is significant.*

In [50], Legoll, Lelievre, and Samaey suggest a multiscale parareal scheme for singularly perturbed ODEs in which the fast dynamics is dissipative, i.e., the dynamics relaxes rapidly to a low-dimensional manifold. One of the main contributions of [50] is the understanding that the slow and fast parts of the dynamics need to be addressed separately. They suggest two approaches: the first is a straightforward application of parareal, which is shown to converge but loses accuracy as the system becomes more singular. Their second approach assumes that the system is split into slow and fast variables, or that a change of variables splitting the system is given. This approach may be applied to HiOsc systems, but it is relatively restrictive as in many examples and applications such a splitting is not known. Dai et al. [20] suggest an application of the parareal framework to Hamiltonian systems. They consider two main approaches. The first is a time-reversible iteration scheme (applied together with time-reversible fine and coarse integrators). The second projects solutions at coarse time segments onto the constant energy manifold. The two approaches are also combined together. The first approach is specific to Hamiltonian dynamics and not to general HiOsc problems. The projection method cannot be applied to the HiOsc case because the main difficulty is not with the approximation of slow variables (or constants of motion), but with the fast phase. In addition, since the methods presented in [20] are not multiscale, their accuracy and efficiency are expected to deteriorate when the frequencies of oscillations are large. Combining the symmetric approach of Dai et al. with our alignment method for Hamiltonian systems may be an interesting application, but is beyond the scope of the current manuscript. In particular, the alignment algorithms should also be made symmetric, similar to the ideas of Dai et al. Applications of parareal methods to Hamiltonian dynamics is also analyzed in [28]. Additional approaches to use symplectic integrators with applications to molecular dynamics include [12, 13, 40]. Finally, Haut and Wingate [33] suggest a parareal method for PDEs with linear HiOsc forcing, As in [50], their method applies exact knowledge of the fast variable (the phase in the HiOsc case) to design a convergent parareal scheme. In this respect, the method proposed here goes further and is also applicable to nonlinear HiOsc systems. However, in this paper the discussion is restricted to the ODE case. One of the main goals of the current paper is to design a convergent parareal algorithm that does not require explicit knowledge of the fast and slow variables.

We begin with a short overview of the parareal method within the context of ODEs and test its performance on a simple example HiOsc system.

1.1. The parareal method for ODEs. Consider the initial value problem

$$\dot{u} = f(t,u), \quad u(0) = u_0, \tag{1.1}$$

where $u \in \mathbb{R}^d$ and $t \in [0,T]$. We assume that f is sufficiently smooth. Let H denote an intermediate time step, $0 < H < T$, and $N = T/H$ an integer. Suppose that we are given two approximate integrators for (1.1): a cheap coarse integrator with low accuracy denoted $\mathcal{C}$, and a fine, high accuracy integrator which is relatively expensive in terms of efficiency, denoted $\mathcal{F}$. Furthermore, Denote by u_n^k the approximation for $u(nH)$ at the kth iteration. For all iterations, the initial values are the same, $u_0^k = u_0, k = 0, 1, \ldots$. The parareal method to (1.1) is as follows.

Algorithm 1.1.

1. Initialization: Construct the zeroth iteration approximation using a chosen coarse integrator:

$$u_0^0 = u_0 \text{ and } u_n^0 = \mathcal{C}_H u_{n-1}^0, \quad n = 1, \ldots, N.$$

2. Iterations: $k = 1 \ldots K$

$$u_0^k = u_0 \text{ and } u_n^k = \mathcal{C}_H u_{n-1}^k + \mathcal{F}_H u_{n-1}^{k-1} - \mathcal{C}_H u_{n-1}^{k-1}, \quad n = 1, \ldots, N. \tag{1.2}$$

Note that the calculation of the fine integrator $\mathcal{F}_H u_{n-1}^{k-1}$ in (1.2) requires only the initial condition u_{n-1}^{k-1}, which depends on the previous iteration. Hence, for each k, $\mathcal{F}_t u_{n-1}^{k-1}$, $0 < t \leq H$, $n = 1, 2, \cdots, N$, can be computed in parallel. The solution computed by the accurate but expensive integrator is a fixed point. Indeed, when the iteration is sufficiently large, the solution u_n^k become identical to it, $u_n^k = \mathcal{F}_{nH} u_0$ for all $n \leq k$. In fact, (1.2) can be regarded as a fixed-point iteration. In [53], it is proved that under some sufficient conditions of f, with a first order accruate coarse integrator,

$$|u_n^k - u(nH)| \leq C(H^k + E_f), \tag{1.3}$$

where E_f is the global error in solving the full ODE using the fine propagator, and C depends on the derivatives of the solutions.

1.2. Parareal and HiOsc ODEs. We consider HiOsc ODEs given in the singular perturbation form

$$\dot{u} = \epsilon^{-1} f_1(u) + f_0(u) \tag{1.4}$$

with initial condition $u(0) = u_0 \in D \subset \mathbb{R}^d$, where D is a domain uniformly bounded in ϵ. The parameter $0 < \epsilon \leq \epsilon_0 \ll 1$ characterizes the separation of time scales—the fast scale involves oscillations with frequencies of order ϵ^{-1} while the computational time domain is $[0,T]$ with T independent of ϵ. Throughout the paper we assume that f_1, f_0 are sufficiently smooth, and that for each $u_0 \in D$, $u(t)$ is uniformly bounded in ϵ in the time interval $[0,T]$. Furthermore, we assume that the Jacobian of f_1 has only purely imaginary eigenvalues in D, which are bounded away from 0 and independent

of ϵ. These settings typically imply that the computational complexity of direct non-multiscale methods is at least $\mathcal{O}(\epsilon^{-1})$ in order to have a reasonable accuracy in the computed solutions. Using an order p solver, the truncation error is proportional to the $p+1$ time derivative of the solution, which is of order $\mathcal{O}(\epsilon^{-(p+1)})$. As a result, the parareal error for HiOsc systems is bounded explicitly with respect to ϵ as follows:

$$|u_n^k - u(nH)| \leq C \left\{ E_f + \left[\epsilon^{-1} \left(\epsilon^{-1} H \right)^p \right]^k \right\}. \tag{1.5}$$

An immediate consequence is that H has to be $o(\epsilon)$, even when applying A-stable or symplectic methods. See, for example, the conclusion in [20].

This simple example reveals the reason why a naive implementation of the parareal approach may not be effective for integrating HiOsc problems: Both stability and accuracy restrictions require that the coarse integrator take steps of order ϵ. As a result, the number of coarse steps is $\mathcal{O}(\epsilon^{-1})$ and the method may take $\mathcal{O}(\epsilon^{-1})$ iterations to converge.

2. Fast oscillations and parareal. In order to facilitate the presentation of the main algorithms, we shall first describe the setting for the underlying multiscale methods.

The literature on efficient numerical integration of problems with separated time scale is rapidly growing. For HiOsc ODEs, recent approaches include envelope methods [55], FLow AVeraging integratORS [58], Young measure [10, 11] and equation free approaches [43], Magnus methods [17, 36], Filon methods [38, 44], spectral methods [37, 51], asymptotic expansions [19, 39], and the heterogeneous multiscale methods [1, 22, 23]. For a recent review, see [25].

Typically, multiscale methods tackle the computational difficulty in solving HiOsc ODEs by taking advantage of scale separation, and aim at computing only the slowly varying properties of the oscillatory solutions. It requires that enough information about the influence of fast scales on the slower scale dynamics can be obtained by performing localized simulations over short times, and thereby better efficiency is achieved. The numerical complexity of these methods is therefore much smaller than direct simulations of the given systems with HiOsc solutions. For example, [5] presents multiscale algorithms that compute the effective behavior of HiOsc dynamical systems by using slow variables that are predetermined either analytically or numerically. More precisely, we define a slow variable for the system (1.4) with solution $u(t;\epsilon)$.

Definition 2.1. *A smooth function $a(t,\epsilon)$ is called* slow *to order $\nu \geq 1$ if $|d^\nu a/dt^\nu| \leq C$ in $t \in [0,T]$ for some constants C and T independent of $\epsilon \in (0,\epsilon_0]$, $\epsilon_0 > 0$. A smooth function $\xi(u) : D \to \mathbb{R}$ is called a* slow variable with respect to $u(t)$ if $\xi(t) = \xi(u(t;\epsilon))$ is slow to order 1.

See [6, 10, 11, 27, 29, 47, 48, 49] for similar definitions and applications. In this paper, we will work with the following main assumption.

Assumption 2.2. There exists a diffeomorphism $\Psi : u \to (\xi(u), \phi(u))$, independent of ϵ, separating slow and fast variables such that (ξ, ϕ) along the trajectories of (1.4) satisfies an ODE of the form

$$\begin{cases} \dot{\xi} = g_0(\xi,\phi), & \xi(0) = \xi(u_0), \\ \dot{\phi} = \epsilon^{-1} g_1(\xi) + g_2(\xi,\phi), & \phi(0) = \phi(u_0), \end{cases} \tag{2.1}$$

where $\xi \in \mathbb{R}^{d-n}$, $\phi \in \mathbb{R}^n$, and $0 < \epsilon \leq \epsilon_0 \ll 1$ is a small parameter. We assume that for fixed slow coordinates ξ, the fast variable ϕ is ergodic[1] with respect to an invariant manifold, which is diffeomorphic to an n-dimensional torus, $\mathbb{T}^n$.

Using ergodicity, one can invoke a theory of averaging [56], which implies that the dynamics of slow variables can be approximated ($\mathcal{O}(\epsilon)$ in the sup norm for $0 \leq t \leq T$, $T = \mathcal{O}(1)$) by an averaged equation of the form

$$\begin{aligned} &\dot{\bar{\xi}} = F(\bar{\xi}), \quad F(\bar{\xi}) = \int g_0(\xi, \phi) d\phi_\xi, \\ &\bar{\xi}(0) = \xi(u_0), \end{aligned} \tag{2.2}$$

where $d\phi_\xi$ denotes the invariant measure for ϕ at fixed ξ. For example, perturbed integrable Hamiltonian systems constitute a wide class of dynamical systems that satisfy this assumption. From now on, we shall refer to ϕ as the phase of u.

The main objective of many multiscale methods is efficient numerical approximations of $\xi(u(t))$ only. The general strategy of our algorithm is based on such multiscale methods for HiOsc ODEs that only resolve the macroscopic behavior of a system as specified by the slow variables [3, 4, 5, 6, 7, 8, 26, 57]. In this respect, the algorithms listed above are different from other multiscale methods that resolve all scales of the dynamics, for example, multilevel methods or high order asymptotic expansions [17, 18, 19, 47, 48].

It is possible to design a parareal algorithm for computing *only the averaged slow variables* using multiscale integrators as both the coarse and fine integrators. Such an approach is essentially a parareal scheme for the averaged equation. However, this is not the point of this paper—here we are interested in the possibility of creating a parareal algorithm that computes *every variable*, including the fast phase information.

We consider the problem of using a multiscale integrator in the coarse integration, and provide the stability of the corresponding coupling of multiscale-fine integrators under the parareal framework. Since the error bound stated in (1.5) still formally applies in this case, one cannot expect convergence of $u(t)$ unless some additional improvement is made to the chosen existing multiscale scheme.

Consider the following simple example of a linear oscillator with a growing amplitude,

$$\dot{u} = (1 + i\epsilon^{-1})u, \quad u(0) = 1. \tag{2.3}$$

The trajectory of $u(t) = e^{(1+i\epsilon^{-1})t}$ is an expanding spiral in the complex plane. It is easily verified that $|u(t)| = e^t$ is a slow variable. For convenience of the discussion, we assume that the fine/microscopic solver is exact, i.e., $\mathcal{F}_t u = e^{(1+i\epsilon^{-1})t}u$, and that the coarse/macroscopic solver is *exact in the slow variables,* i.e., any function of $|u|$ is computed without error *but the phase of u may be wrong.* We write the macroscopic solution as $\mathcal{C}_t U = e^t e^{i(t\epsilon^{-1}+\theta_t)}U$, where $\theta_t \in [0, 2\pi)$ denotes the error in the phase that is produced by the macroscopic solver. Applying Algorithm 1.1 we obtain

$$u_3^{(1)} = u(3H)\left(1 + \mathcal{O}(\theta_H^2)\right), \quad u_3^{(2)} = u(3H)\left(1 + \mathcal{O}(\theta_H^3)\right). \tag{2.4}$$

This simple exercise shows that *the naive iterations improve the accuracy of the macroscopic solution if θ_H is small.* However, in a typical HiOsc method, θ_H is

[1]By ergodic, we mean that any trajectory of $\phi(t)$ can get arbitrarily close to any point in the invariant manifold. In particular, this implies the existence of a unique invariant distribution and Birkhoff's ergodic theorem.

not necessarily small. In general, θ_H can be any value in $[0, 2\pi]$ and the parareal iterations may diverge.

In the following sections, we show that by aligning the phase of the coarse and fine solvers, it is possible to design parareal algorithms using multiscale coarse integrators.

3. Multiscale parareal. In this section, we introduce the main contribution of this paper—accurate and convergent parareal algorithms that use multiscale methods as coarse integrators. Two parareal coupling schemes are presented. The first focuses on approximating only the slow variables using what we call a "local alignment" strategy. The second coupling scheme achieves sup-norm convergence in the state variable, $u \in \mathbb{R}^d$ using a "forward alignment" strategy in addition to the local alignment. Here "alignment" refers to adjustment of the numerical solutions computed by the coarse multiscale integrator, based on comparison with solutions computed in parallel by the fine integrator. We show that if local and forward alignments can be implemented satisfying the prescribed accuracy objectives, the coupling schemes are convergent in the sense mentioned above.

3.1. Multiscale coarse integrator. For the remainder of this paper, we shall assume that the coarse propagator is a multiscale method that only approximates the slow variables. In order to emphasize this point, the multiscale coarse integrator will be denoted $\mathcal{M}_t$ in place of $\mathcal{C}_t$. Similarly to the assumption in a standard parareal method, we assume that

$$|\delta\mathcal{M}_t x - \delta\mathcal{M}_t y| \leq (1 + tCH)|\xi(x) - \xi(y)|. \tag{3.1}$$

However, with a multiscale coarse integrator, (3.1) implies that stability only in the slow variables is guaranteed. Accordingly, we propose to modify the coarse multiscale integrator by further adjusting the fast variable (a fast phase in the case of HiOsc problems). In terms of slow-fast coordinates, the multiscale integrator will be stable in the slow coordinates due to (3.1) while stability in the fast variable will be enforced by aligning trajectories with respect to a common reference phase.

We first elucidate our prescription for a proper local alignment. For simplicity, u_0 will denote the value computed by the coarse multiscale integrator at some time t_n that we wish to adjust. v_0 will denote the reference value for the adjustment of u_0. In practice, at the kth iteration, v_0 will come from the fine integrator applied to the values computed in the $(k-1)$th iteration.

Local alignment:
Given u_0 and v_0 such that $\xi(u_0) = \xi(v_0) + \Delta\xi$.
Find a point $\tilde{w}_0$ such that $\left|\tilde{w}_0 - \Psi^{-1}(\xi(u_0), \phi(v_0))\right| = \mathcal{O}(\Delta\xi)$.

In other words, the local alignment procedure replaces the u_0 by a new point $\tilde{w}_0$ that has the same (to order $\Delta\xi$) slow coordinates, i.e., ξ values, as u_0, and approximately the same phase as v_0. A trivial solution to the local alignment problem is to set $\tilde{w}_0 := v_0$. However, this is not an adequate strategy in the generalization described in section 3.2 that enables convergent approximation of the state variables, including the fast phases.

Notation 3.1. We denote such a local alignment procedure as $\tilde{w}_0 = \mathcal{S}_0(u_0; v_0)$.

Given a local alignment algorithm $\mathcal{S}_0$, we propose the modified parareal scheme.

ALGORITHM 3.2.

1. Initialization: (Construct the zeroth iteration approximation):

$$u_0^0 = u_0 \quad \text{and} \quad u_n^0 = \mathcal{M}_H u_{n-1}^0, \quad n = 1, \dots, N.$$

2. Iterations: $k = 1 \dots K$
   (a) Parallel fine integrations for $n = k, \dots, N$,

$$u_{F,n}^k = \mathcal{F}_H u_{n-1}^{k-1}.$$

   (b) Parareal correction: For $n = k, \dots, N$,
   (3.2)

$$u_0^k = u_0 \quad \text{and} \quad u_n^k = \mathcal{S}_0(\mathcal{M}_H u_{n-1}^k; u_{F,n}^k) + u_{F,n}^k - \mathcal{S}_0(\mathcal{M}_H u_{n-1}^{k-1}; u_{F,n}^k).$$

In each iteration we first calculate all fine scale integrations. Then, the results of the multiscale integrators are aligned with the fine scale ones. In the following, we prove that using Algorithm 3.2, all slow variables converge to their limiting value given by the fine scale approximation. We consider a first order multiscale integrator with local phase alignment.

THEOREM 3.3. *Let $K \le N/2 = T/2H$. Then, for all $k \le K$,*

$$\sup_{n=0,\dots,N} \left| \xi(u_n^k) - \xi\left(\mathcal{F}_{nH} u_0\right) \right| \le C H^k.$$

Proof. We recall the assumption that there exists a diffeomorphism $\Psi : u \to (\xi(u), \phi(u))$ such that $\xi \circ u(t)$ are slow while $\phi \circ u(t)$ are fast. The variables (ξ, ϕ) are only used in the analysis but *not* in the numerical algorithm.

Denote $\delta\mathcal{S}_0(\mathcal{M}_t u_1; u^*) = \mathcal{S}_0(\Phi_t u_1; u^*) - \mathcal{S}_0(\mathcal{M}_t u_1; u^*)$, where Φ denotes the flow map (propagator) associated with the ODE, i.e., the exact solution at time t with initial condition u_1. We have

$$\Psi \circ \delta\mathcal{S}_0(\mathcal{M}_t u_1; u^*) - \Psi \circ \delta\mathcal{S}_0(\mathcal{M}_t u_2; u^*) = (\delta\xi, \delta\phi) \tag{3.3}$$

such that $|\delta\xi| \le (1 + Ct)|\xi(u_1) - \xi(u_2)|$ but $|\delta\phi| = \mathcal{O}(\epsilon)$ which is the accuracy of local alignment. Comparing with conventional methods as a coarse integrator and the related estimate (1.5), the slow part is *controlled by the local phase alignment* in Algorithm 3.2 just like in the nonsingular case, while the rapidly changing phase is incorrect but does not affect the accuracy of the slow variables.

The slow variables of u_n^k in (3.2) are

$$\xi(u_n^k) = \xi(\mathcal{S}_0(\mathcal{M}_H u_{n-1}^k; u_{F,n}^k)) + \xi(u_{F,n}^k) - \xi(\mathcal{S}_0(\mathcal{M}_H u_{n-1}^{k-1}; u_{F,n}^k)),$$

which is valid with the local alignment. We may thus think of the multiscale integrator combined with the local alignment as a coarse integrator with first order accuracy for the slow variables. For shorthand, we denote by $\mathcal{M}_H u_{n-1}^k$ the combined $\mathcal{S}_0(\mathcal{M}_H u_{n-1}^k; u_{F,n}^k)$. Thus,

$$\begin{aligned}
&\xi(u_n^k) - \xi(\mathcal{F}_{nH} u_0) \\
&= \left[\xi(\mathcal{M}_H u_{n-1}^k) + \xi(\mathcal{F}_H u_{n-1}^{k-1}) - \xi(\mathcal{M}_H u_{n-1}^{k-1}) - \xi(\mathcal{F}_H \mathcal{F}_{(n-1)H} u_0)\right] \\
&= \left[\xi(\mathcal{M}_H u_{n-1}^k) - \xi(\mathcal{M}_H \mathcal{F}_{(n-1)H} u_0)\right] + \left[\xi(\delta\mathcal{M}_H \mathcal{F}_{(n-1)H} u_0) - \xi(\delta\mathcal{M}_H u_{n-1}^{k-1})\right] \\
&\quad + \left[\xi(\delta\mathcal{F}_H u_{n-1}^{k-1}) - \xi(\delta\mathcal{F}_H \mathcal{F}_{(n-1)H} u_0)\right].
\end{aligned}$$

Using (3.3), for every slow variable ξ, we have that

$$\begin{aligned}&\left|\xi(u_n^k) - \xi(\mathcal{F}_{nH}u_0)\right| \\ &\leq (1+CH)\left|\xi(u_{n-1}^k) - \xi(\mathcal{F}_{(n-1)H}u_0)\right| + C(E_f+H)H\left|\xi(u_{n-1}^{k-1}) - \xi(\mathcal{F}_{(n-1)H}u_0)\right|.\end{aligned}$$

Denoting $\theta_n^k = (1+CH)^{k-n}(E_f+H)^{-k}H^{-k}\left|\xi(u_n^k) - \xi\left(\mathcal{F}_{nH}u_0\right)\right|$ and following the same procedure as in [53], we have for the slow variable

$$\sup_{n=0,\ldots,N}\left|\xi(u_n^k) - \xi\left(\mathcal{F}_{nH}u_0\right)\right| \leq C(E_f+CH)^k \leq CH^k. \tag{3.4}$$ ∎

3.2. Phase continuity in the coarse and fine scale simulations. We next consider convergence of the parareal approximation to the exact solutions. The main idea is to enforce consistency in the fine scale solutions between neighboring coarse time intervals. We may rephrase this problem as the following.

> ***Forward alignment of step size*** H***:***
> Given u_0, v_0, and $u_1 = \mathcal{F}_H u_0$ such that $\xi(u_0) - \xi(v_0) = \mathcal{O}(\epsilon)$.
> Let $w_0 = \Psi^{-1}(\xi(u_0), \phi(v_0))$ and $w_1 = \mathcal{F}_H w_0$.
> Find a point $\tilde{w}_1$ using u_1 such that $\xi(\tilde{w}_1) = \xi(w_1) + \mathcal{O}(\epsilon)$ and $\phi(\tilde{w}_1) = \phi(w_1) + \mathcal{O}(H^2)$.

In the problem of forward alignment, if w_0 is a point with the same slow variable as u_0 and phase as v_0, then a forward alignment procedure constructs an order H^2 approximation of $w_1 = \mathcal{F}_H w_0$, the right endpoint of a coarse interval. See Figure 1(A) for a schematic sketch.

Notation 3.4. We denote such a forward alignment procedure as $\tilde{w}_1 = \mathcal{S}_H^\lambda(u_1; u_0, v_0)$, where λ are precomputed parameters to be used in the alignment.

The forward alignment procedure can be trivially accomplished simply by setting $\tilde{w}_1$ to $\mathcal{F}_H v_0$ or to w_1. However, this would require the additional computation of $\mathcal{F}_H v_0$, and so this trivial "fix" has a computational cost of sequentially solving the entire system with the fine integrator. In practice, for the purpose of parallel in time computations, one needs to do so with a computational cost that is lower than running the fine scale solver sequentially. Hence, we need to estimate the solution to the given ODE with the given initial condition v_0 by certain simple operations performed on the fine scale solutions already computed in parallel. In the following section, we shall

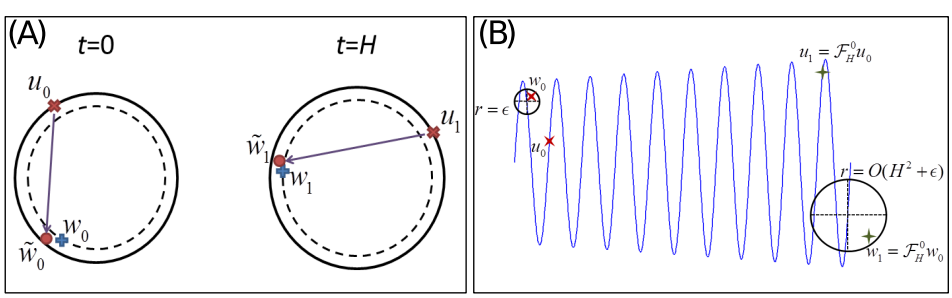

FIG. 1. *Local and forward alignments.* (A) *At* $t = 0$, *given two points* u_0 *and* v_0, *we wish to approximate* $w_0 = \mathcal{S}_0(u_0; v_0)$*—a point that has the same slow variables as* u_0 *and the same phase as* v_0. *At* $t = H$, *we approximate the point* $w_1 = \mathcal{F}_H w_0$. (B) *At* $t = 0$, *a small* $\mathcal{O}(\epsilon)$ *of* $\tilde{w}_0$ *yields at* $t = H$ *a larger* $\mathcal{O}(H^2 + \epsilon)$ *error. The center of the circles represent* $\tilde{w}_0$ *and* $\tilde{w}_1$.

describe a forward alignment algorithm for the special case of HiOsc ODEs, in which, for fixed slow variables, the fast phase is periodic. The method applies only a local exploration by means of minimal additional fine scale computation of the solution around u_0 and u_1. In particular, its efficiency is independent of ϵ.

To summarize, we present the complete multiscale-parareal algorithm in Algorithm 3.5. Recall that $\mathcal{M}_H$ is a multiscale method that only approximates the slow variables. For the fast phase, using local and forward alignments, we propagate the needed phase adjustments sequentially along with the parareal correction.

Local and forward alignment steps (step 2(c)i and ii, respectively) create a point $\tilde{u}_{F,n}^{k-1}$ at the end of each coarse segment according to which all points in the current corrector iteration can be aligned. Since the error in each forward alignment is of order H^2, we find that the overall phase is continuous up to a global $\mathcal{O}(H)$ error. We conclude that the aligned coarse multiscale method provides a globally $\mathcal{O}(H)$ approximation of both slow and fast variables, i.e., it approximates the solution in the sup norm.

Before describing numerical methods for local and forward alignments of HiOsc ODEs, we address the convergence of the algorithm.

ALGORITHM 3.5. *Full multiscale-parareal algorithm.*

1. Initialization: Construct the zeroth iteration approximation:

$$u_0^0 = u_0 \quad \text{and} \quad u_n^0 = \mathcal{M}_H u_{n-1}^0, \quad n = 1, \dots, N.$$

2. Iterations: $k = 1, \dots, K$

(a) Parallel fine integrations for $n = k, \dots, N$:

$$u_{F,n}^{k-1} := \mathcal{F}_H u_{n-1}^{k-1}.$$

(b) Header: Set the initial condition $u_0^k = u_0$ and for $n = 1, \dots, k$, set $u_n^k = u_{F,n}^{k-1}$.

(c) Parareal step: Set the initial reference point $u^* = u_k^k$ and for $n = k+1, \dots, N$,

i. Locally align the previous u_{n-1}^{k-1} with the current reference point u^*:

$$\tilde{u}_{n-1}^{k-1} = \mathcal{S}_0(u_{n-1}^{k-1}; u^*).$$

ii. Align forward to the end of the coarse segment:

$$\tilde{u}_{F,n}^{k-1} = \mathcal{S}_H^{\lambda}(u_{F,n}^{k-1}; u_{n-1}^{k-1}, u^*).$$

iii. Corrector:

$$u_n^k = \mathcal{S}_0(\mathcal{M}_H u_{n-1}^k; \tilde{u}_{F,n}^{k-1}) + \tilde{u}_{F,n}^{k-1} - \mathcal{S}_0(\mathcal{M}_H \tilde{u}_{n-1}^{k-1}; \tilde{u}_{F,n}^{k-1}).$$

iv. Update the reference point $u^* = u_n^k$ and repeat.

3.3. Convergence of Algorithm 3.5. Convergence of Algorithm 3.5 in the state variable is obtained in two steps. First, following section 3.1 and Theorem 3.3, all slow variables converge to their values obtained by the fine scale integrators,

$$\sup_{n=0,\dots,N} |\xi(u_n^k) - \xi(u(nH))| \le C(H^k + E_f),$$

where C is a constant that is independent of ϵ. In particular, if $E_f = \mathcal{O}(\epsilon)$, then, following $\mathcal{O}(\log(\epsilon))$ iterations, the error in the slow variables is of order ϵ. As a result, after a few (typically one or two) iterations, the assumptions underlying forward alignment, that the error in the slow variables is of order ϵ, holds (more precisely, $\xi(u_0) - \xi(v_0) = \mathcal{O}(\epsilon)$). We may thus think of the adopted multiscale combined with the local/forward alignment algorithm as a coarse integrator with first order accuracy for all state variables. Hence, the conventional parareal proof of convergence as in [53] holds. More precisely, suppose that, following the phase alignment, the set $\{w_0, \ldots, w_N\}$ is computed and updated in every iteration. The following estimates hold for $j = 0, \ldots, N$,

$$\begin{aligned} |\xi(w_j) - \xi(u(jH))| &= \mathcal{O}(E_m), \\ |\phi(w_j) - \phi(u(jH))| &= \mathcal{O}(H), \\ |w_j - u(jH)| &= \mathcal{O}(H) + \mathcal{O}(E_m), \end{aligned}$$

where E_m is the error of the aligned multiscale method in approximating the slow variables.

Assume further the stability properties for the fine and aligned-multiscale coarse propagators as in [53]. Then after one parareal iteration,

$$\sup_{n=0,\ldots,N} \left|u_n^1 - u(nH)\right| \leq C(\epsilon^{-1}E_m + E_f).$$

Thus if a first order accurate multiscale integrator is used, i.e., $E_m = \mathcal{O}(H)$, we have

$$\sup_{n=0,\ldots,N} \left|\xi(u_n^k) - \xi(u(nH))\right| \leq C(H^k + E_f),$$

$$\sup_{n=0,\ldots,N} \left|u_n^k - u(nH)\right| \leq C(\epsilon^{-1}H^k + E_f).$$

The accuracy of slow variables is improved by a factor of H per parallel iteration (compare with the diverging factor of $(\epsilon^{-1}H)^k$ in (1.5)).

Remark 3.6. In [50], Legoll, Lelievre, and Samaey propose a multiscale parareal algorithm for stiff ODEs in which the fast dynamics is dissipative, i.e., trajectories quickly converge to lower-dimensional manifolds. Unlike the HiOsc case, a naive application of the parareal methodology to stiff dissipative systems converges. However, it is not very efficient and suffers from similar difficulties as discussed earlier. To circumvent these difficulties, Legoll, Lelievre, and Samaey [50] suggest a correction step that allows a consistent approximation of the fast-slow dynamics with parareal. This work assumes that the system is split into slow and fast variables or, alternatively, that a change of variables that splits the system into such coordinates is given explicitly. Essentially, the idea of [50] is to set the fast component of the multiscale solver with that obtained from the fine one. This step may be viewed as a simple alignment method. Indeed, if the (ξ, ϕ) coordinates are know, then the same approach can also be applied to the HiOsc case. In contrast, the method presented in the following section is seamless in the sense that it does not require knowing the slow or the fast variables.

4. Phase alignment strategies. In this section, we describe a numerical method for both local and forward alignments as defined in the previous section for the special

case of HiOsc ODEs in which, for fixed values of the slow variables ξ, the dynamics of the fast phase ϕ is periodic.

In the Algorithm 3.5, v_0 and u_0 will correspond to u_n^k and u_n^{k-1}, the solutions computed at the current and the previous iterations, respectively. The assumption is that v_0 is the more accurate approximation of the solution at the time $t = nH$, particularly in the phase variable. The goal is that from the available information, u_0, v_0, and $u_1 := \mathcal{F}_H u_0$, we estimate $v_1 := \mathcal{F}_H v_0$ at $t = (n+1)H$ in order to make corrections in the phase of u_1. We also emphasize that in the subsequent time steps, $\mathcal{F}_H u_0$ is always available because of the prior parallel fine integrations. Now w_0, as defined in section 3.2, is a point on the same slow coordinates as u_0 but has the same phase as v_0. Consequently $w_1 := \mathcal{F}_H w_0$ is a good estimate of v_1. In this section, we propose a strategy that move u_0 to w_0, and u_1 to a state that is within $\mathcal{O}(\epsilon)$ to w_1, *without computing* $\mathcal{F}_H w_0$ *or* $\mathcal{F}_H v_0$. Our goal is to describe a method that finds a point $\tilde{w}_0$ such that $|w_0 - \tilde{w}_0| = \mathcal{O}(\epsilon)$ (local alignment) and a second point $\tilde{w}_1$ such that $|w_1 - \tilde{w}_1| = \mathcal{O}(H^2 + \epsilon)$ (forward alignment).

In addition to Assumption 2.2, we assume the following.

Assumption 4.1. The fast variable $\phi \in \mathbb{R}$ and $g_0(\xi, \phi)$ is 1-periodic in ϕ.

For fixed ξ, the time derivative of ϕ may depend on the slow variables, i.e., the periodicity in time of $g_0(\xi, \phi(t))$ is of order ϵ and depends on ξ. Accordingly, it is denoted $\epsilon\tau(\xi)$, where τ is a smooth, slow function. Note that this does not mean that the oscillation in the original state variables are linear because the transformation Ψ is, in general, nonlinear.

4.1. At time t move u_0 closer to w_0. Assuming that $\xi(u_0) - \xi(v_0) = \mathcal{O}(\epsilon)$, we may use v_0 instead of w_0, which is not known. Denote

$$J(t; u_0, v_0) = |\mathcal{F}_t u_0 - v_0|^2 .$$

We look for the local minima of $J(t; u_0, v_0)$ closest to $t = 0$ (by the periodicity assumption, such local minima exist):

$$\begin{aligned} 0 = J'(t; u_0, v_0)(t) &= 2\left(\mathcal{F}_t u_0 - v_0\right) \cdot \frac{d}{dt}\left(\mathcal{F}_t u_0 - v_0\right) \\ &= 2\left(\mathcal{F}_t u_0 - v_0\right) \cdot \left(\frac{\partial \Psi^{-1}}{\partial \xi}\dot{\xi} + \frac{\partial \Psi^{-1}}{\partial \phi}\dot{\phi}\right). \end{aligned}$$

To leading order in ϵ, we have

$$(\mathcal{F}_t u_0 - v_0) \cdot (\partial \Psi^{-1}/\partial \phi) = \mathcal{O}(\epsilon). \tag{4.1}$$

In other words, the phase of $\mathcal{F}_t u_0$ is close to that of v_0, $\phi(\mathcal{F}_t u_0) = \psi_0 + \mathcal{O}(\epsilon)$ and therefore also to the phase of w_0. We denote the "first" two local minima

$$-\epsilon\tau_0 + \mathcal{O}(\epsilon^2) < t_0^- < 0 < t_0^+ < \epsilon\tau_0 + \mathcal{O}(\epsilon^2),$$

where $\tau_0 = \tau(\xi_0)$. Consider $\tilde{w}_0^{\pm} = \mathcal{F}_{t_0^{\pm}} u_0$, and $\tilde{w}_0 = \mathcal{S} u_0 = \lambda_+ \tilde{w}_0^+ + \lambda_- \tilde{w}_0^-$ with weights $\lambda = (\lambda_+, \lambda_-)$ independent of ϵ. Equation (4.1) implies that, for any linear combination such that $\lambda_+ + \lambda_- = 1$, $|\tilde{w}_0 - v_0| = \mathcal{O}(\epsilon)$, i.e., $\tilde{w}_0$ defined above is a valid choice in the local alignment procedure. Therefore, we define the following.

Local alignment:

$$\mathcal{S}_0(u_0; v_0) = \lambda_+ \mathcal{F}_{t_0^+} u_0 + \lambda_- \mathcal{F}_{t_0^-} u_0.$$

In the numerical implementation of the local alignment $\mathcal{S}_0$, we use a simple algorithm which solves the l_2 minimization of $J(t)$ with a small $\mathcal{O}(\epsilon)$ step size and adaptively increasing the search domain until an appropriate minimum is found. Denoting the time where the minimum is attained by t^*, we improve the accuracy by a quadratic interpolation through the neighboring points of t^*. This algorithm achieves $\mathcal{O}(\epsilon)$ accuracy with an efficiency that is independent of ϵ. See [9] for details and [54] for further references on other efficient minimization techniques.

4.2. At time $t + H$ move u_1 closer to w_1. We would like to do the same at $t = H$, i.e., move $u_1 = \mathcal{F}_H u_0$ to $w_1 = \mathcal{F}_H w_0$. The main difficulty is that we cannot expect that the solution has oscillations of constant periodicities. We denote $\tau_1 = \tau(\xi_1) = \tau_0 + (\xi_1 - \xi_0)(\partial\tau/\partial\xi) + \mathcal{O}(H^2)$. In analogy to the procedure at $t = 0$, we find the first two minimizers of

$$J(t; u_1, w_1) = |\mathcal{F}_t u_1 - w_1|^2 ,$$

such that $-\epsilon\tau_1 + \mathcal{O}(\epsilon^2) < t_1^- < 0 < t_1^+ < \epsilon\tau_1 + \mathcal{O}(\epsilon^2)$. Let

$$\tilde{w}_1 = \lambda_+ \tilde{w}_1^+ + \lambda_- \tilde{w}_1^- \text{ with } \tilde{w}_1^\pm = \mathcal{F}_{t_1^\pm} u_1 = w_1 + \mathcal{O}(\epsilon).$$

Then, for any constants λ_+, λ_- we have that $|w_1 - \tilde{w}_1| = \mathcal{O}(\epsilon)$. *The problem is that we do not know w_1 and therefore cannot find $t_1^\pm$.* One option is to use $t_0^\pm$ instead and choose weights $\lambda_\pm$ that minimize the error. This requires us to relate $t_0^\pm$ and $t_1^\pm$.

Denote

$$\begin{aligned} \Psi u_0 &= (\xi_0, \phi_0) , & \Psi u_1 &= \Psi \mathcal{F}_H u_0 = (\xi_1, \phi_1) , \\ \Psi v_0 &= (\eta_0, \psi_0) , & & \\ \Psi w_0 &= (\xi_0, \psi_0) , & \Psi w_1 &= \Psi \mathcal{F}_H w_0 = (\xi_1, \psi_1) . \end{aligned}$$

Without loss of generality, we assume that $\psi_0 > \phi_0$ and $|\psi_0 - \phi_0| < 1$. Similarly, assume $\psi_1 > \phi_1$ and $|\psi_1 - \phi_1| < 1$. Then, to leading order in ϵ,

$$\begin{aligned} t_0^+ &= (\psi_0 - \phi_0)\epsilon\tau_0, & t_1^+ &= (\psi_1 - \phi_1)\epsilon\tau_1, \\ t_0^- &= -(1 - \psi_0 + \phi_0)\epsilon\tau_0, & t_1^- &= -(1 - \psi_1 + \phi_1)\epsilon\tau_1. \end{aligned}$$

Next, denote the solution of (ξ, ϕ) with initial condition (ξ_0, ϕ_0) as $\xi(t; \xi_0, \phi_0)$ and $\phi(t; \xi_0, \phi_0)$, i.e., $(\xi(t; \xi_0, \phi_0), \phi(t; \xi_0, \phi_0)) = \Psi \mathcal{F}_t u_0$. Using the averaging principle (2.2), we can write

$$\xi(t; \xi_0, \phi_0) = \bar{\xi}(t) + \epsilon\gamma(t/\epsilon, \xi) + \mathcal{O}(\epsilon^2),$$

where $\bar{\xi}(t)$ is a slow function that does not depend on the phase and $\gamma(s, \xi)$ is independent of ϵ and is $\tau(\xi)$-periodic in s with zero average, $\int_0^{\tau(\xi)} \gamma(s, \xi) ds = 0$.

$$\begin{aligned} \phi(H; \xi_0, \phi_0) &= \phi_0 + \int_0^H \left[\epsilon^{-1} g_1(\xi(t; \xi_0, \phi_0)) + g_2(\xi(t; \xi_0, \phi_0))\right] dt \\ &= \phi_0 + \int_0^H \left[\epsilon^{-1} g_1(\bar{\xi}(t)) + g_2(\bar{\xi}(t))\right] dt + \int_0^H g_1'(\bar{\xi}(t))\gamma(t/\epsilon, \bar{\xi}(t)) dt + \mathcal{O}(\epsilon) \\ &= \phi_0 + F(\xi_0; \epsilon) + \mathcal{O}(\epsilon) \end{aligned}$$

for some function F that depends only on ξ_0 and ϵ, but not on the initial phase ϕ_0. In particular, we note that

$$\phi(w_1) - \psi_0 = \phi(\mathcal{F}_H w_0) - \psi_0 = \phi(H; \xi_0, \psi_0) - \psi_0 = F(\xi_0; \epsilon) + \mathcal{O}(\epsilon).$$

Similarly,

$$\phi_1 - \phi_0 = \phi(\mathcal{F}_H u_0) - \phi_0 = \phi(H; \xi_0, \phi_0) - \phi_0 = F(\xi_0; \epsilon) + \mathcal{O}(\epsilon).$$

Hence, $\phi(w_1) - \psi_0 = \phi_1 - \phi_0 = \mathcal{O}(\epsilon)$. In other words, starting at w_0 instead of u_0 introduces a phase shift that is practically constant. We have then

$$\begin{aligned} t_1^+ &= (\psi_1 - \phi_1)\epsilon\tau_1 = (\psi_0 - \phi_0)\,\epsilon\left[\tau_0 + \frac{\partial \tau}{\partial \xi}\cdot(\xi_1 - \xi_0)\right] + \mathcal{O}(\epsilon H^2 + \epsilon^2) \\ &= t_0^+ + H t_0^+ \Delta + \mathcal{O}(\epsilon H^2 + \epsilon^2), \end{aligned}$$

where $\Delta = \frac{1}{H\tau_0}\frac{\partial \tau}{\partial \xi}\cdot(\xi_1 - \xi_0) = \mathcal{O}(1)$. Similarly,

$$t_1^- = t_0^- + H t_0^- \Delta + \mathcal{O}(\epsilon H^2 + \epsilon^2).$$

Consider

$$\mathcal{F}_{t_0^\pm} u_1 = \mathcal{F}_{-H t_0^\pm \Delta}\mathcal{F}_{t_1^\pm} u_1 + \mathcal{O}(\epsilon H^2 + \epsilon^2) = \mathcal{F}_{-H t_0^\pm \Delta}\tilde{w}_1^\pm + \mathcal{O}(\epsilon H^2 + \epsilon^2).$$

Expanding around $\tilde{w}_1^\pm$, $\mathcal{F}_{t_0^\pm} u_1 = \tilde{w}_1^\pm - \frac{H}{\epsilon}\delta t_0^\pm \Delta + \mathcal{O}(H^2 + \epsilon^2)$ for some vector $\delta \in \mathbb{R}^d$ independent of ϵ. Therefore,

$$\begin{aligned} \mathcal{S}_H^\lambda(u_1; u_0, v_0) &= \lambda_+ \mathcal{F}_{t_0^+} u_1 + \lambda_- \mathcal{F}_{t_0^-} u_1 \\ &= \left(\lambda_+ \tilde{w}_1^+ + \lambda_- \tilde{w}_1^-\right) + \frac{H}{\epsilon}\delta\left(\lambda_+ t_0^+ + \lambda_- t_0^-\right)\Delta + \mathcal{O}(H^2 + \epsilon^2) \\ &= w_1 + \frac{H}{\epsilon}\delta\left(\lambda_+ t_0^+ + \lambda_- t_0^-\right)\Delta + \mathcal{O}(H^2 + \epsilon). \end{aligned} \tag{4.2}$$

Finally we see that with the choice

$$\lambda_+ = \frac{-t_0^-}{t_0^+ - t_0^-} \quad , \quad \lambda_- = \frac{t_0^+}{t_0^+ - t_0^-},$$

the first order terms cancel out. Thus, we obtain a second order accurate forward alignment to w_1. See Figure 2 for the error of (4.2) for the simple example in (2.3). In the following, we shall denote $\lambda = (\lambda_+, \lambda_-)$.

Algorithm 3.5 applies the convex combination (4.2) in forward alignment. Indeed, convergence in the state variable heavily relies on this step because the new point $\tilde{u}_{F,n}^{k-1}$ after the forward alignment is assigned as the reference for local alignment in the next coarse interval. See step 2(c)iii. We emphasize that taking $\mathcal{S}_H^\lambda$ as (4.2) may shift the slow coordinates of the resulting u_1 from what was computed by the multiscale coarse integrator and assumed to be accurate. In the next subsection, we propose a more elaborate method to further improve the accuracy of the forward alignment step.

4.3. Improving accuracy in forward alignment. Here, the idea is that we identify the convex combination with the point which divides the trajectory of (1.4) originating from $\mathcal{F}_{t_0^+} u_1$ and ending close to $\mathcal{F}_{t_0^-} u_1$ by a proportion of λ_- to λ_+. Since there are two orientations of $\mathcal{F}_t(\mathcal{F}_{t_0^+} u_1)$ defined by forward and backward in

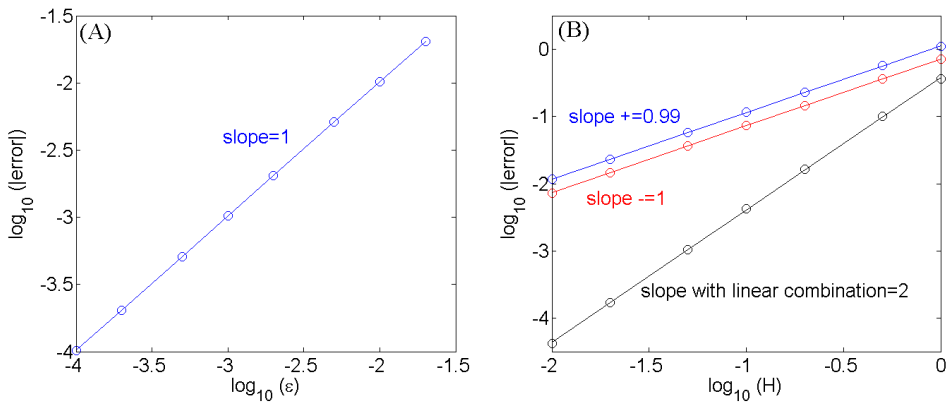

FIG. 2. *The error in correcting the phase at the end of one coarse segment for the expanding spiral* (2.3). (A) *In local alignment, the phase at the end of a coarse segment* u_0 *is aligned with* v_0 *with an* $\mathcal{O}(\epsilon)$ *error.* (B) *Blue: forward correction with* s_+*; red: backward correction with* s_-*; and black: a linear combination of shifts using the forward alignment algorithm defined in section* 4. *With the proposed linear combination, the error is of order* H^2.

time integrations, the modified convex combination will provide us with two points depending on the orientations, and we will choose the one closer to $\mathcal{S}_H^\lambda u_1$.

First, we propose to find the first two local minimizers of

$$J(t; \mathcal{F}_{t_0^+} u_1, \mathcal{F}_{t_0^-} u_1) = |\mathcal{F}_t(\mathcal{F}_{t_0^+} u_1) - \mathcal{F}_{t_0^-} u_1|^2,$$

such that $-\epsilon\tau_1 < \Gamma_p^- < 0 < \Gamma_p^+ < \epsilon\tau_1$. Denoting

$$t_0^{++} = t_0^+ + \lambda_- \Gamma_p^+ \quad \text{and} \quad t_0^{+-} = t_0^+ + \lambda_- \Gamma_p^-, \tag{4.3}$$

we again find the first local minimizers of

$$\begin{aligned} J(t; \mathcal{F}_{t_0^-} u_1, \mathcal{F}_{t_0^{++}} u_1) &= |\mathcal{F}_t(\mathcal{F}_{t_0^-} u_1) - \mathcal{F}_{t_0^{++}} u_1|^2, \\ J(t; \mathcal{F}_{t_0^-} u_1, \mathcal{F}_{t_0^{+-}} u_1) &= |\mathcal{F}_t(\mathcal{F}_{t_0^-} u_1) - \mathcal{F}_{t_0^{+-}} u_1|^2, \end{aligned} \tag{4.4}$$

such that $-\epsilon\tau_1 < \Gamma_*^- < 0$ and $0 < \Gamma_*^+ < \epsilon\tau_1$, and denote them by

$$t_0^{--} = t_0^- + \Gamma_*^-, \; t_0^{-+} = t_0^- + \Gamma_*^+. \tag{4.5}$$

With local minimizers of (4.4), the phases between $\mathcal{F}_{t_0^{++}}$ and $\mathcal{F}_{t_0^{--}}$, and between $\mathcal{F}_{t_0^{+-}}$ and $\mathcal{F}_{t_0^{-+}}$ are the same. Now, define the new weights using t_0's in (4.3) and (4.5):

$$\lambda_{++} = \frac{-t_0^{--}}{t_0^{++} - t_0^{--}}, \; \lambda_{+-} = \frac{-t_0^{-+}}{t_0^{+-} - t_0^{-+}}, \; \lambda_{--} = \frac{t_0^{++}}{t_0^{++} - t_0^{--}}, \; \lambda_{-+} = \frac{t_0^{+-}}{t_0^{+-} - t_0^{-+}}.$$

The convex combination (4.2) is now modified as

$$\begin{aligned} \mathcal{S}_H^{\lambda_1}(u_1; u_0, v_0) &= \lambda_{++} \mathcal{F}_{t_0^{++}} u_1 + \lambda_{--} \mathcal{F}_{t_0^{--}} u_1, \\ \mathcal{S}_H^{\lambda_2}(u_1; u_0, v_0) &= \lambda_{+-} \mathcal{F}_{t_0^{+-}} u_1 + \lambda_{-+} \mathcal{F}_{t_0^{-+}} u_1. \end{aligned}$$

Here, we note that

$$\xi(\mathcal{S}_H^{\lambda_1}(u_1; u_0, v_0)) = \xi(u_1) + \mathcal{O}(\epsilon), \;\; \xi(\mathcal{S}_H^{\lambda_2}(u_1; u_0, v_0)) = \xi(u_1) + \mathcal{O}(\epsilon).$$

In words, the modified convex combinations $\mathcal{S}_H^{\lambda_1}(u_1; u_0, v_0)$ and $\mathcal{S}_H^{\lambda_2}(u_1; u_0, v_0)$ guarantee the accuracy of order ϵ in the slow variables of u_1.

Now, we propose to implement the forward alignment $\mathcal{S}_H^{\lambda}(u_1; u_0, v_0)$ as follows.

Forward alignment:

1. Set the reference point using (4.2), $\hat{u} := \lambda_+ \mathcal{F}_{t_0^+} u_1 + \lambda_- \mathcal{F}_{t_0^-} u_1$.
2. Compute two modified convex combinations with opposite orientations,

$$\lambda_{++} \mathcal{F}_{t_0^{++}} u_1 + \lambda_{--} \mathcal{F}_{t_0^{--}} u_1, \ \lambda_{+-} \mathcal{F}_{t_0^{+-}} u_1 + \lambda_{-+} \mathcal{F}_{t_0^{-+}} u_1.$$

3. Denote by $\mathcal{S}_H^{\lambda}(u_1; u_0, v_0)$ the combination closer to $\hat{u}$.

Remark 4.2. Let $\mathcal{F}^0$ be a fine integrator for the unperturbed system of (1.4). Then by definition, $\xi(\mathcal{F}_t^0 u_0) = \xi(u_0)$ for all $t > 0$ but $\phi(\mathcal{F}_t^0 u_0) \neq \phi(u_0)$ for some $t > 0$. If $\mathcal{F}^0$ is given explicitly, one can achieve more accurate local and forward alignments by using it in the minimization of $J(t)$.

5. Numerical examples. In this section, the accuracy of the suggested multiscale-parareal methods are studied in several typical examples. To this end, we briefly discuss our choice of multiscale solver used as a coarse integrator.

The simulations performed in this paper are computed using a symmetric algorithm based on the multiscale algorithm introduced in [2]. This method only assumes the existence of slow variables but does not use its explicit form. We now briefly elucidate some of the needed information about the algorithm. The multiscale propagator $\mathcal{M}_H$ of step size H will involve *fine scale propagators* $\mathcal{F}$ *for*

$$\dot{\bar{u}} = \epsilon^{-1} f_1(\bar{u}) + K(t/\eta) f_0(\bar{u}, t, \epsilon^{-1} t), \tag{5.1}$$

where the filter K *satisfies adequate smoothness and moment conditions, as well as* $\mathcal{F}^0$ *for the "unperturbed" equation:*

$$\dot{\bar{u}} = \epsilon^{-1} f_1(\bar{u}). \tag{5.2}$$

Unless specified otherwise, the following symmetric coarse propagator, denoted $\mathcal{M}_H$, originally proposed and analyzed in [45], is applied throughout the section,

$$\mathcal{M}_H = \left(1 - \frac{H}{2\eta}\right) \mathcal{F}_\eta^0 + \frac{H}{2\eta} \mathcal{F}_{-\eta}^0 \mathcal{F}_{2\eta}. \tag{5.3}$$

5.1. Stellar orbits in a galaxy. The following system is taken from the theory of stellar orbits in a galaxy [41, 42]:

$$\begin{cases} \ddot{r}_1 + a^2 r_1 = \epsilon r_2^2, \\ \ddot{r}_2 + b^2 r_2 = 2\epsilon r_1 r_2. \end{cases}$$

Following a change of variables $\mathbf{x} = [x_1, v_1, x_2, v_2]^T = [r_1, \frac{d}{d\tau} r_1/a, r_2, \frac{d}{d\tau} r_2/b]^T$ and after a rescaling of time, $t = \epsilon\tau$, the system can be written in the following form

$$\begin{aligned} \dot{x}_1 &= \epsilon^{-1} a v_1, \qquad \dot{v}_1 = -\epsilon^{-1} a x_1 + x_2^2/a, \\ \dot{x}_2 &= \epsilon^{-1} v_2, \qquad \dot{v}_2 = -\epsilon^{-1} x_2 + 2 x_1 x_2/b. \end{aligned} \tag{5.4}$$

Initial conditions are $(x_1, v_1, x_2, v_2)(0) = (1, 0, 1, 0)$. Resonance of oscillatory modes take effect in the lower order term when $a = 2$ and $b = 1$. Using the algorithm proposed in [5], three independent slow variables are identified as

$$\xi_1 = x_1^2 + v_1^2, \quad \xi_2 = x_2^2 + v_2^2, \quad \xi_3 = x_1 x_2^2 + 2v_1 x_2 v_2 - x_1 v_2^2. \tag{5.5}$$

The example falls under the category of HiOsc systems in which two stiff harmonic oscillators are coupled. The local $\mathcal{O}(H^2)$ error introduced by the first order coarse multiscale integrator is realized.

The system (5.4) is integrated using Algorithm 3.5, applying the global alignment algorithm described in section 4 to ensure convergence in the state variable. The symmetric Poincaré method is used as a coarse solver, using the trajectory of (5.4) as the flow $\mathcal{F}$ and the one of (5.4) without a lower order perturbation as the flow $\mathcal{F}^0$. We stress that the numerical approximation is obtained without using our knowledge that the system can be decomposed into the three slow variables ξ_1, ξ_2, and ξ_3 and a fast phase-like variable ϕ. These decompositions are only used in order to explain the fast-slow structure in the dynamics.

Figure 3 shows the absolute error in the state variable of the entire trajectory. Initially, the absolute error is large because the inaccurate slow variables create a jump in the phase between coarse time segments. However, after four iterations the error in the state variable drops below ϵ, which is the theoretical limit possible with multiscale methods on their own. Parameters are detailed in Table 1. The fine integrator is the ODE45 method, and the coarse integrator is the Poincaré second order multiscale method. A C^3 kernel with $p = 1$ is used for the filtered equation.

Finally, we make a brief remark on systems with multiple fast frequencies, for example, by taking $a = \sqrt{2}$ and $b = 1$ in (5.4). In this case, the dynamics is described by the two slow variables ξ_1 and ξ_2 in (5.5) and two fast variables related to the phases of the two oscillators [5]. Unlike the resonant case $a = 2$ with fixed ξ_1 and ξ_2, the fast dynamics is ergodic with respect to a two-dimensional torus. As a result, Assumption 4.1 is not satisfied and the alignment algorithms of section 4 cannot

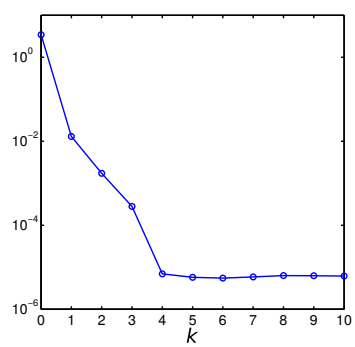

FIG. 3. *Stellar orbits in a galaxy, example* 5.1. *The absolute error in the state variables* (x_1, v_1, x_2, v_2) *as a function of iteration is depicted.*

TABLE 1
Parareal parameters in example 5.1.

ϵ	T	H	h_{fine}	$\eta_{Poincaré}$	$h_{Poincaré}$	RelTol, AbsTol (ODE45)
10^{-3}	14	0.5	$\epsilon/100$	20ϵ	$\epsilon/10$	10^{-13}, 10^{-11}

TABLE 2
The absolute error in the slow variables for a nonresonant slow-fast system described in example 5.1.

k	1	2	4	8	9
	3.28e-2	7.59e-3	1.04e-3	7.82e-4	5.25e-4

be applied. Suppose, however, that the system is given in the slow-fast coordinates explicitly. As discussed in section 3, one trivial strategy for local alignment is to simply reset the values of fast variables obtained by the coarse integrator with the values computed by the fine one at the corresponding time step. With this simple alignment strategy, slow variables do not change (i.e., $\Delta\xi = 0$) and the assumptions underlying Theorem 3.3 hold. In particular, the approximation for the slow variables will converge. Table 2 lists numerical results for the nonresonant $a = \sqrt{2}$ case. Since the slow variables are known, the midpoint rule ((5.6) and (5.7) in [5]) was chosen instead of the Poincaré method as the macrosolver. Note that although convergence of this macrosolver for the nonresonant case was not proven in [5], their results can be extended for quasi-periodic trajectories using [21]. All numerical parameters are the same as in Table 1.

5.2. Nonlinear oscillators. Consider the following example of a Voltera–Lotka oscillator with slowly varying frequency and amplitude,

$$\begin{aligned}\dot{x} &= \epsilon^{-1}x(1 - zy),\\ \dot{y} &= \epsilon^{-1}zy(x - 1),\\ \dot{z} &= 0.2x.\end{aligned}$$

Initial conditions are $(x, y, z)(0) = (1, 2.9, 1)$. For fixed z, (x, y) is a Voltera–Lotka oscillator whose period is of order ϵ. The period and amplitude of (x, y) depend on a parameter z, which is given by the time integral of x. As a result, z is a slow variable. It is easily verified that the first integral of the oscillator is also slow,

$$I = x - \log(x) + y - \log(y)/z.$$

Again, we stress that the slow variables are only used in order to demonstrate the results of the method. They are *not* used in the numerical approximation. In addition, Figure 4(A) shows the slow variable level set, $\{u \in \mathbb{R}^3 : I(u) = I(x(t_n), y(t_n), z(t_n))\}$, projected onto x-y plane. In contrast to the previous examples, the level set of the slow variable I is not a circle. As a result, $J(t)$ may have several local minima and we need to find the first local minima which is close to the global minimum of J within a few periods. Parameters are given in Table 3. The fine integrator is the ODE45 method, and the coarse integrator is the Poincaré second order multiscale method. A C^3 kernel with $p = 1$ is used for the filtered equation.

5.3. Resonant swinging spring. The swinging spring (or elastic pendulum) is a three-dimensional mechanical system in which a mass is suspended from a fixed point by a spring. It has been shown [35] that under specific resonance conditions, the pendulum exhibits stepwise precession as energy is transfered between the oscillatory modes. The equations of motion of the spring are given by

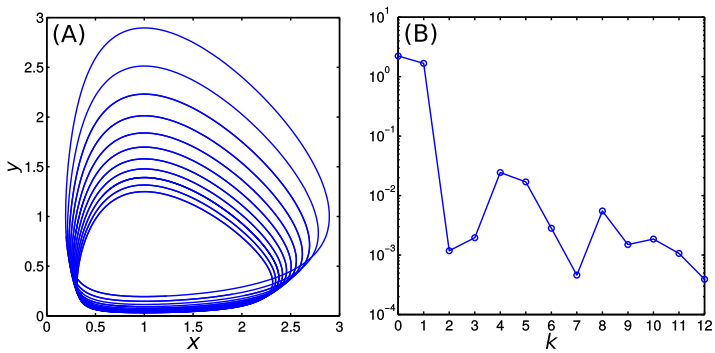

FIG. 4. *A Lotka–Voltera oscillator with slowly varying frequency and amplitude, example* 5.2. (A) *The level sets of the slow variable I projected onto the x-y plane.* (B) *The absolute error in the state variables as a function of iteration.*

TABLE 3
Parareal parameters in example 5.2.

ϵ	T	H	h_{fine}	$\eta_{Poincaré}$	$h_{Poincaré}$	RelTol, AbsTol (ODE45)
10^{-3}	10	1/2	$\epsilon/200$	30ϵ	$\epsilon/10$	10^{-13}, 10^{-10}

$$\begin{aligned}\ddot{x} + \omega_R^2 x &= \lambda x z,\\ \ddot{y} + \omega_R^2 y &= \lambda y z,\\ \ddot{z} + \omega_Z^2 z &= 0.5\lambda(x^2 + y^2),\end{aligned}$$

where ω_R and ω_Z are the frequencies of pendular and elastic oscillations, and λ depends on w_R and the length of the spring. When $\omega_Z = 2\omega_R$, the spring rotates twice in the x–y projection before returning to the initial position, and the motion becomes periodic. This is the 2:1:1 resonance dynamics which is investigated in [35].

After the change of variables $(x_1, x_2, x_3, x_4, x_5, x_6) = (\omega_R x, \dot{x}, \omega_R y, \dot{y}, \omega_Z z, \dot{z})$, the system can be written in the following form,

$$\begin{aligned}\dot{x}_1 &= \omega_R x_2,\\ \dot{x}_2 &= -\omega_R x_1 + \lambda \omega_R^{-1}\omega_Z^{-1} x_1 x_5,\\ \dot{x}_3 &= \omega_R x_4,\\ \dot{x}_4 &= -\omega_R x_3 + \lambda \omega_R^{-1}\omega_Z^{-1} x_3 x_5,\\ \dot{x}_5 &= \omega_Z x_6,\\ \dot{x}_6 &= -\omega_Z x_5 + 0.5\lambda\omega_R^{-2}(x_1^2 + x_3^2).\end{aligned} \tag{5.6}$$

Assuming an HiOsc scenario, we take $\omega_R = \pi\epsilon^{-1}$, $\omega_Z = 2\pi\epsilon^{-1}$, and $\lambda = 0.75\omega_Z^2$. Following [5], we identify five independent slow variables: three oscillatory energies, a 1:1 resonance between (x_1, x_2) and (x_3, x_4), and a 2:1 resonance between (x_1, x_2) and (x_5, x_6). Initial conditions are set as $(x_1, x_2, x_3, x_4, x_5, x_6)(0) = (0.6, 0, 0, 0.489, 1.2, 0)$. Figure 5 presents numerical results using the multiscale parareal Algorithm 3.5. Parameters are tabulated in Table 4. The fine integrator is the ODE45 method, and the coarse integrator is the Poincaré second order multiscale method. A C^1 kernel with $p = 1$ is used.

5.4. Passage through resonance. One of the fundamental assumptions underlying multiscale approaches such as Poincaré and other methods is a spectral gap in

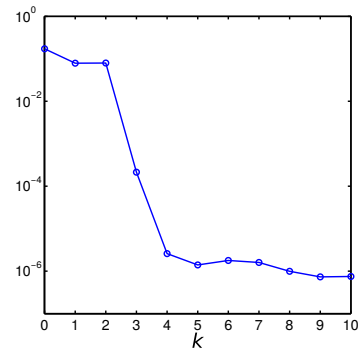

FIG. 5. *The resonant swinging spring, example* 5.3. *The absolute error in the state variables as a function of iteration.*

TABLE 4
Parareal parameters in example 5.3.

ϵ	T	H	h_{fine}	$\eta_{Poincaré}$	$h_{Poincaré}$	RelTol, AbsTol (ODE45)
10^{-3}	20	1/2	$\epsilon/50$	10ϵ	$\epsilon/10$	10^{-13}, 10^{-11}

the spectrum of the Jacobian of the equations of motion. When this assumption fails, for example, due to a temporary passage through resonance, the assumption (3.1) may not hold; the resonance and typical multiscale methods fail. However, the applicability of the multiscale algorithms can be extended by the parareal approach described above, i.e., by resolving all scales of the dynamics.

We consider the following example:

$$\begin{aligned}\dot{x} &= -2\pi\epsilon^{-1}f(z)y + 0.5\sin(z)x,\\ \dot{y} &= 2\pi\epsilon^{-1}f(z)x,\\ \dot{z} &= 1,\end{aligned}$$

where $f(z) = \tanh\left(50(z-4.5)\right)$. Initial conditions are $(1,0,0)$. In words, $f(z)$ changes smoothly from -1 to 1, vanishing at $z = 4.5$. Hence, the frequency of oscillation undergoes fast oscillations with varying frequency, except close to $t = 4.5$. At this time, $f(z)$ vanishes and the system is no longer HiOsc. More precisely, trajectories go through a transition layer. Its width in this example is of order ϵ. The two slow variables are $I = x^2 + y^2$ and z.

Figure 6(A) shows the values of the state variables with $\epsilon = 10^{-4}$. Due to the resonance, the Poincaré method fails to capture the correct evolution of the slow variables when crossing the singular point $t = 4.5$. However, combining with parareal, the fine solution of parareal integrates the equation across the resonance and allows the multiscale method to proceed beyond the singularity. In Figure 6(B), the absolute error in the state variable does not decrease with iterations because the accuracy of phase alignment relies on the scale separation which does not exists near $t = 4.5$. We show, however, that the convergence in the state variable can be achieved with a slight modification of Algorithm 3.5. Figure 6(C) is obtained by skipping phase alignments, step 2(c)i and ii, and replacing step (c)iii with the naive correction (1.2) near $t = 4.5$. See Table 5 for parameters used in numerical solution.

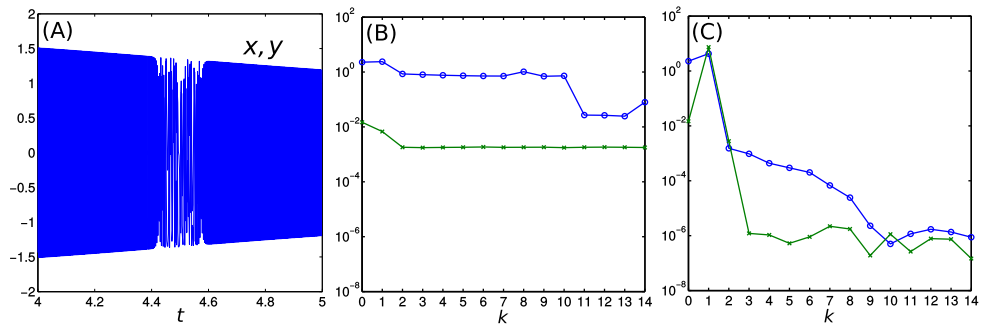

FIG. 6. *Passage through resonance, example* 5.4. (A) *The solution of* $x(t)$ *and* $y(t)$ *with* $t \in [4, 5]$ *and* $\epsilon = 10^{-4}$. *The frequency function vanishes at* $t = 4.5$ *and solutions lose their HiOsc nature.* (B) *The absolute errors of both state (circles) and slow (crosses) variables as a function of iterations with phase alignment at all times.* (C) *The absolute errors with phase alignment turned off near* $t = 4.5$. *Convergence in the state variable is achieved.*

TABLE 5
Parareal parameters in example 5.4.

ϵ	T	H	h_{fine}	$\eta_{Poincaré}$	$h_{Poincaré}$	RelTol, AbsTol (ODE45)
10^{-4}	7	1/4	$\epsilon/200$	15ϵ	$\epsilon/10$	10^{-13}, 10^{-11}

Remark 5.1. This example demonstrates that the applicability of the multiscale-parareal coupling using alignments may be wider than the convergence theory proved in section 4.

5.5. The Fermi–Pasta–Ulam problem. We consider a chain of $2k$ springs on a line, connected with alternating soft nonlinear and stiff linear springs with both ends fixed. This problem has been used as a benchmark for testing the long-time performance of geometric integrators [32]. See also [14, 15, 31] and references therein for related recent work. The model is derived from the following Hamiltonian:

$$H(p,q) = \frac{1}{2}\sum_{i=1}^{2k} p_i^2 + \frac{1}{4}\epsilon^{-2}\sum_{i=1}^{k}(q_{2i} - q_{2i-1})^2 + \sum_{i=1}^{k}(q_{2i+1} - q_{2i})^4.$$

Using the change of variables given in [5], the system can be written as

$$\begin{cases} \dot{y}_i = u_i, \\ \dot{x}_i = \epsilon^{-1} v_i, \\ \dot{u}_i = -(y_i - \epsilon x_i - y_{i-1} - \epsilon x_{i-1})^3 + (y_{i+1} - \epsilon x_{i+1} - y_i - \epsilon x_i)^3, \\ \dot{v}_i = -\epsilon^{-1} x_i + (y_i - \epsilon x_i - y_{i-1} - \epsilon x_{i-1})^3 + (y_{i+1} - \epsilon x_{i+1} - y_i - \epsilon x_i)^3. \end{cases} \tag{5.7}$$

Since the ends are fixed, $y_0 = x_0 = y_{k+1} = x_{k+1} = 0$.

In this section, we endeavor solving (5.7) in the $\mathcal{O}(\epsilon^{-1})$ time scale using the proposed methods. The aim is to expose the limit of the various different algorithmic components of the proposed methods. Clearly, the time scale in which we compute the solutions of (5.7) is out of the scope of the analysis that we presented earlier.

System (5.7) is solved in $\mathbb{R}^8$ with $k = 2$, and it admits seven slow variables—they are the total energies of the stiff springs, $I_i = x_i^2 + v_i^2$ for $i = 1, 2$, the relative phases

between the stiff springs, $\phi = x_1x_2 + v_1v_2$, and all the degrees of freedom which are related to the soft springs: y_i and u_i, $i = 1, 2$. The nontrivial energy transfer and the relative phase take place in the very long ϵ^{-1} time scale.

In Figure 7(A), we present the maximum errors in the state variable in a long time interval $[0, T] = [0, \epsilon^{-1}/2]$ with $\epsilon = 10^{-3}$. The system (5.7) is integrated using Algorithm 3.5, applying the global alignment algorithm described in section 4 to achieve convergence in the state variable. The results are computed with the initial conditions $y_1 = x_1 = y_2 = x_2 = 1$, $(u_1, v_1, u_2, v_2) = (0, 1.2, 1, 0)$, and with the parameters given in Table 6. The Verlet method was used as the fine integrator and the Poincaré second order method (Verlet macrosolver and ODE45 microsolver) as the multiscale integrator. The parareal results are compared with the solutions of the fine integrator using h_{fine} given in Table 6. We further point out that the errors will decrease as ϵ becomes smaller.

In Figure 7(B), we present the maximum errors computed using the same set of conditions and parameters, *except that the phase alignments are computed by the numerical solutions of the modified equation*:

(5.8)
$$\begin{cases} \dot{y}_i = 0 \cdot u_i, \\ \dot{x}_i = \epsilon^{-1} v_i, \\ \dot{u}_i = 0 \cdot \left\{ -(y_i - \epsilon x_i - y_{i-1} - \epsilon x_{i-1})^3 + (y_{i+1} - \epsilon x_{i+1} - y_i - \epsilon x_i)^3 \right\}, \\ \dot{v}_i = -\epsilon^{-1} x_i + 0 \cdot \left\{ (y_i - \epsilon x_i - y_{i-1} - \epsilon x_{i-1})^3 + (y_{i+1} - \epsilon x_{i+1} - y_i - \epsilon x_i)^3 \right\} \end{cases}$$

with $y_0 = x_0 = y_{k+1} = x_{k+1} = 0$. This is a slightly more general type of "unperturbed" equations as those defined in section 5. This approach may be thought of as phase alignments with some slow variables constrained. A detailed discussion about such a topic is out of the scope of this paper, and shall be addressed in a future work.

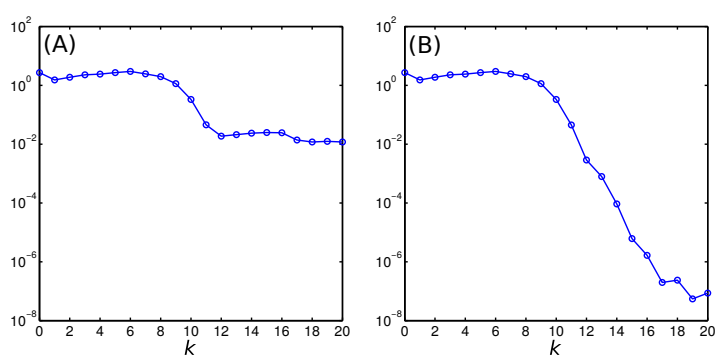

FIG. 7. *The Fermi–Pasta–Ulam problem, example* 5.5. *The absolute error in the state variables as a function of iteration is depicted.* (A) *Phase alignments with the full system* (5.7). (B) *Phase alignments with the unperturbed equation explained in the paragraph.*

TABLE 6
Parareal parameters in example 5.5.

ϵ	T	H	h_{fine}	$\eta_{Poincaré}$	$h_{Poincaré}$	RelTol, AbsTol (ODE45)
10^{-3}	$\epsilon^{-1}/2$	$1/4$	$\epsilon/20$	10ϵ	$\epsilon/20$	10^{-12}, 10^{-10}

Remark 5.2. We observe that in this study, phase alignments by solutions of (5.8) yield significantly superior results. Indeed, our experience indicated that for the problems that can be solved by the Poincaré method described in [2], it is generally better to use the so-called "unperturbed" equations for phase alignments, as the slow variables are constrained and would not deteriorate the performance of the phase alignment steps. However, we chose to present a more general setup in this paper, avoiding any specific choice of coarse integrators in describing the algorithms.

6. Summary. The paper describes two approaches to incorporate multiscale integrators as coarse integrators in parareal methods. The first, presented in section 3.1, approximates all the slow variables. However, the numerical approximation of the state variables u_n^k does not converge to the true solution $u(nH)$. This parareal-multiscale combination has several advantages compared to other multiscale schemes:

- Increased stability and lowered sensitivity to the choice of parameters.
- Increased accuracy: The accuracy of slow variables may be smaller than $\mathcal{O}(\epsilon)$, which is a theoretical limit for Poincaré and other multiscale methods that are based on averaging or homogenization principles.
- Applicability to systems with moderate scale separation. Most multiscale methods are more efficient than conventional, nonmultiscale schemes if the separation in scale is large enough, i.e., if ϵ is sufficiently small. However, they typically become less efficient or unstable at intermediate values of ϵ.
- The method may be used in situations in which the dynamics looses its multiscale structure in a short transition layer, for example, due to passage through resonance; see example 5.4.

The second approach, presented in section 3.2, computes a convergent approximation to all state variables in the system. This algorithm requires the phase alignment procedure, described in section 4 in addition to the steps needed in our first algorithm. We prove that the accuracy of the scheme in the sup norm after K iterations is of order $\epsilon^{-1}H^K + E_f$, where H is the a coarse step size. In particular, the number of iterations to achieve a given error tolerance is logarithmic in ϵ^{-1}.

The computational cost of the method can be grouped into two contributions. The first is the cost of the fine integrator invoked at each parareal iteration. It is proportional to $H\epsilon^{-1}$ for each iteration. The second contribution comes from the overhead of multiscale integrators and phase alignment. While this contribution is independent of ϵ, it grows linearly with the number of coarse step sizes, H^{-1}. Hence, there is a trade off in choosing H. With a large scale separation $\epsilon \ll 1$, the first contribution dominates and, assuming maximal parallelization is available, it is advantageous to use a relatively small H, even if the multiscale method allows larger steps. The two contributions balance if one takes $H = \sqrt{\epsilon}$, which implies a computational cost of order $K\epsilon^{-1/2}$. In contrast, parallel methods using conventional integrators will require at least $\mathcal{O}(\epsilon^{-1})$ steps.

The general approach for a multiscale-parareal coupling using alignments is not limited to the proposed particular implementation. While the alignment method described above assumes a single-frequency fast variable, different alignment algorithms may be possible in the multifrequency case. For example, if the slow-fast system is given explicitly, one could simply set the fast variable in the coarse integrator to coincide with the fine one. The related experiment was given in section 5.1. Seamless methods for systems with multiple high frequencies, which do not require full knowledge of the fast and slow coordinates, are beyond the scope of this paper.

REFERENCES

[1] A. Abdulle, W. E, B. Engquist, and E. Vanden-Eijnden, *The heterogeneous multiscale method*, Acta Numer., 21 (2012), pp. 1–87.

[2] G. Ariel, B. Engquist, S. Kim, Y. Lee, and R. Tsai, *A multiscale method for highly oscillatory dynamical systems using a Poincaré map type technique*, J. Sci. Comput., 54 (2013), pp. 247–268, http://dx.doi.org/10.1007/s10915-012-9656-x.

[3] G. Ariel, B. Engquist, S. J. Kim, and R. Tsai, *Iterated averaging of three-scale oscillatory systems*, Commun. Math. Sci., 12 (2014), pp. 791–824, http://dx.doi.org/10.4310/CMS.2014.v12.n5.a1.

[4] G. Ariel, B. Engquist, H.-O. Kreiss, and R. Tsai, *Multiscale computations for highly oscillatory problems*, in Multiscale Modeling and Simulation in Science, Lect. Notes Comput. Sci. Eng. 66, Springer, Berlin, 2009, pp. 237–287, http://dx.doi.org/10.1007/978-3-540-88857-4_5.

[5] G. Ariel, B. Engquist, and R. Tsai, *A multiscale method for highly oscillatory ordinary differential equations with resonance*, Math. Comp., 78 (2009), pp. 929–956.

[6] G. Ariel, B. Engquist, and R. Tsai, *Numerical multiscale methods for coupled oscillators*, Multiscale Model. Simul., 7 (2009), pp. 1387–1404.

[7] G. Ariel, B. Engquist, and R. Tsai, *A reversible multiscale integration method*, Commun. Math. Sci., 7 (2009), pp. 595–610.

[8] G. Ariel, B. Engquist, and R. Tsai, *Oscillatory systems with three separated time scales – analysis and computation*, in Numerical Analysis of Multiscale Computations, Lect. Notes Comput. Sci. Eng. 82, Springer, Berlin, 2011, pp. 23–45.

[9] G. Ariel, S. J. Kim, and R. Tsai, *Parareal methods for highly oscillatory ordinary differential equations*, preprint, arXiv:1503.02094, 2015.

[10] Z. Artstein, I. G. Kevrekidis, M. Slemrod, and E. S. Titi, *Slow observables of singularly perturbed differential equations*, Nonlinearity, 20 (2007), pp. 2463–2481.

[11] Z. Artstein, J. Linshiz, and E. S. Titi, *Young measure approach to computing slowly advancing fast oscillations*, Multiscale Model. Simul., 6 (2007), pp. 1085–1097, http://dx.doi.org.ezproxy.lib.utexas.edu/10.1137/070687219.

[12] C. Audouze, M. Massot, and S. Volz, *Symplectic Multi-time Step Parareal Algorithms Applied to Molecular Dynamics*, preprint, hal-00358459, 2009.

[13] G. Bal and Q. Wu, *Symplectic parareal*, in Domain Decomposition Methods in Science and Engineering XVII, Lect. Notes Comput. Sci. Eng. 60, Springer, Berlin, 2008, pp. 401–408, doi:10.1007/978-3-540-75199-1_51.

[14] D. Bambusi and A. Ponno, *On metastability in FPU*, Comm. Math. Phys., 264 (2006), pp. 539–561.

[15] D. Bambusi and A. Ponno, *Resonance, metastability and blow up in FPU*, in The Fermi-Pasta-Ulam Problem, Lecture Notes in Phys. 728, Springer, Berlin, 2008, pp. 191–205.

[16] A. Blouza, L. Boudin, and S. M. Kaber, *Parallel in time algorithms with reduction methods for solving chemical kinetics*, Commun. Appl. Math. Comput. Sci., 5 (2010), pp. 241–263.

[17] F. Casas and A. Iserles, *Explicit Magnus expansions for nonlinear equations*, J. Phys. A, 39 (2006), pp. 5445–5461, doi:10.1088/0305-4470/39/19/S07.

[18] M. Condon, *Efficient computation of delay differential equations with highly oscillatory terms*, ESAIM Math. Model. Numer. Anal., 46 (2012), pp. 1407–1420, doi:10.1051/m2an/2012004.

[19] M. Condon, A. Deano, and A. Iserles, *On second-order differential equations with highly oscillatory forcing terms*, R. Soc. Lond. Proc. A Math. Phys. Eng. Sci., 486 (2010), pp. 1809–1828, doi:10.1098/rspa.2009.0481.

[20] X. Dai, C. Le Bris, F. Legoll, and Y. Maday, *Symmetric parareal algorithms for Hamiltonian systems*, ESAIM Math. Model. Numer. Anal., 47 (2013), pp. 717–742.

[21] S. Das and J. A. Yorke, *Super convergence of ergodic averages for quasiperiodic orbits*, preprint, arXiv:1506.06810v3, 2016.

[22] W. E and B. Engquist, *The heterogeneous multiscale methods*, Commun. Math. Sci., 1 (2003), pp. 87–132.

[23] W. E, B. Engquist, X. Li, W. Ren, and E. Vanden-Eijnden, *Heterogeneous multiscale methods: A review*, Commun. Comput. Phys., 2 (2007), pp. 367–450.

[24] S. Engblom, *Parallel in time simulation of multiscale stochastic chemical kinetics*, SIAM Multiscale Model. Simul., 8 (2009), pp. 46–68.

[25] B. Engquist, A. Fokas, E. Hairer, and A. Iserles, *Highly Oscillatory Problems*, Cambridge University Press, New York, 2009.

[26] B. Engquist and Y.-H. Tsai, *Heterogeneous multiscale methods for stiff ordinary differential equations*, Math. Comp., 74 (2005), pp. 1707–1742.

[27] I. Fatkullin and E. Vanden-Eijnden, *A computational strategy for multiscale chaotic systems with applications to Lorenz 96 model*, J. Comput. Phys., 200 (2004), pp. 605–638.

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

[29] C. W. Gear and I. G. Kevrekidis, *Constraint-defined manifolds: A legacy code approach to low-dimensional computation*, J. Sci. Comput., 25 (2005), pp. 17–28.

[30] D. Givon, R. Kupferman, and A. M. Stuart, *Extracting macroscopic dynamics: Model problems and algorithms*, Nonlinearity, 17 (2004), pp. R55–R127.

[31] E. Hairer and C. Lubich., *On the energy distribution in Fermi-Pasta-Ulam lattices*, Arch. Rational Mech. Anal. 205 (2010), pp. 993–1029.

[32] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration*, Springer Ser. Comput. Math., Springer, Berlin, 2002.

[33] T. Haut and B. Wingate, *An asymptotic parallel-in-time method for highly oscillatory PDEs*, SIAM J. Sci. Comput. 36 (2014), pp. A693–A713.

[34] L. He, *The reduced basis technique as a coarse solver for parareal in time simulations*, J. Comput. Math, 28 (2010), pp. 676–692.

[35] D. D. Holm and P. Lynch, *Stepwise precession of the resonant swinging spring*, SIAM J. Appl. Dyn. Syst., 1 (2002), pp. 44–64, doi:10.1137/S1111111101388571.

[36] A. Iserles, *Think globally, act locally: Solving highly-oscillatory ordinary differential equations*, Appl. Numer. Math., 43 (2002), pp. 145–160, doi:10.1016/S0168-9274(02)00122-8.

[37] A. Iserles, *On the numerical quadrature of highly-oscillating integrals I: Fourier transforms*, IMA J. Numer. Anal., 24 (2004), pp. 365–391, doi:10.1093/imanum/24.3.365.

[38] A. Iserles and S. P. Norsett, *On quadrature methods for highly oscillatory integrals and their implementation*, BIT, 44 (2004), pp. 755–772, doi:10.1007/s10543-004-5243-3.

[39] A. Iserles and S. P. Norsett, *Efficient quadrature of highly oscillatory integrals using derivatives*, R. Soc. Lond. Proc. A Math. Phys. Eng. Sci., 461 (2005), pp. 1383–1399, doi:10.1098/rspa.2004.1401.

[40] H. Jiménez-Pérez and J. Laskar, *A time-parallel algorithm for almost integrable Hamiltonian systems*, preprint, arXiv:1106.3694, 2011.

[41] J. Kevorkian and J. D. Cole, *Perturbation Methods in Applied Mathematics*, Appl. Math. Sci. 34, Springer, New York, 1980.

[42] J. Kevorkian and J. D. Cole, *Multiple Scale and Singular Perturbation Methods*, Appl. Math. Sci. 114, Springer, New York, 1996.

[43] I. G. Kevrekidis and G. Samaey, *Equation-free multiscale computation: Algorithms and applications*, Annu. Rev. Phys. Chem., 60 (2009), pp. 321–344.

[44] M. Khanamiryan, *Quadrature methods for highly oscillatory linear and nonlinear systems of ordinary differential equations: Part I*, BIT, 48 (2008), pp. 743–761, doi:10.1007/s10543-008-0201-0.

[45] S. J. Kim, *Numerical Methods for Highly Oscillatory Dynamical Systems using Multiscale Structure*, PhD thesis, University of Texas at Austin, Austin, TX, 2013.

[46] S. J. Kim and H. Zhou, *A multiscale computation for highly oscillatory dynamical systems using empirical mode decomposition (EMD)–type methods*, Multiscale Modeling & Simulation, 14 (2016), pp. 534–557, doi:10.1137/15M1023038.

[47] H.-O. Kreiss, *Problems with different time scales for ordinary differential equations*, SIAM J. Numer. Anal., 16 (1979), pp. 980–998.

[48] H.-O. Kreiss, *Problems with different time scales*, Acta Numer., 1 (1992), pp. 101–139.

[49] H.-O. Kreiss and J. Lorenz, *Manifolds of slow solutions for highly oscillatory problems*, Indiana Univ. Math. J., 42 (1993), pp. 1169–1191.

[50] F. Legoll, T. Lelievre, and G. Samaey, *A micro-macro parareal algorithm: Application to singularly perturbed ordinary differential equations*, SIAM J. Sci. Comput., 35 (2013), pp. A1951–A1986.

[51] D. Levin, *Fast integration of rapidly oscillatory functions*, J. Comput. Appl. Math., 67 (1996), pp. 95–101, doi:10.1016/0377-0427(94)00118-9.

[52] J.-L. Lions, Y. Maday, and G. Turinici, *A "parareal" in time discretization of PDE's*, C. R. Acad. Sci. Paris Ser. I Math., 332 (2001), pp. 661–668.

[53] Y. Maday, *The parareal in time algorithm*, in Substructuring Techniques and Domain Decomposition Methods, Comput. Sci. Eng. Tech. Ser. 24, Saxe-Coburg, Stirlingshire, UK, 2010, pp. 19–42.

[54] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. Springer Ser. Oper. Res. Financ. Eng., Springer, New York, 2006.

[55] R. L. Petzold, O. J. Laurent, and Y. Jeng, *Numerical solution of highly oscillatory ordinary differential equations*, Acta Numer., 6 (1997), pp. 437–483.

[56] J. A. Sanders, F. Verhulst, and J. Murdock, *Averaging Methods in Nonlinear Dynamical Systems*, 2nd. ed., Appl. Math. Sci. 59, Springer, New York, 2007.

[57] R. Sharp, Y.-H. Tsai, and B. Engquist, *Multiple time scale numerical methods for the inverted pendulum problem*, in Multiscale Methods in Science and Engineering, Lect. Notes Comput. Sci. Eng. 44, Springer, Berlin, 2005, pp. 241–261.

[58] M. Tao, H. Owhadi, and J. E. Marsden, *Nonintrusive and structure preserving multiscale integration of stiff ODEs, SDEs, and Hamiltonian systems with hidden slow dynamics via flow averaging*, Multiscale. Mod. Simul., 8 (2010), pp. 1269–1324.