Journal of Computational Physics 436 (2021) 110282

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

Reduced model-based parareal simulations of oscillatory singularly perturbed ordinary differential equations

Laura Grigori [a], Sever A. Hirstoaga [a,*], Van-Thanh Nguyen [a], Julien Salomon [b]

[a] Inria, Project-Team ALPINES, Sorbonne Université and Université de Paris, CNRS, Laboratoire Jacques-Louis Lions (LJLL), 75589 Paris Cedex 12, France

[b] Inria, Project-Team ANGE, Sorbonne Université and Université de Paris, CNRS, Laboratoire Jacques-Louis Lions (LJLL), 75589 Paris Cedex 12, France

ARTICLE INFO

Article history:
Available online 16 March 2021

Keywords:
Parareal algorithm
Two-scale expansion
Multi-scale models
Vlasov characteristics
Electric and magnetic fields

ABSTRACT

We propose a new strategy for solving by the parareal algorithm highly oscillatory ordinary differential equations which are characteristics of a six-dimensional Vlasov equation. For the coarse solvers we use reduced models, obtained from the two-scale asymptotic expansions in [4]. Such reduced models have a low computational cost since they are free of high oscillations. The parareal method allows to improve their accuracy in a few iterations through corrections by fine solvers of the full model. We demonstrate the accuracy and the efficiency of the strategy in numerical experiments of short time and long time simulations of charged particles submitted to a large magnetic field. In addition, the convergence of the parareal method is obtained uniformly with respect to the vanishing stiff parameter.

1. Introduction

In this paper we propose a new coupling strategy in the parareal framework [14,15] to efficiently solve the following six dimensional dynamical system for $0 < \varepsilon \ll 1$

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{x}_\varepsilon}{\mathrm{d}t} = \mathbf{v}_\varepsilon, & \mathbf{x}_\varepsilon(s) = \mathbf{x}, \\ \dfrac{\mathrm{d}\mathbf{v}_\varepsilon}{\mathrm{d}t} = \dfrac{1}{\varepsilon}\big(\mathbf{v}_\varepsilon \times \mathbf{B}_\varepsilon(\mathbf{x}_\varepsilon)\big) + \mathbf{E}(t, \mathbf{x}_\varepsilon), & \mathbf{v}_\varepsilon(s) = \mathbf{v}, \end{cases} \tag{1}$$

where $(\mathbf{x}, \mathbf{v})$ is an initial condition given at the initial time $t = s$. The system in (1) models the dynamics of a charged particle under the influence of an external electro-magnetic field. This is a typical characteristic curve of the Vlasov equation. In this context, $\mathbf{x}_\varepsilon : \mathbb{R} \to \mathbb{R}^3$ stands for the position unknown, $\mathbf{v}_\varepsilon : \mathbb{R} \to \mathbb{R}^3$ for the velocity unknown, and $\mathbf{E} : \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R}^3$ and $\frac{1}{\varepsilon}\mathbf{B}_\varepsilon : \mathbb{R}^3 \to \mathbb{R}^3$ for a given electro-magnetic field. We assume $|\mathbf{B}_\varepsilon| = 1$ and that the mass and the charge particle are both equal to 1. The parameter $1/\varepsilon$ in front of the $\mathbf{v}_\varepsilon \times \mathbf{B}_\varepsilon(\mathbf{x}_\varepsilon)$ term means that the magnetic field is assumed high with respect to the electric term, in view of plasma confinement considerations [10]. More precisely, $1/\varepsilon$ denotes the strength

* Corresponding author.
E-mail addresses: laura.grigori@inria.fr (L. Grigori), sever.hirstoaga@inria.fr (S.A. Hirstoaga), van-thanh.nguyen@inria.fr (V.-T. Nguyen), julien.salomon@inria.fr (J. Salomon).

https://doi.org/10.1016/j.jcp.2021.110282
0021-9991/

of the magnetic field and thus, since the charge-to-mass ratio is assumed to be 1, the cyclotron frequency is also equal to $1/\varepsilon$. The difficulty of the problem is that the large magnetic field introduces a rapid time scale, the rotation of particles around the magnetic field line, which is much smaller than the one driven by the electric field. We are thus faced with a multi-scale problem whose numerical solution by standard methods requires high computational cost, since a standard but accurate enough numerical integrator requires time steps that are of the order of the fastest oscillation. This is an issue to be avoided in applications, and therefore, in this paper we are interested in solving equations in (1) with a time step which is not constraint by ε.

The parareal algorithm is an efficient method performing real time simulations with the help of parallel computing, for the numerical solving of a very large class of time dependent equations. The literature is huge, we cite only [14,15,3,7]. The method involves a fine expensive solver that is only applied in parallel, and a coarse but cheap solver which is used in sequential. A basic way to apply parareal in practice consists in taking large time steps Δt for a coarse solver and in refining the solutions in parallel using smaller time steps δt. This can reduce the computational time if the parareal iterations converge rapidly and if the ratio $\Delta t/\delta t$ is large.

However, when solving stiff equations like (1), regardless of the numerical scheme used for the coarse solver, the time step should satisfy $\Delta t \sim \varepsilon$ to achieve enough accuracy leading to a rapid convergence of the parareal scheme [8]. This is due to the high oscillations in time (with period of order ε) in the solution. Therefore, it can be interesting to use a different model to define the coarse solver in such a way that it remains computationally cheap but with a time step satisfying $\Delta t \gg \varepsilon$. Eventually, it is also important that the coarse solver be accurate enough so that the parareal iterations converge rapidly. In the case of equation (1), it is crucial for the coarse solver to provide an accurate approximation of the high oscillations, since otherwise the solver accumulates large errors, parareal requires a large number of iterations and thus the computational speed-up deteriorates. The purpose of our work is to obtain a convergent parareal algorithm with a large ratio $\Delta t/\delta t$ and a small number of iterations ($k \ll N$, see section 2 for notation).

In this paper, we use the parareal algorithm to efficiently integrate equation (1), by using a reduced model to define the coarse solver. Roughly speaking, such a reduced model reads

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{Y}}{\mathrm{d}t} = f(\mathbf{Y},\mathbf{U}), & \mathbf{Y}(s) = \mathbf{x}, \\ \dfrac{\mathrm{d}\mathbf{U}}{\mathrm{d}t} = g(\mathbf{Y},\mathbf{U}), & \mathbf{U}(s) = \mathbf{v}, \end{cases} \tag{2}$$

where $\mathbf{Y}$, $\mathbf{U}$ are used to approximate $\mathbf{x}_\varepsilon$, $\mathbf{v}_\varepsilon$ thanks to

$$\big(\mathbf{x}_\varepsilon(t), \mathbf{v}_\varepsilon(t)\big) \sim Z\Big((t-s)/\varepsilon, \big(\mathbf{Y}(t), \mathbf{U}(t)\big)\Big) \quad \text{when } \varepsilon \to 0,$$

and where Z is an operator for which an explicit form is to be derived in practice.

Specifically, we illustrate the idea above with an example in a similar context, as detailed in [9]. Thus, if instead of equation (1) we consider

$$\frac{\mathrm{d}\mathbf{u}_\varepsilon}{\mathrm{d}t} + \frac{1}{\varepsilon} L\mathbf{u}_\varepsilon = N(\mathbf{u}_\varepsilon), \quad \mathbf{u}_\varepsilon(0) = u_0,$$

where L is a linear operator and $N(\mathbf{u}_\varepsilon)$ is a specific nonlinear term, then it is well-known that under suitable assumptions, the solution $\mathbf{u}_\varepsilon(t)$ has the asymptotic approximation $\mathbf{u}_\varepsilon(t) = \exp(-\frac{t}{\varepsilon}L)\overline{\mathbf{u}}(t) + \mathcal{O}(\varepsilon)$, where the slowly varying function $\overline{\mathbf{u}}$ is the solution to the reduced model

$$\frac{\mathrm{d}\overline{\mathbf{u}}}{\mathrm{d}t} = \overline{N}(\overline{\mathbf{u}}), \quad \overline{\mathbf{u}}(0) = u_0, \tag{3}$$

where the $\overline{N}(\overline{\mathbf{u}})$ term is obtained by time averaging

$$\overline{N}(\overline{\mathbf{u}}(t)) = \lim_{T\to\infty} \frac{1}{T} \int_0^T e^{\theta L} N(e^{-\theta L}\overline{\mathbf{u}}(t))d\theta.$$

It is important to note that $\overline{\mathbf{u}}(t)$ and its derivatives are formally bounded independently of ε and therefore, large time steps $\Delta t \gg \varepsilon$ can be taken to solve (3) (see [9]).

In our approach, the reduced model is obtained through a two-scale asymptotic expansion and is proved to provide an accurate approximation of the initial equation when the small parameter ε vanishes [4]. More precisely, we use either a zero-th order or a first order two-scale model, depending on the availability of practical equations. Indeed, it is possible that the first order model is too complex to be solved, analytically or numerically and in this case, only the zero-th order model will be considered. These models have two advantages: the low computational cost and the capability to give a good approximation of the high frequency oscillations through the operator Z enclosing the smallest scale, under the assumption that these oscillations are periodic and can be analytically computed [4].

The idea of using a different model for the coarse solver is not new. As an example, a similar approach has been used in chemical kinetics [15], where a reduction of a linear kinetic system with multiple scales was applied for the coarse solving. In [13], a slow manifold projector is used as coarse solver for solving ordinary differential equations with dissipative stiffness. We also mention two contributions closely related to our work. First, a parareal method for PDEs with linear high oscillating term is proposed in [9]. On the basis of a classical averaged model, the method needs exact knowledge of the fast variable to obtain a convergent parareal algorithm. Though our strategy also assumes the period of the fastest motion to be known, the high oscillating term is not necessarily linear in our case (see section 5). Additionally, we consider first-order asymptotic terms which provide a more accurate averaged model and can accelerate the convergence of the parareal method. Second, in the frame of models similar to (1), a multi-scale method for solving the slow evolution was successfully used as coarse solver in [1], without requiring explicit knowledge of the fast and slow variables. However a tedious alignment algorithm is required to achieve convergence of parareal. Specifically, the method needs to make the alignment of the fast phases of the coarse and fine solvers with sufficient efficiency and accuracy. On the contrary, our reduced model accurately synchronizes phase with the fine solution, which allows us to avoid such an alignment algorithm to get the numerical convergence of the parareal method. The drawback of our approach is that the small period of the fast oscillation must be known. However, this particular framework covers several models which solve interesting problems in plasma physics, as illustrated in section 6.

The paper is organized as follows. In section 2, we briefly present the parareal method and describe our strategy as applied in the context of stiff ordinary differential equations. In section 3, we introduce the two-scale asymptotic expansion at the base of the reduced models and we justify their use as coarse solvers of our parareal algorithm. In sections 4 and 5, we present three ODE models which enter into the general form (1) and derive their first order and zero-th order reduced equations following [4]. The full equations under consideration apply to different models in plasma physics: isotope separation by ion cyclotron resonance, storing of charged particles in a Penning trap, and an example of charged particle confinement by strong variable magnetic field. For these models, we present in section 6 numerical experiments that show that the parareal strategy provides accurate results together with computational efficiency through parallelism.

2. The parareal algorithm

Introduced in 2001 [14], the parareal (parallel in real time) algorithm displays its advantage by covering various fields of applications where it exploits very efficiently parallel computing over a large number of processors to solve problems in real time constraint context. Since its conception, the algorithm has been intensively analyzed [2,8,16,6]. Let us briefly recall this approach. Consider the simple time dependent problem

$$\frac{\mathrm{d}u}{\mathrm{d}t} = f(u) \ \text{ in } (0,T), \ u(0) = \mathrm{u}_0. \tag{4}$$

The time interval $[0,T]$ is decomposed into N uniform time slices $[T_n, T_{n+1}]$, for $n \in \{0,\dots,N-1\}$. Let $\mathcal{F}(T_{n+1}, T_n, U_n)$ denote the fine solver, which gives a very accurate approximation of the solution at time T_{n+1} with the initial solution U_n at time T_n and let $\mathcal{G}(T_{n+1}, T_n, U_n)$ denote the coarse solver, which gives a coarse approximation of the solution at time T_{n+1} also with the initial solution U_n at time T_n. The coarse solver is to be chosen in such a way that, its cost is much lower than the one of the fine solver. A popular strategy consists in using the approximation method considered in the fine solver but with a larger time step [8]. Alternatively, one can use an approximation method with lower accuracy, or even use a different model from the original problem as long as it can give a reasonable coarse and fast approximation of the solution of the original problem [15].

In this paper, we follow the latter approach and focus on the idea of using a reduced model of the original problem for the coarse solver. For that reason, the coarse solver $\mathcal{G}(T_{n+1}, T_n, U_n)$ is always assigned to the solution of the reduced model (2) and the fine solver $\mathcal{F}(T_{n+1}, T_n, U_n)$ is always assigned to the (approximated) solution of the original problem (1). In addition, we let the coarse propagator perform a single time step per time slice $[T_n, T_{n+1}]$.

The parareal algorithm aims at computing a sequence $(U_n^k)_{k,n}$ of approximations of $u(T_n)$ for $n \in \{0,\dots,N\}$ for every k in the following way. At the first step, the initial approximation U_n^0 at coarse time points $0 = T_0 < T_1 < \cdots < T_N = T$ can be computed sequentially using the coarse solver that reads

$$U_{n+1}^0 = \mathcal{G}(T_{n+1}, T_n, U_n^0), \ U_0^0 = u_0,$$

and then for $k = 0, 1, \dots$ with $U_0^{k+1} = u_0$, the parareal algorithm computes a more accurate approximation

$$U_{n+1}^{k+1} = \mathcal{G}(T_{n+1}, T_n, U_n^{k+1}) + \mathcal{F}(T_{n+1}, T_n, U_n^k) - \mathcal{G}(T_{n+1}, T_n, U_n^k).$$

In this iteration, the terms $\mathcal{F}(T_{n+1}, T_n, U_n^k)$ have the largest computational cost. Therefore, all these fine computations could be performed in parallel over each interval $[T_n, T_{n+1}]$, the main goal of parareal being to speed up the computing time. However, in order to achieve a real speed-up, the algorithm should converge in a number of iterations significantly smaller than the number of time intervals.

3. Two-scale asymptotic expansion

In this section, we summarize the principles and the main result of two-scale asymptotic expansion allowing to obtain reduced models. The equation (1) is a particular instance of the more general singularly perturbed dynamical system

$$\frac{\mathrm{d}\mathcal{X}_\varepsilon}{\mathrm{d}t} = \mathbf{a}(t, \mathcal{X}_\varepsilon) + \frac{1}{\varepsilon}\mathbf{b}(t, \mathcal{X}_\varepsilon), \qquad \mathcal{X}_\varepsilon(s) = \mathcal{X}, \tag{5}$$

where $\mathcal{X}_\varepsilon : \mathbb{R} \to \mathbb{R}^d$ and $\mathbf{a}$ and $\mathbf{b}$ are given fields satisfying suitable assumptions and s plays the role of the initial time. Following [4], we briefly recall the asymptotic two-scale expansion method in order to approximate the solution $\mathcal{X}_\varepsilon(t)$ when $\varepsilon \to 0$. Under regularity assumptions on $\mathbf{a}$ and $\mathbf{b}$ and assuming the solution $\mathbf{Z}(t; \theta, \mathbf{z})$ to equation

$$\frac{\mathrm{d}\mathbf{Z}}{\mathrm{d}\theta} = \mathbf{b}(t, \mathbf{Z}), \qquad \mathbf{Z}(t; 0, \mathbf{z}) = \mathbf{z} \tag{6}$$

to be periodic in θ, for every $t \in \mathbb{R}$ and every $\mathbf{z} \in \mathbb{R}^d$, it is proved in [4] that $\mathcal{X}_\varepsilon$ admits the following two-scale expansion in time

$$\mathcal{X}_\varepsilon(t) = \mathcal{X}^0\Big(t, \frac{t-s}{\varepsilon}\Big) + \varepsilon\, \mathcal{X}^1\Big(t, \frac{t-s}{\varepsilon}\Big) + \varepsilon^2\, \mathcal{X}^2\Big(t, \frac{t-s}{\varepsilon}\Big) + \dots \tag{7}$$

when $\varepsilon \to 0$ and where the functions $\mathcal{X}^i(t, \theta)$ are periodic in θ for every $i \in \mathbb{N}$. In this setting, ordinary differential equations characterizing the terms of the expansion (7) are derived in [4, Theorems 1.1 & 1.3]. In addition, strong convergence theorems are proved, justifying the approximation results asserting that, *e.g.*, at the zero-th order we have

$$\mathcal{X}_\varepsilon(t) \sim \mathcal{X}^0\Big(t, \frac{t-s}{\varepsilon}\Big), \quad \text{when } \varepsilon \to 0,$$

and at the first order,

$$\mathcal{X}_\varepsilon(t) \sim \mathcal{X}^0\Big(t, \frac{t-s}{\varepsilon}\Big) + \varepsilon\, \mathcal{X}^1\Big(t, \frac{t-s}{\varepsilon}\Big), \quad \text{when } \varepsilon \to 0.$$

For the sake of completeness, we give below the result concerning the two-scale limit model or the zero-th order approximation [4, Theorem 1.1] in the case of a six dimensional space ($d = 6$).

Theorem 3.1. *We assume that*[1] $\mathbf{a} \in \big(C_b^1(\mathbb{R} \times \mathbb{R}^6)\big)^6$ *and* $\mathbf{b} \in \big(C_b^2(\mathbb{R} \times \mathbb{R}^6)\big)^6$. *Assume also that the solution of* (6) *is* 2π*-periodic in* θ*, for every* $t \in \mathbb{R}$ *and every* $\mathbf{z} \in \mathbb{R}^6$*. Then, for every initial condition* $\mathcal{X} \in \mathbb{R}^6$*, every* $\varepsilon > 0$*, and every* $\Delta S > 0$*, the solution* $\mathcal{X}_\varepsilon$ *of* (5) *exists on* $[s, s + \Delta S]$*, is unique and satisfies*

$$\lim_{\varepsilon \to 0} \sup_{t \in [s, s+\Delta S]} \left| \mathcal{X}_\varepsilon(t) - \mathcal{X}^0\left(t, \frac{t-s}{\varepsilon}\right) \right| = 0, \tag{8}$$

where $|\cdot|$ *stands for the Euclidean norm on* $\mathbb{R}^6$ *and* $\mathcal{X}^0$ *satisfies*

$$\mathcal{X}^0(t, \theta) = \mathbf{Z}\big(t; \theta, \mathcal{Y}^0(t)\big) \tag{9}$$

and where $\mathcal{Y}^0$ *is the solution to*

$$\frac{\mathrm{d}\mathcal{Y}^0}{\mathrm{d}t} = \alpha(t, \mathcal{Y}^0), \qquad \mathcal{Y}^0(s) = \mathcal{X}, \tag{10}$$

with α *defined by*

$$\alpha(t, \mathcal{Y}) = \frac{1}{2\pi} \int_0^{2\pi} \{\nabla \mathbf{Z}(t; \theta, \mathcal{Y})\}^{-1} \Big\{ \mathbf{a}(t, \mathbf{Z}(t; \theta, \mathcal{Y})) - \frac{\partial \mathbf{Z}}{\partial t}(t; \theta, \mathcal{Y}) \Big\}\, d\theta.$$

Remark 3.2. We remark that the limit model in (10) does not contain high oscillations in time so that cheap numerical schemes can be used to compute $\mathcal{Y}^0$. Then, when $\mathbf{Z}$ is known in (6), we obtain the term $\mathcal{X}^0$ by (9), as an approximation of the solution $\mathcal{X}_\varepsilon$ in the sense of (8). Though obtained at a low computational cost, the approximation $\mathcal{X}^0$ still contains information about the high oscillations in the solution, through the operator $\mathbf{Z}$.

These facts underline that the solution to the limit model given by (9)-(10) is a good candidate for a coarse solving in the parareal framework.

[1] C_b^m stands for the space of continuous functions with bounded derivatives to the order m.

In the subsequent sections, we develop this framework for equations of the type of equation (1), by using the notation $\mathcal{X} = (\mathbf{x}, \mathbf{v})^T$, where, as in classical mechanics, $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)^T$ stands for the position vector and $\mathbf{v} = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)^T$ for the velocity vector. In this setting, it is important to note the particular form of the system (6). The solution $\mathbf{Z}$ captures only the rotation of the particle velocity following the magnetic field: $\mathbf{Z} = (\mathbf{x_Z}, \mathbf{v_Z})^T$ is the solution to

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{x_Z}}{\mathrm{d}\theta} = 0, & \mathbf{x_Z}(0) = \mathbf{x}, \\ \dfrac{\mathrm{d}\mathbf{v_Z}}{\mathrm{d}\theta} = \mathbf{v_Z} \times \mathbf{B}(\mathbf{x_Z}), & \mathbf{v_Z}(0) = \mathbf{v}. \end{cases}$$

This motion is assumed to be 2π-periodic in the theory we use. We denote in the sequel the cyclotron period in time by $P = 2\pi\varepsilon$ and the cyclotron frequency by $1/\varepsilon$, which are associated to the full system (1).

4. The case of a constant magnetic field

In this section we consider equation (1) provided with a constant magnetic field $\mathbf{B}_\varepsilon = \overrightarrow{e_1}$, where $\{\overrightarrow{e_1}, \overrightarrow{e_2}, \overrightarrow{e_3}\}$ is the frame of $\mathbb{R}^3$ and with a given external electric field. In this way, the term $\mathbf{v}_\varepsilon \times \mathbf{B}_\varepsilon(\mathbf{x}_\varepsilon)$ in the velocity equation of (1) writes $(\mathbf{v}_\varepsilon)^\perp = (0, (\mathbf{v}_\varepsilon)_3, -(\mathbf{v}_\varepsilon)_2)^T$. Thus, we can see that the basic assumption of periodicity of the solution of (6) is satisfied. The common feature of the test cases we treat in this section is that we can compute analytically the solutions of equation (1) and of the corresponding reduced model. Therefore, when applying the parareal algorithm we will be able to use the exact flows for the fine and the coarse solvers.

4.1. A uniform time varying electric field

In this section, we take an electric field which is only highly oscillating in time. In this case, system (1) writes

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{x}_\varepsilon}{\mathrm{d}t} = \mathbf{v}_\varepsilon, & \mathbf{x}_\varepsilon(s) = \mathbf{x}, \\ \dfrac{\mathrm{d}\mathbf{v}_\varepsilon}{\mathrm{d}t} = \dfrac{1}{\varepsilon}(\mathbf{v}_\varepsilon)^\perp + \mathbf{E}\Big(\dfrac{t}{\varepsilon}\Big), & \mathbf{v}_\varepsilon(s) = \mathbf{v}, \end{cases} \tag{11}$$

where $\mathbf{E}$ has the form $\mathbf{E}(\tau) = \big(E_1, E_2(\tau), E_3(\tau)\big)^T$, with $E_1 \in \mathbb{R}$ and E_2, E_3 are 2π-periodic functions, see [4, Section 3.1]. This system can be used for modelling ion cyclotron resonance with application in isotope separation in plasmas, see [5] and the references therein. In magnetized plasmas, the ions are heated by an oscillating perpendicular electric field at frequencies corresponding to the ion cyclotron frequency. Thus, the cyclotron resonance leads to a growth of the amplitude of motion in time. In the sequel, we consider for illustration the following electric field

$$E_1 = 0, \quad E_2(\tau) = \sin(\tau), \quad E_3(\tau) = \cos(\tau). \tag{12}$$

However, all the following results can be derived in a similar form for a general electric field with the above properties. Next, we need the following matrices denoted by

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, R(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}, \mathcal{R}(\theta) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sin\theta & 1-\cos\theta \\ 0 & \cos\theta - 1 & \sin\theta \end{pmatrix}. \tag{13}$$

It is convenient to put the solution of (11)-(12) in the form

$$\begin{pmatrix} \mathbf{x}_\varepsilon(t) \\ \mathbf{v}_\varepsilon(t) \end{pmatrix} = \mathcal{A} \begin{pmatrix} \mathbf{x} \\ \mathbf{v} \end{pmatrix} + \mathcal{B}, \tag{14}$$

where the 6×6 matrix $\mathcal{A}$ and the vector $\mathcal{B}$ are given by

$$\mathcal{A} = \begin{pmatrix} I_3 & (t-s)P + \varepsilon\mathcal{R}\big(\frac{t-s}{\varepsilon}\big) \\ O_3 & R\big(\frac{t-s}{\varepsilon}\big) \end{pmatrix} \text{ and}$$

$$\mathcal{B} = \begin{pmatrix} \varepsilon(t-s)\big(0, -\cos(t/\varepsilon), \sin(t/\varepsilon)\big)^T + \varepsilon^2\big(0, \sin(t/\varepsilon) - \sin(s/\varepsilon), \cos(t/\varepsilon) - \cos(s/\varepsilon)\big)^T \\ (t-s)\big(0, \sin(t/\varepsilon), \cos(t/\varepsilon)\big)^T \end{pmatrix},$$

with I_3 the 3×3 identity matrix and O_3 the 3×3 zero matrix.

Next, we derive the reduced model for equation (11). We apply [4, Theorems 3.1, 3.2] to equations (11)-(12) to obtain the first order two-scale model. The approximation of the solution is

$$\mathbf{G}(t) = \begin{pmatrix} \mathbf{x}^0\Big(t, \dfrac{t-s}{\varepsilon}\Big) \\ \mathbf{v}^0\Big(t, \dfrac{t-s}{\varepsilon}\Big) \end{pmatrix} + \varepsilon \begin{pmatrix} \mathbf{x}^1\Big(t, \dfrac{t-s}{\varepsilon}\Big) \\ \mathbf{v}^1\Big(t, \dfrac{t-s}{\varepsilon}\Big) \end{pmatrix}, \tag{15}$$

where the terms in the expansion are given by

$$\begin{pmatrix} \mathbf{x}^0(t,\theta) \\ \mathbf{v}^0(t,\theta) \end{pmatrix} = \begin{pmatrix} \mathbf{y}^0(t) \\ R(\theta)\,\mathbf{u}^0(t) \end{pmatrix} \tag{16}$$

and

$$\begin{pmatrix} \mathbf{x}^1(t,\theta) \\ \mathbf{v}^1(t,\theta) \end{pmatrix} = \begin{pmatrix} \mathbf{y}^1(t) + \mathcal{R}(\theta)\,\mathbf{u}^0(t) \\ R(\theta)\,\mathbf{u}^1(t) + R(\theta)\Big(\displaystyle\int_0^\theta d\sigma - \frac{\theta}{2\pi}\int_0^{2\pi} d\sigma\Big)\big(R(-\sigma)\mathbf{E}(\sigma)\big) \end{pmatrix}. \tag{17}$$

Then, in the particular case of the electric field in (12), we have that $\big(\mathbf{y}^0(t), \mathbf{u}^0(t)\big)$ is solution to

$$\frac{\mathrm{d}\mathbf{y}^0}{\mathrm{d}t} = \begin{pmatrix} (\mathbf{u}^0)_1 \\ 0 \\ 0 \end{pmatrix}, \frac{\mathrm{d}\mathbf{u}^0}{\mathrm{d}t} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ and } \begin{cases} \mathbf{y}^0(s) = \mathbf{x}, \\ \mathbf{u}^0(s) = \mathbf{v}, \end{cases} \tag{18}$$

with $(\mathbf{x}, \mathbf{v})$ the initial condition in (11) and that $\big(\mathbf{y}^1(t), \mathbf{u}^1(t)\big)$ is solution to

$$\frac{\mathrm{d}\mathbf{y}^1}{\mathrm{d}t} = \begin{pmatrix} (\mathbf{u}^1)_1 \\ -1 \\ 0 \end{pmatrix}, \frac{\mathrm{d}\mathbf{u}^1}{\mathrm{d}t} = 0 \text{ and } \begin{cases} \mathbf{y}^1(s) = 0, \\ \mathbf{u}^1(s) = 0. \end{cases} \tag{19}$$

Equations (18)-(19) are easy to solve, their solutions are

$$\begin{cases} \mathbf{y}^0(t) = \big(\mathbf{v}_1(t-s) + \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\big)^T, \\ \mathbf{u}^0(t) = \big(\mathbf{v}_1, \mathbf{v}_2, (t-s) + \mathbf{v}_3\big)^T, \end{cases}$$

and respectively

$$\begin{cases} \mathbf{y}^1(t) = \big(0, -(t-s), 0\big)^T, \\ \mathbf{u}^1(t) = \big(0, 0, 0\big)^T. \end{cases}$$

Replacing these formulas in (16)-(17) and getting the result in (15) we obtain the analytical form of the first-order two-scale approximation $\mathbf{G}(t)$. However, it is interesting to write $\mathbf{G}$ as the solution of the original system was derived in equation (14). We have

$$\mathbf{G}(t) = \mathcal{A}\begin{pmatrix} \mathbf{x} \\ \mathbf{v} \end{pmatrix} + \mathcal{C}, \tag{20}$$

where the matrix $\mathcal{A}$ is as in (14) and $\mathcal{C}$ is given by

$$\mathcal{C} = \begin{pmatrix} \varepsilon(t-s)\big(0, -\cos((t-s)/\varepsilon), \sin((t-s)/\varepsilon)\big)^T \\ (t-s)\big(0, \sin((t-s)/\varepsilon), \cos((t-s)/\varepsilon)\big)^T \end{pmatrix}.$$

Remark 4.1. We notice that in the general case where $\mathbf{E}$ has the form

$$\mathbf{E}(\tau) = \big(E_1, E_2(\tau), E_3(\tau)\big)^T, \text{ with } E_1 \in \mathbb{R} \text{ and } 2\pi\text{-periodic functions } E_2, E_3,$$

the solutions of the full model and of the reduced one keep similar expressions to those in (14) and (20) respectively. More precisely, the matrix $\mathcal{A}$ will be the same, the difference appearing in the vectors $\mathcal{B}$ and $\mathcal{C}$ which will contain averages in the fast variable against $\sin(\cdot)$ and $\cos(\cdot)$ of the functions E_2 and E_3.

4.2. A non uniform stationary electric field

In this part, we consider an electric field which is not dependent of time but of space and we use the framework in [4, Section 3.2]. In this case, system (1) writes

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{x}_\varepsilon}{\mathrm{d}t} = \mathbf{v}_\varepsilon, & \mathbf{x}_\varepsilon(s) = \mathbf{x}, \\ \dfrac{\mathrm{d}\mathbf{v}_\varepsilon}{\mathrm{d}t} = \dfrac{1}{\varepsilon}(\mathbf{v}_\varepsilon)^\perp + \mathbf{E}(\mathbf{x}_\varepsilon), & \mathbf{v}_\varepsilon(s) = \mathbf{v}. \end{cases} \tag{21}$$

Then, following a standard strategy we can find an explicit form of a linear application $\mathbf{E}$ leading to highly oscillating solution but which is bounded in time. Nevertheless, by taking in (21) the electric field given by

$$\mathbf{E}(\mathbf{x}) = c \begin{pmatrix} -\mathbf{x}_1 \\ \mathbf{x}_2/2 \\ \mathbf{x}_3/2 \end{pmatrix}, \tag{22}$$

with an arbitrary constant $c > 0$, the system describes the dynamics of a charged particle in an ideal Penning trap [12] (we fix to 1 both the charge and the mass of the particle). Under the condition $\varepsilon < \sqrt{1/(2c)}$, the solution of (21)-(22) is

$$\begin{aligned} \mathbf{x}_\varepsilon(t) &= \begin{pmatrix} c_1 \cos(\sqrt{c}\,(t-s)) + c_2 \sin(\sqrt{c}\,(t-s)) \\ a_1 \sin(a_\varepsilon(t-s)) - a_2 \cos(a_\varepsilon(t-s)) + b_1 \sin(b_\varepsilon(t-s)) - b_2 \cos(b_\varepsilon(t-s)) \\ a_1 \cos(a_\varepsilon(t-s)) + a_2 \sin(a_\varepsilon(t-s)) + b_1 \cos(b_\varepsilon(t-s)) + b_2 \sin(b_\varepsilon(t-s)) \end{pmatrix}, \\ \mathbf{v}_\varepsilon(t) &= \frac{\mathrm{d}\mathbf{x}_\varepsilon}{\mathrm{d}t}(t), \end{aligned} \tag{23}$$

where

$$a_\varepsilon = \frac{1 + \sqrt{1 - 2c\varepsilon^2}}{2\varepsilon}, \quad b_\varepsilon = \frac{1 - \sqrt{1 - 2c\varepsilon^2}}{2\varepsilon}, \tag{24}$$

and $a_1, a_2, b_1, b_2, c_1, c_2$ are constants to be found from the initial condition.

Remark 4.2.

1. A Penning trap is a device for storing charged particles using a homogeneous magnetic field and an inhomogeneous quadrupole electric field. The constant c in (22) entails the geometry of the trap and the voltage between the electrodes, while $1/\varepsilon$ is the magnitude of the magnetic field. The condition for having a stable periodic trajectory [12] is

$$\frac{1}{\varepsilon} > \sqrt{2c}. \tag{25}$$

   Otherwise, the particle escapes from the trap due to a magnetic field which is weaker than the electric field. This corresponds to a solution with growing amplitude of motion in time.
2. We notice that the three frequencies $\sqrt{c}$, a_ε, and b_ε are denoted in literature [12] by ω_x, ω_+, and ω_- respectively, and they verify the relation

$$\omega_\pm = \frac{1}{2}\left(\omega_{cy} \pm \sqrt{\omega_{cy}^2 - 2\omega_x^2}\right),$$

   where ω_{cy} is the cyclotron frequency. In our notation $\omega_{cy} = 1/\varepsilon$.
3. It is clear that the motion in the $\overrightarrow{e_1}$ direction is decoupled from the motion in the other two directions. More precisely, a charged particle performs in an ideal Penning trap three independent motions with characteristic frequencies: a modified cyclotron motion (at frequency ω_+), the axial motion (at frequency ω_x), and the magnetron motion or the $\mathbf{E} \times \mathbf{B}$ drift (at frequency ω_-).
4. We have $a_\varepsilon \sim \frac{1}{\varepsilon}$ and $b_\varepsilon \sim \varepsilon$ when $\varepsilon \to 0$. Therefore, the solution in (23) oscillates in time at three scales, $2\pi\varepsilon$, 1 and $2\pi/\varepsilon$. In addition, we can identify initial conditions leading to solutions which are oscillating at the desired scale(s) by equating to zero the corresponding coefficients.

Next, we derive the reduced model for equation (21). More precisely, we apply [4, Theorem 3.3] to write the specific first order two-scale model to the system (21)-(22). Recalling the formula in [4, Theorem 3.3], the first-order approximation of the solution to the model (21)-(22) is given by

$$\mathbf{G}(t) = \begin{pmatrix} \mathbf{x}^0\Big(t, \dfrac{t-s}{\varepsilon}\Big) \\ \mathbf{v}^0\Big(t, \dfrac{t-s}{\varepsilon}\Big) \end{pmatrix} + \varepsilon \begin{pmatrix} \mathbf{x}^1\Big(t, \dfrac{t-s}{\varepsilon}\Big) \\ \mathbf{v}^1\Big(t, \dfrac{t-s}{\varepsilon}\Big) \end{pmatrix}, \tag{26}$$

where, as in section 4.1, the terms in the expansion are given by

$$\begin{pmatrix} \mathbf{x}^0(t,\theta) \\ \mathbf{v}^0(t,\theta) \end{pmatrix} = \begin{pmatrix} \mathbf{y}^0(t) \\ R(\theta)\,\mathbf{u}^0(t) \end{pmatrix} \tag{27}$$

and

$$\begin{pmatrix} \mathbf{x}^1(t,\theta) \\ \mathbf{v}^1(t,\theta) \end{pmatrix} = \begin{pmatrix} \mathbf{y}^1(t) + \mathcal{R}(\theta)\,\mathbf{u}^0(t) \\ R(\theta)\,\mathbf{u}^1(t) + \mathcal{R}(\theta)\,\mathbf{E}(\mathbf{y}^0(t)) \end{pmatrix}. \tag{28}$$

Then, in the particular case of the electric field in (22), we have that $\big(\mathbf{y}^0(t), \mathbf{u}^0(t)\big)$ is solution to

$$\frac{\mathrm{d}\mathbf{y}^0}{\mathrm{d}t} = \begin{pmatrix} (\mathbf{u}^0)_1 \\ 0 \\ 0 \end{pmatrix}, \frac{\mathrm{d}\mathbf{u}^0}{\mathrm{d}t} = \begin{pmatrix} -c(\mathbf{y}^0)_1 \\ 0 \\ 0 \end{pmatrix} \text{ and } \begin{cases} \mathbf{y}^0(s) = \mathbf{x}, \\ \mathbf{u}^0(s) = \mathbf{v}, \end{cases} \tag{29}$$

with $(\mathbf{x}, \mathbf{v})$ the initial condition in (21) and that $\big(\mathbf{y}^1(t), \mathbf{u}^1(t)\big)$ is solution to

$$\frac{\mathrm{d}\mathbf{y}^1}{\mathrm{d}t} = \begin{pmatrix} (\mathbf{u}^1)_1 \\ \frac{c}{2}(\mathbf{y}^0)_3 \\ -\frac{c}{2}(\mathbf{y}^0)_2 \end{pmatrix}, \frac{\mathrm{d}\mathbf{u}^1}{\mathrm{d}t} = \begin{pmatrix} -c(\mathbf{y}^1)_1 \\ -\frac{c}{2}(\mathbf{u}^0)_3 \\ \frac{c}{2}(\mathbf{u}^0)_2 \end{pmatrix} \text{ and } \begin{cases} \mathbf{y}^1(s) = 0, \\ \mathbf{u}^1(s) = 0. \end{cases} \tag{30}$$

Equations (29)-(30) are easy to solve, their solutions are

$$\begin{cases} \mathbf{y}^0(t) = \big(\mathbf{x}_1 \cos(\sqrt{c}\,(t-s)) + \frac{\mathbf{v}_1}{\sqrt{c}} \sin(\sqrt{c}\,(t-s)),\ \mathbf{x}_2,\ \mathbf{x}_3\big)^T, \\ \mathbf{u}^0(t) = \big(-\mathbf{x}_1\sqrt{c}\sin(\sqrt{c}\,(t-s)) + \mathbf{v}_1\cos(\sqrt{c}\,(t-s)),\ \mathbf{v}_2,\ \mathbf{v}_3\big)^T, \end{cases}$$

and respectively

$$\begin{cases} \mathbf{y}^1(t) = \big(0,\ \frac{c}{2}\mathbf{x}_3(t-s),\ -\frac{c}{2}\mathbf{x}_2(t-s)\big)^T, \\ \mathbf{u}^1(t) = \big(0,\ -\frac{c}{2}\mathbf{v}_3(t-s),\ \frac{c}{2}\mathbf{v}_2(t-s)\big)^T. \end{cases}$$

Replacing (27)-(28) in (26), we obtain

$$\mathbf{G}(t) = \begin{pmatrix} \mathbf{y}^0(t) \\ R(\frac{t-s}{\varepsilon})\mathbf{u}^0(t) \end{pmatrix} + \varepsilon \begin{pmatrix} \mathbf{y}^1(t) + \mathcal{R}(\frac{t-s}{\varepsilon})\mathbf{u}^0(t) \\ R(\frac{t-s}{\varepsilon})\mathbf{u}^1(t) + \mathcal{R}(\frac{t-s}{\varepsilon})\mathbf{E}(\mathbf{y}^0(t)) \end{pmatrix}, \tag{31}$$

and thus, getting the analytic expressions of $\mathbf{y}^0, \mathbf{u}^0, \mathbf{y}^1, \mathbf{u}^1, \mathbf{E}$ and of matrices R and $\mathcal{R}$ in the above formula leads to the analytic form of the approximation $\mathbf{G}(t)$ to the solution $\big(\mathbf{x}_\varepsilon(t), \mathbf{v}_\varepsilon(t)\big)$ when ε is small enough and at any time $t \in [s, s+\Delta S]$. The obtained formula will be used in section 6 for the coarse solver.

5. The case of a variable magnetic field

In this section we study the case of a magnetic field with a strong part which is variable and a bounded part which is constant (see [4, Section 3.4]). In addition, we restrict to the case without electric field. More precisely, we consider equation (1) in the form

$$\begin{cases} \dfrac{\mathrm{d}\mathbf{x}_\varepsilon}{\mathrm{d}t} = \mathbf{v}_\varepsilon, & \mathbf{x}_\varepsilon(s) = \mathbf{x}, \\ \dfrac{\mathrm{d}\mathbf{v}_\varepsilon}{\mathrm{d}t} = \dfrac{1}{\varepsilon}\big(\mathbf{v}_\varepsilon \times \mathcal{M}(\mathbf{x}_\varepsilon)\big) + \mathbf{v}_\varepsilon \times \overrightarrow{e_3}, & \mathbf{v}_\varepsilon(s) = \mathbf{v}, \end{cases} \tag{32}$$

where

$$\mathcal{M}(\mathbf{x}) = \frac{1}{\sqrt{\mathbf{x}_1{}^2 + \mathbf{x}_2{}^2}} \begin{pmatrix} -\mathbf{x}_2 \\ \mathbf{x}_1 \\ 0 \end{pmatrix}.$$

We first notice that the assumption on the 2π-periodicity of the solution $\mathbf{Z}$ to equation (6) is satisfied. Then, unlike the test cases in section 4, we do not have an analytic expression for the solution of equation (32). The reduced model we will use

in the parareal method for this case, is the two-scale limit and not the first order approximation. The reason is that using the first order term in the asymptotic expansion becomes almost impossible due to its complex form (see [4, Theorem 3.6 & Appendix A]).

Next, we detail the two-scale limit model approximating equation (32) when $\varepsilon \to 0$. Following [4, Theorem 3.6], the limit term in the expansion is given by

$$\begin{pmatrix} \mathbf{x}^0(t,\theta) \\ \mathbf{v}^0(t,\theta) \end{pmatrix} = \begin{pmatrix} \mathbf{y}^0(t) \\ \mathcal{Z}_{\mathbf{v}}\big(t;\theta,\mathbf{y}^0(t),\mathbf{u}^0(t)\big) \end{pmatrix},$$

where the components of $\mathcal{Z}(t;\theta,\mathbf{x},\mathbf{v}) = \big(\mathcal{Z}_{\mathbf{x}}(t;\theta,\mathbf{x},\mathbf{v}), \mathcal{Z}_{\mathbf{v}}(t;\theta,\mathbf{x},\mathbf{v})\big)^T$ are $\mathcal{Z}_{\mathbf{x}}(t;\theta,\mathbf{x},\mathbf{v}) = \mathbf{x}$ and $\mathcal{Z}_{\mathbf{v}}(t;\theta,\mathbf{x},\mathbf{v}) = C(\theta,\mathbf{x})\mathbf{v}$, with

$$C(\theta,\mathbf{x}) = \begin{pmatrix} \frac{\mathbf{x}_1{}^2\cos\theta+\mathbf{x}_2{}^2}{\mathbf{x}_1{}^2+\mathbf{x}_2{}^2} & \frac{\mathbf{x}_1\mathbf{x}_2(\cos\theta-1)}{\mathbf{x}_1{}^2+\mathbf{x}_2{}^2} & -\frac{\mathbf{x}_1\sin\theta}{\sqrt{\mathbf{x}_1{}^2+\mathbf{x}_2{}^2}} \\ \frac{\mathbf{x}_1\mathbf{x}_2(\cos\theta-1)}{\mathbf{x}_1{}^2+\mathbf{x}_2{}^2} & \frac{\mathbf{x}_2{}^2\cos\theta+\mathbf{x}_1{}^2}{\mathbf{x}_1{}^2+\mathbf{x}_2{}^2} & -\frac{\mathbf{x}_2\sin\theta}{\sqrt{\mathbf{x}_1{}^2+\mathbf{x}_2{}^2}} \\ \frac{\mathbf{x}_1\sin\theta}{\sqrt{\mathbf{x}_1{}^2+\mathbf{x}_2{}^2}} & \frac{\mathbf{x}_2\sin\theta}{\sqrt{\mathbf{x}_1{}^2+\mathbf{x}_2{}^2}} & \cos\theta \end{pmatrix}$$

and where $\big(\mathbf{y}^0(t),\mathbf{u}^0(t)\big)$ is solution to

$$\frac{\mathrm{d}\mathbf{y}^0}{\mathrm{d}t} = \overline{A}(\mathbf{y}^0)\mathbf{u}^0, \quad \frac{\mathrm{d}\mathbf{u}^0}{\mathrm{d}t} = \overline{\beta}(\mathbf{y}^0,\mathbf{u}^0) \quad \text{and} \quad \begin{cases} \mathbf{y}^0(s) = \mathbf{x}, \\ \mathbf{u}^0(s) = \mathbf{v}, \end{cases} \tag{33}$$

with $(\mathbf{x},\mathbf{v})$ the initial condition in (32) and with

$$\overline{A}(\mathbf{y}) = \frac{1}{\mathbf{y}_1{}^2+\mathbf{y}_2{}^2}\begin{pmatrix} \mathbf{y}_2{}^2 & -\mathbf{y}_1\mathbf{y}_2 & 0 \\ -\mathbf{y}_1\mathbf{y}_2 & \mathbf{y}_1{}^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ \overline{\beta}(\mathbf{y},\mathbf{u}) = \begin{pmatrix} \frac{\mathbf{u}_2(\mathbf{u}_1\mathbf{y}_2-\mathbf{u}_2\mathbf{y}_1)}{\mathbf{y}_1{}^2+\mathbf{y}_2{}^2} \\ \frac{\mathbf{u}_1(\mathbf{u}_2\mathbf{y}_1-\mathbf{u}_1\mathbf{y}_2)}{\mathbf{y}_1{}^2+\mathbf{y}_2{}^2} \\ 0 \end{pmatrix}.$$

Thus, in this case, the approximation $\mathbf{G}(t)$ to the solution $\big(\mathbf{x}_\varepsilon(t),\mathbf{v}_\varepsilon(t)\big)$ when ε is small enough, is obtained first by solving the system (33) and then

$$\mathbf{G}(t) = \begin{pmatrix} \mathbf{y}^0(t) \\ C\big(\frac{t-s}{\varepsilon},\mathbf{y}^0(t)\big)\mathbf{u}^0(t) \end{pmatrix}. \tag{34}$$

6. Numerical results

First, in section 6.1 we analyze the time interval of validity and the accuracy of the reduced models for each test case. Then, we present numerical experiments illustrating the convergence of the parareal algorithm. In all the cases we consider, we obtained the numerical convergence with a number of parareal iterations K much smaller than the number N of the time slices of the interval $[0,T]$.

The reduced models that we use are zero-th or first order approximations of the initial stiff equation until a final time of order 1. The parareal algorithm allows us to perform simulations in long times, of order $1/\varepsilon$ or larger, by using the reduced model on time intervals where the latter is proved to be valid. For each case, we study the convergence of the algorithm when ε is fixed and also when making the parameter ε vanishing. This last issue is meaningful from the application viewpoint, since in realistic plasma physics phenomena such parameters are not fixed to a single value during the simulation but they can decrease in time.

6.1. Validity of the reduced models

The theorems from [4] prove convergence over time intervals of length 1 of the original models to the reduced models when the parameter ε vanishes. Therefore we cannot expect, in theory, that the approximation be valid over intervals of length $1/\varepsilon$ or larger. In addition, to the best of our knowledge, there are no estimates for the rate of convergence. In this section, we consequently assess numerically the quality of approximation of the reduced models in valid final times, *i.e.* in times of order $\mathcal{O}(1)$. We then check how large the final time can be such that the reduced models still provide satisfactory approximations. To this end, we plot the relative error

$$\mathrm{Error}(T_n) = \frac{\|\mathcal{G}_n - \mathcal{X}(T_n)\|_1}{\|\mathcal{X}(T_n)\|_1}, \tag{35}$$

where $\|\cdot\|_1$ stands for the ℓ_1 norm in $\mathbb{R}^6$, $\mathcal{G}_n$ stands for the reduced model solution at time T_n and $\mathcal{X}(T_n)$ stands for the original model solution at time T_n. Recall that $\mathcal{G}_n$ and $\mathcal{X}(T_n)$ have analytic forms for the test case in section 4.2, whereas

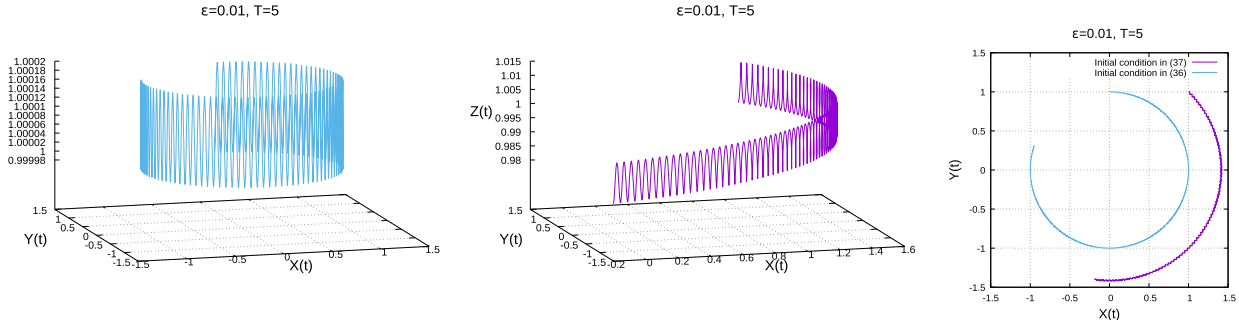

Fig. 1. The position trajectories until final time 5 of two particles: (36) at the left panel and (37) at the centre, following the model in (32). The projection of their motion on the perpendicular plane to $\vec{e_3}$ is at the right panel.

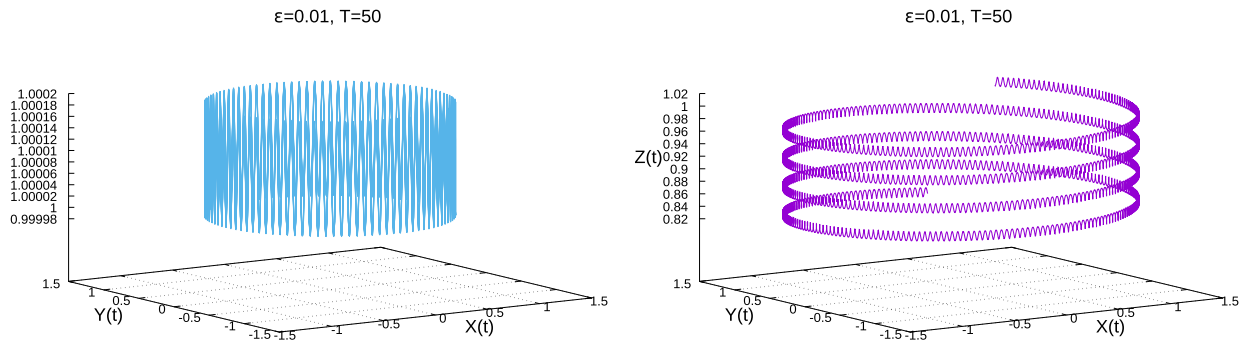

Fig. 2. The position trajectories until final time 50 of two particles: (36) at the left panel and (37) at the right, following the model in (32).

numerical approximations are used for both for the test case in section 5. Next, we do not discuss the case described in section 4.1 since writing the solutions of the original and the reduced models in the forms (14) and (20) respectively, shows that parareal convergence occurs after one iteration (see next section).

For the case considered in section 4.2 we recall that both the original and the first-order reduced models have analytic solutions given by (23) and (31). First, we remark that the exact solutions corresponding to the values of $\varepsilon \in \{0.1, 0.04\}$ are not well-approximated by the reduced model, see Fig. 3. We can see that beyond the final time $T = 50$ the approximations are not acceptable anymore. In contrast, for $\varepsilon = 0.01$ or smaller, the relative error is below 0.1 till the final time $T = 2500$, meaning that ε is small enough so that the reduced model provides a good approximation. Thus, if $\varepsilon = 0.01$, we obtain an acceptable relative error at the final time $T = 2500$, which means almost 40000 cyclotron orbits.

We now consider the test case of section 5. Here, we solve both models numerically, since no analytic expressions of their solutions are available. More precisely, we solve the system (32) by the symmetric and volume-preserving method G^4 of order 4 described in [11] and the limit model in (33) by the explicit Runge-Kutta 4 method. We use for the limit model approximation a time step equal to 0.625, whereas for the original dynamics we use a time step about $2\pi\varepsilon/80$ to accurately solve the cyclotron motion. We consider two initial conditions

$$\mathbf{x} = (0, 1, 1)^T, \quad \mathbf{v} = (1, \varepsilon, 0)^T \tag{36}$$

and

$$\mathbf{x} = (1, 1, 1)^T, \quad \mathbf{v} = (1, \varepsilon, 0)^T. \tag{37}$$

Let us do some qualitative remarks about the trajectories of both particles. First, notice that the solutions obtained with these initial conditions behave differently: for the first particle, the solution oscillates at two time scales (a rapid oscillation of order ε and a slower oscillation of order 1) whereas for the second one, the solution entails additionally a slow motion, consisting of a linear drift in the $\vec{e_3}$ direction (see Fig. 1). Also, the amplitude of the rapid oscillation in position in the $\vec{e_3}$ direction is of order ε^2 for one particle and of order ε for the other.

Then, as we can deduce from (33), the two-scale limit model does not capture the motion in the $\vec{e_3}$ direction, providing only an approximation of the projected motion on the perpendicular plane to $\vec{e_3}$. Thus, the limit model misses the $\vec{e_3}$-drift motion of the particle in (37). Eventually, the right panel in Fig. 1 shows that the planar angular velocity of the particle in (36) is larger than that of the particle in (37). We plot the relative error of the reduced (limit) model for several values of ε in Fig. 4. We can see that the behaviour of the error displays significant difference between these two initial conditions. More precisely, we observe that at final time $T = 100$, the reduced model does not provide a good approximation of the

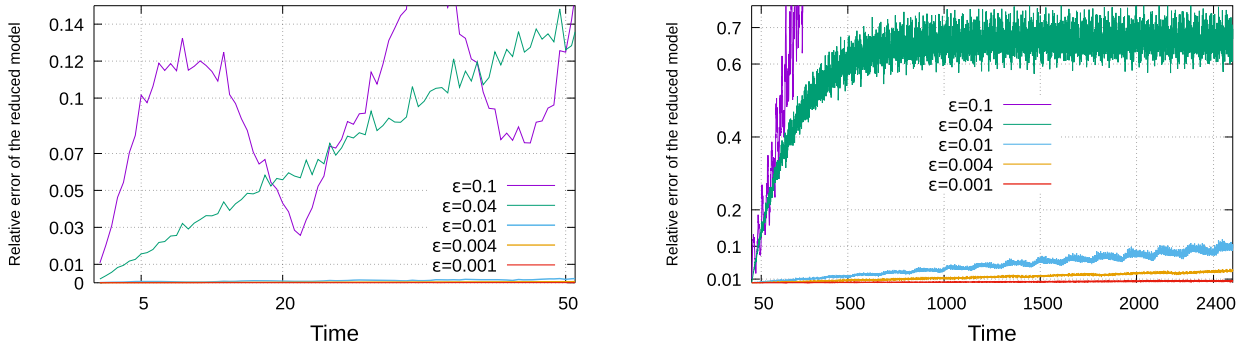

Fig. 3. Evolution with respect to time of the relative errors of the reduced model solution in (31) with respect to the solution in (23) with the initial condition in (38), in short time (at left) and long time (at right), for several values of ε.

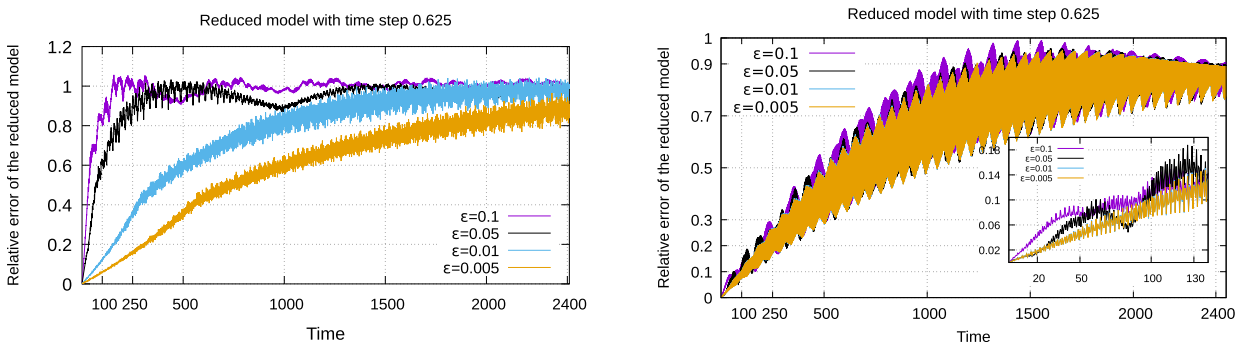

Fig. 4. Evolution with respect to time of the relative errors of the numerical approximation of the reduced model in (34) with respect to the numerical solution of (32) with the initial condition in (37) (at left) and that in (36) (at right), for several values of ε.

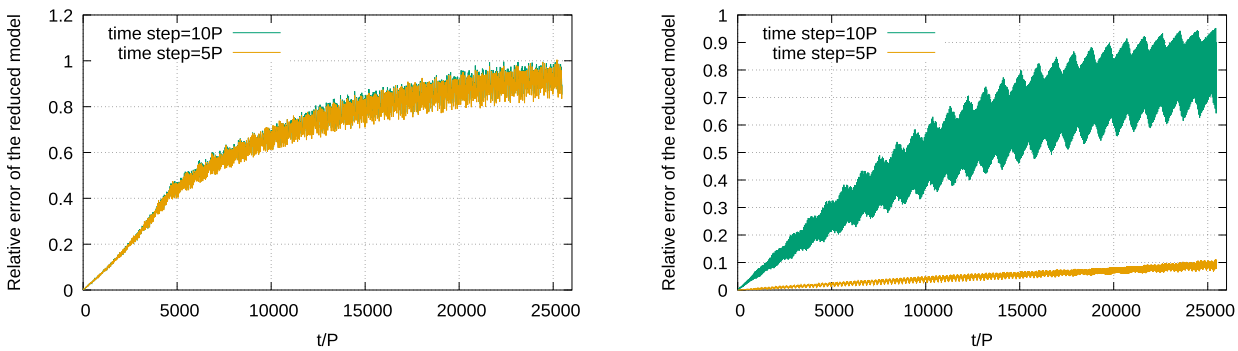

Fig. 5. Evolution with respect to time of the relative errors of the reduced model in (34) with respect to the solution of (32) with the initial condition in (37) (at left) and that in (36) (at right). The fast cyclotron period is denoted by $P = 2\pi\varepsilon$ where $\varepsilon = 0.01$. Two time steps for the reduced model are used: $0.625 \sim 10P$ and $0.3125 \sim 5P$.

original model when $\varepsilon \in \{0.1, 0.05\}$ in the case of the initial condition given by (37). On the contrary, when the initial condition is given by (36), the error is acceptable. However, for both particles, we deduce from Fig. 4 that the errors are large for times of order 2500, for any value of ε. In addition, when diminishing the time step for the numerical solver of the reduced model, we observe that the error drastically decreases when using the initial condition in (36). This result does not hold for the initial condition in (37), see Fig. 5.

In conclusion, we obtained for the first test case that the reduced model accurately approximates the original model in large times if ε is sufficiently small. For the second test case the reduced model fails to approximate the original dynamics in long times for every considered value of ε. We show in the next sections that the parareal algorithm allows to enhance the situation.

Table 1
Numbers of cyclotron periods ($P = 2\pi/a_\varepsilon$) enclosed in a time step of the coarse solver for several values of N. We have $T = 5$, $\Delta t = T/N$, and $\varepsilon = 0.01$.

	$N=2$	$N=4$	$N=8$	$N=16$	$N=32$	$N=64$	$N=128$
$\Delta t/P$	39.79	19.89	9.95	4.97	2.49	1.24	0.62

6.2. The test cases with strong constant magnetic field

First, we discuss about the case in section 4.1 of a uniform but time varying electric force in equation (11). Assume we fix an initial condition and we fix ε to a small value, say $\varepsilon = 0.01$. Then we use the exact solutions in (14) and (20) for the fine propagator $\mathcal{F}$ and respectively the coarse solver $\mathcal{G}$. We observe that the parareal algorithm writes in this case

$$U_{n+1}^{k+1} = \mathcal{F}(T_{n+1}, T_n, U_n^{k+1}), \quad \forall n \in \{0, \dots, N-1\}, \forall k \geq 0.$$

In particular, for $k = 1$ we have

$$U_{n+1}^{1} = \mathcal{F}(T_{n+1}, T_n, U_n^{1}), \quad \forall n \in \{0, \dots, N-1\}$$

and therefore, since the exact flows are used for the propagators $\mathcal{F}$ and $\mathcal{G}$, the parareal algorithm provides an exact solution in one iteration. Though easy to solve, this test case underlines the strength of the strategy: thanks to the writing of the original and reduced flows as (14) and (20) respectively, the use of the reduced model through the parareal algorithm leads to high accuracy in one iteration, whereas the error of the reduced model alone is very large (of order 1, following our simulations when ε is fixed to $\varepsilon = 0.01$).

We now treat the case in section 4.2. We consider the initial condition

$$\mathbf{x} = (1, 1, 1)^T, \quad \mathbf{v} = (1, 1, 1)^T \tag{38}$$

for solving the model (21)-(22). We set $c = 2$ and we vary ε verifying (25). The solution issued from this initial condition oscillates at three definite time scales (see Remark 4.2).

• We first fix $\varepsilon = 0.01$. As a first approach, we apply the parareal method in a standard way, meaning that we use for the coarse propagator $\mathcal{G}$ the classical Runge-Kutta 4 method for the initial model, with a bigger time step than that for the fine propagator. However, the coarse time step still needs to solve the smallest scale in order to have stability and reasonable accuracy. In this case, we have only to investigate the needed number of the parareal iterations for achieving convergence. More precisely, we first set the final time $T = 2\pi\varepsilon$ (one rapid oscillation), $N \in \{8, 16\}$ (larger N is not interesting), $\Delta t = T/N$, the number of coarse time steps on each time slice $MG = 1$ and the number of fine time steps on each time slice $MF = 80/N$. Thus, the fine time step $\delta t = T/80$ is fixed with respect to N and additionally is small enough for capturing the smallest scale. We plot at the top of Fig. 6 the relative error (in $L^\infty[0, T]$) between the solution $\mathcal{X}(t_n)$ obtained with the fine solver and the parareal solution U_n^k, as a function of the number k of parareal iterations

$$\text{Error}(k) = \frac{\max_{n\in\{1,\dots,N\}} \|U_n^k - \mathcal{X}(t_n)\|_1}{\max_{n\in\{1,\dots,N\}} \|\mathcal{X}(t_n)\|_1}, \tag{39}$$

where $\|\cdot\|_1$ stands for the ℓ_1 norm in $\mathbb{R}^6$. We obtain convergence of the algorithm for small k (4 or 5), however in a case of a too small T from the application point of view. When taking a larger final time $T = 8\pi\varepsilon$ (4 oscillations) with $MF = 320/N$ and $\delta t = T/320$, we have convergence of parareal for k very close to N (see the bottom of Fig. 6 for $N \in \{8, 16\}$). We can conclude that this parareal strategy provides convergence after $k \lesssim N$ iterations and with a ratio $\Delta t/\delta t \sim 1$, which is not an interesting approach.

• We now propose to use for the coarse solver $\mathcal{G}$ the reduced model in section 4.2 and we thus make use of the analytic expression in (31). In addition, we use for the fine solver $\mathcal{F}$ the explicit form of the solution in (23). We start by illustrating the convergence of the algorithm in short time simulations. We fix the final time $T = 5$ and the interval $[0, T]$ is partitioned in $N \in \{2, 4, 8, 16, 32, 64, 128\}$ sub-intervals. The big time step is thus $\Delta t = T/N$. It is interesting to note the size of the coarse time step Δt with respect to the small cyclotron period P, when N varies (see Table 1). Larger is $\Delta t/P$, larger is the ratio $\Delta t/\delta t$ and thus, cheaper is the coarse propagator.

We plot in Fig. 7 the relative error in $L^\infty[0, T]$ defined in (39) by taking the solution in (23) for $\mathcal{X}(t_n)$. The case $k = 0$ corresponds to the relative error of the solution of the reduced model with respect to the exact solution of the original model. We obtained convergence of the algorithm after a maximum of 6 iterations for all the considered values of N.

• We now analyze the behaviour of the parareal algorithm when ε decreases, in which case the reduced model becomes a more accurate approximation for the initial equation. We display in Fig. 10 the relative errors illustrating the convergence of the parareal algorithm. We plot for each value of N in the set $\{8, 16, 32, 64\}$ the errors for several values of ε at final time $T = 500\varepsilon$ which corresponds to approximately 80 cyclotron periods. As expected, the initial errors of the parareal

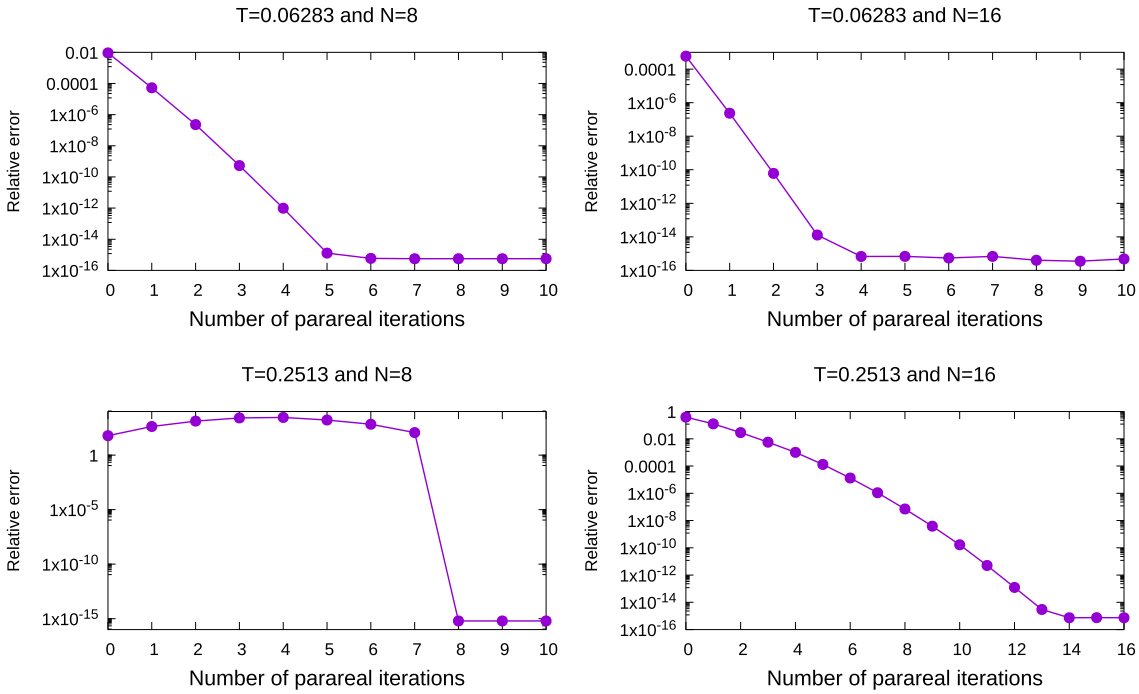

Fig. 6. Convergence rate of standard parareal algorithm for the test case in section 4.2.

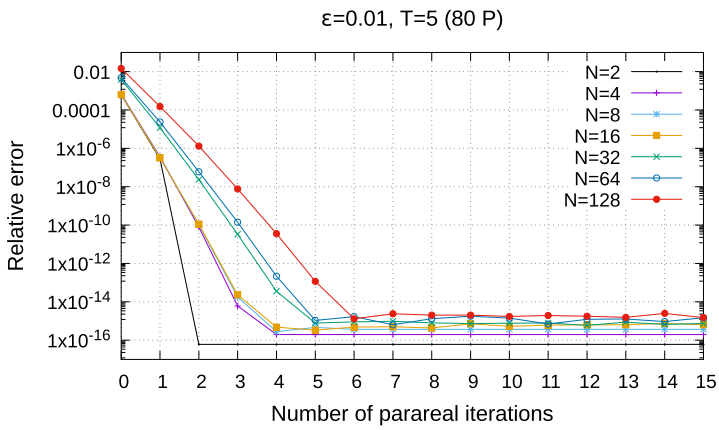

Fig. 7. Convergence of the parareal algorithm for the Penning trap test case at short final time $T = 5$. The fast cyclotron period is denoted by $P \sim 2\pi\varepsilon$.

method (*i.e.* $k = 0$) are decreasing when ε becomes smaller. Also, the smaller is ε, the faster is the convergence of the parareal algorithm since the better is the approximation of the reduced model. We already observed in section 6.1 that for $\varepsilon \in \{0.1, 0.04\}$, the reduced model induces a much bigger error than for the other smaller values of ε. However, with the parareal strategy we obtain acceptable convergence results for $\varepsilon = 0.1$, which entails a $\mathcal{O}(1)$ error for the reduced model: the parareal method convergences after $k = 10$ (resp. $k = 14$) iterations when $N = 32$ (resp. $N = 64$). Except for the values of $\varepsilon \in \{0.1, 0.04\}$, the convergence of the parareal algorithm for all the considered values of N is obtained after a maximum of $k = 5$ iterations. We also emphasize the achievement of an uniform error with respect to ε.

• Then, we consider the more challenging case of a long time simulation (of order $1/\varepsilon$). We fix the final time $T = 600 \sim 2\pi/b_\varepsilon$, where b_ε is defined in (24) and we take N in the set $\{120, 240, 480, 960\}$. As previously, we plot in Fig. 8 the relative errors between the exact solution and the parareal solution, as a function of the number k of parareal iterations. For this case, we can conclude with underlying the strength of using the parareal algorithm. The reduced model is not proved to be an approximation of the initial model in time of order $1/\varepsilon$. However, in a few numbers of parareal iterations we obtain high accuracy by applying the reduced model on valid intervals. Thus, if parallelism is to be used, the computational cost in the case of $N = 480$ (resp. $N = 960$) could drastically be reduced, achieving a round-off error in only 7 parareal iterations. In this case, a time slice includes approximately 20 (resp. 10) rapid oscillations.

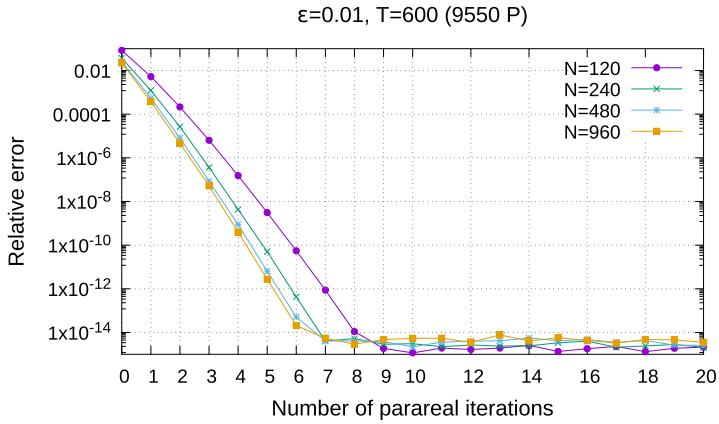

Fig. 8. Convergence of the parareal algorithm for the Penning trap test case at final time $T = 600$. The fast cyclotron period is denoted by $P \sim 2\pi\varepsilon$.

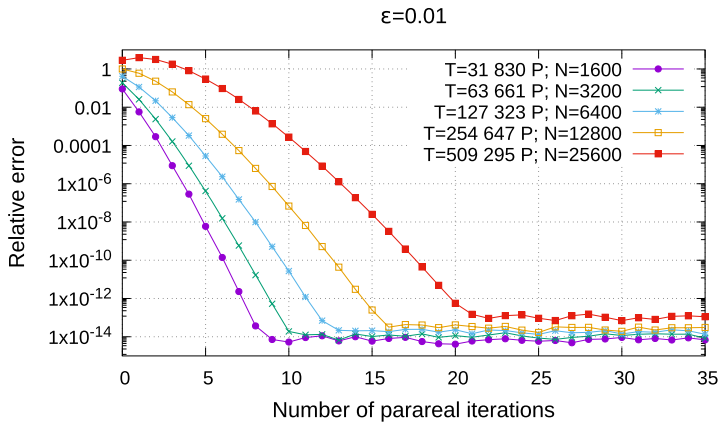

Fig. 9. Convergence of the parareal algorithm for the Penning trap test case when the coarse time step is kept constant to 1.25 and the final time T is increasing with N. The fast cyclotron period is denoted by $P \sim 2\pi\varepsilon$.

• Finally, we show the outcome of much longer simulations, where we keep the coarse time step fixed while the final time is increased with N. This framework is relevant for applications where one needs to integrate over very long times. We fix $\varepsilon = 0.01$ and the coarse time step to $\Delta t = 1.25$. The final time T is chosen in the set $\{2000, 4000, 8000, 16000, 32000\}$, see Fig. 9. Setting $N = 25600$, we observe that when $T = 32000$, *i.e.* T larger than 500000 cyclotron orbits, the convergence of the parareal algorithm is obtained in $k = 21$ iterations, with an error around 10^{-13}. In our opinion this is an excellent result which is due to the accuracy of the reduced model. Being of first order, the model provides good approximations of the slow motion and of the fast oscillation.

6.3. The test case with strong variable magnetic field

We now consider solving the problem in (32) with the initial conditions in (36) and (37). As in the previous section, we discuss the results of our simulations when $\varepsilon = 0.01$ in final times of order 1 and $1/\varepsilon$, and then we perform simulations in short final times by varying the values of ε. We recall that we use a symmetric and volume-preserving scheme and the classical Runge-Kutta 4 method for the models in (32) and (33) respectively, for the fine and respectively the coarse propagators. However, while the $\mathcal{F}$ propagator needs a time step δt which is a fraction of the rapid oscillation ($P \sim 2\pi\varepsilon$), the $\mathcal{G}$ propagator is computed with a time step Δt much larger than $2\pi\varepsilon$ (see typical values in Table 1).

• We first set $\varepsilon = 0.01$. We fix the final time $T = 5$ and we partitioned the interval $[0, T]$ in $N \in \{2, 4, 8, 16, 32, 64, 128\}$ sub-intervals. The coarse time step is $\Delta t = T/N$ and the fine time step is fixed to $\delta t = T/6400$, which is sufficiently small to solve the rapid oscillation. We plot at the top of Fig. 11 the relative error defined in (39) of the parareal solution with respect to the reference one computed with the $\mathcal{F}$ propagator. We observe convergence after a maximum of $k = 9$ iterations when $N \in \{32, 64, 128\}$ which could lead to satisfactory speed-up if parallel processing is set up.

• Then, we plot in Fig. 12 the errors of the parareal algorithm when ε vanishes. For a fixed number of time slices of $[0, T = 500\varepsilon]$ we perform simulations when ε goes in $\{0.01, 0.04, 0.01, 0.004, 0.001, 0.0004, 0.0001\}$. We find numerically

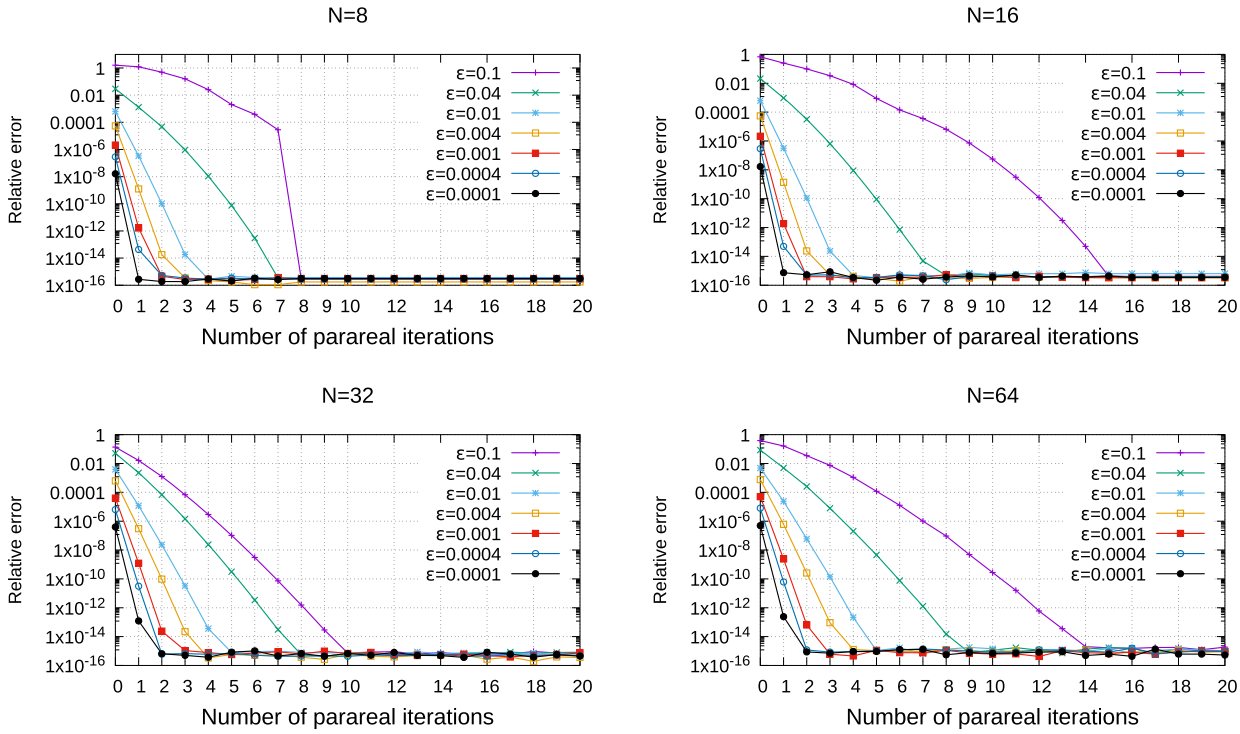

Fig. 10. Convergence of the parareal algorithm in short final times, $T = 500\varepsilon \sim 80P$, for several small values of ε, for the test case in section 4.2.

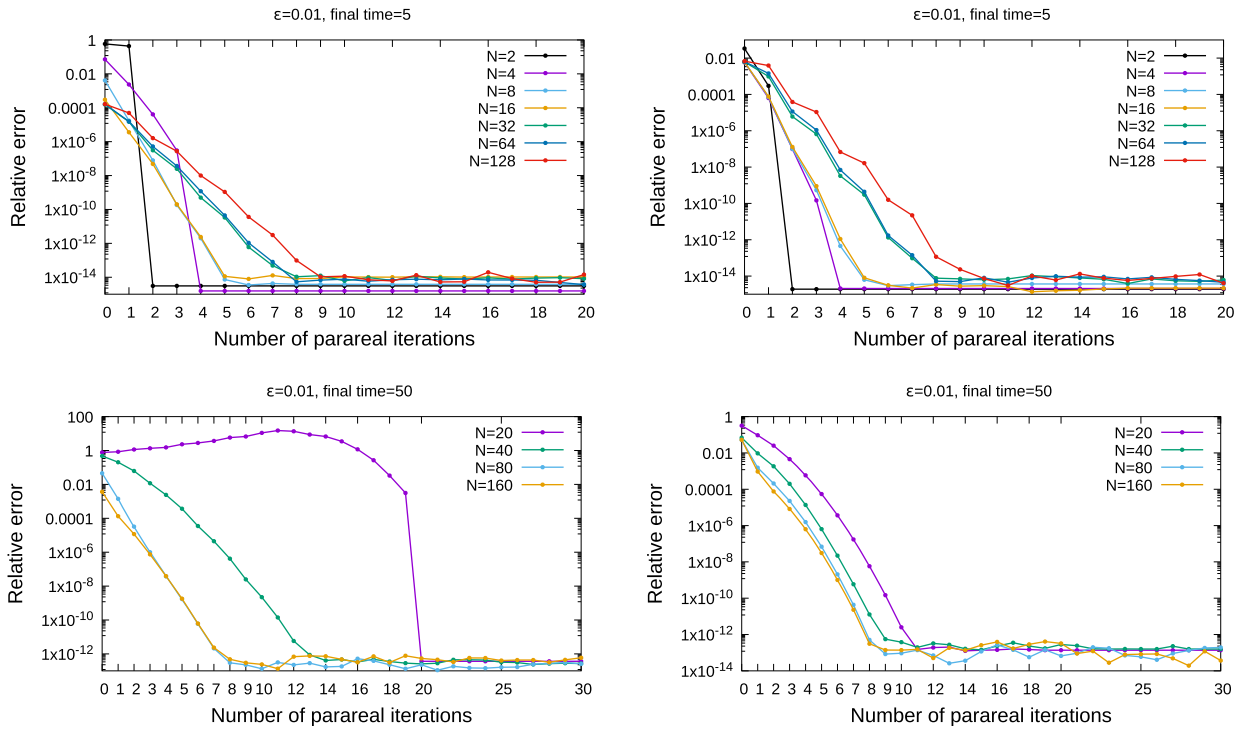

Fig. 11. Convergence of the parareal algorithm for the test case in section 5 in short final time $T = 5 \sim 80P$ (at the top) for the initial conditions in (36) (left panel) and (37) (right panel) and in longer time $T = 50$ (at the bottom).

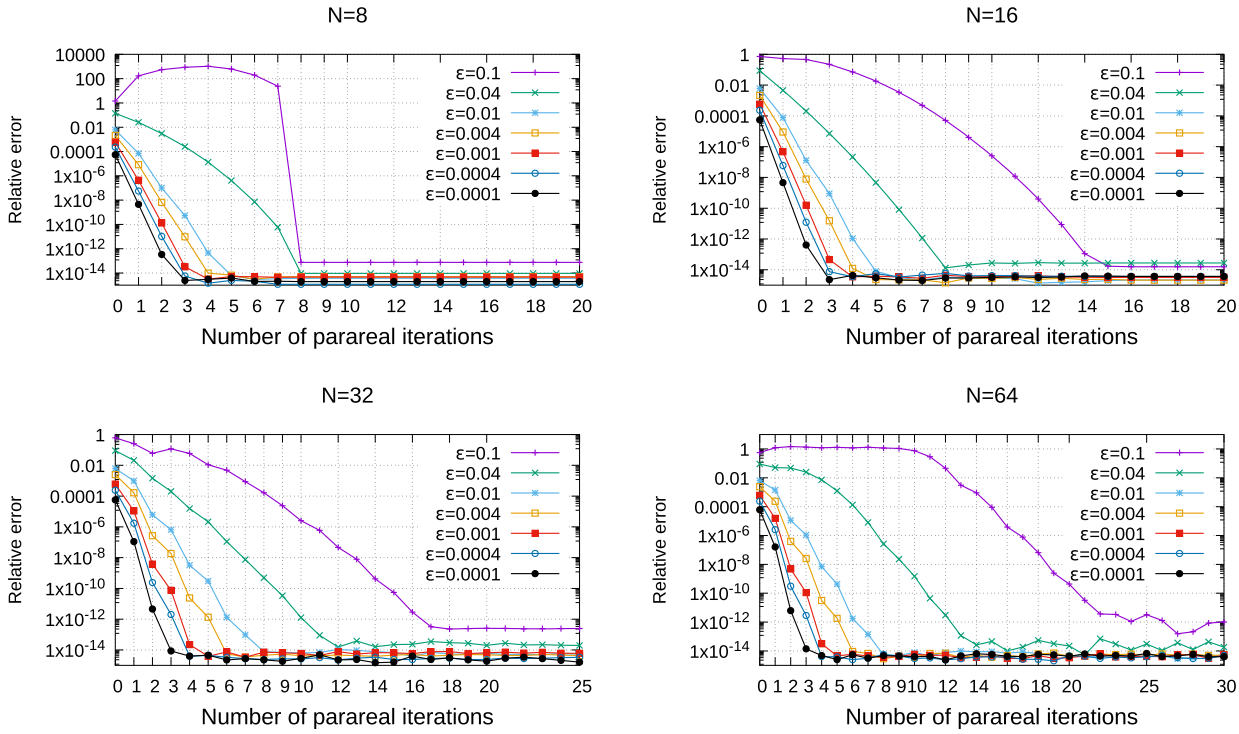

Fig. 12. Convergence of the parareal algorithm in short final times, $T = 500\varepsilon \sim 80P$, for several small values of ε, for the initial condition in (37), for the test case in section 5.

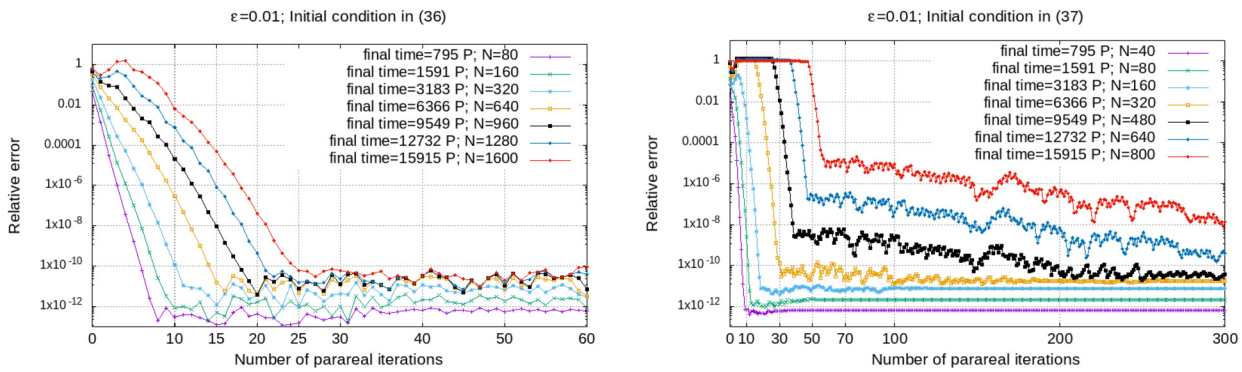

Fig. 13. Convergence of the parareal algorithm for the test case in section 5 when the coarse time step Δt is kept constant and the final time is increasing with N. At the left panel: the initial condition in (36) and $\Delta t = 0.625$. At the right panel: the initial condition in (37) and $\Delta t = 1.25$. The fast cyclotron period is denoted by $P = 2\pi\varepsilon$.

the property of smaller errors with smaller ε, for every k, due to the smaller error of the reduced model with respect to the initial equation. We obtain unsatisfactory results when $\varepsilon = 0.1$, since we recall from the left panel of Fig. 4 that the reduced model is not a good approximation at $T = 50$ for this case. On the contrary, the value of $\varepsilon = 0.04$ leads to satisfactory parareal results, when $N \in \{32, 64\}$. For the other smaller values of the parameter, we observe convergence of the algorithm for all N after a maximum of $k = 9$ iterations. As for the test case in the previous section, we obtain uniform error with respect to ε.

• However, the most interesting problem is that of a long time simulation. We now fix $T = 50$ and we take N in the set $\{20, 40, 80, 160\}$. A bigger value of T can be treated similarly, since the trajectories of both particles evolve as until $T = 50$, with a linear drift in the $\overrightarrow{e_3}$ direction for the particle in (37) (see Fig. 2). The fine time step is set to $\delta t = T/64000$. We plot the relative error at the bottom of Fig. 11. We obtained when N is small much larger errors of the parareal algorithm for the particle in (36) because of its larger perpendicular angular velocity, as mentioned above. Indeed, when N is small, *i.e.* when the time step is big, the error of the limit model is too large so that the parareal method (or the fine solver) can catch a convenient accuracy in a small number of iterations. At the top of Fig. 11, $N = 2$ means a coarse time step of almost 40 rapid oscillations; we have the same remark for $N = 20$ at the bottom of the figure. Nevertheless, we note that in the

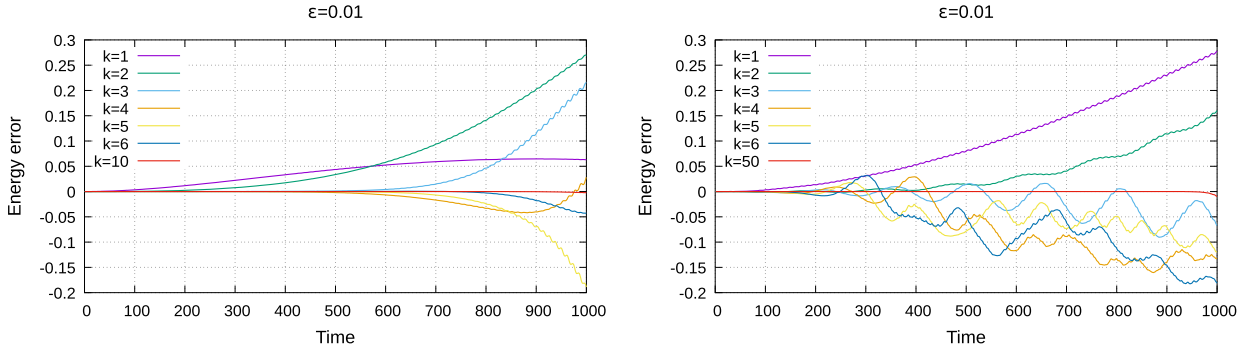

Fig. 14. Evolution of the energy error defined in (40) of the parareal algorithm for the test case in section 5, until the final time $T = 1000 = 15915P$ with $P = 2\pi\varepsilon$. Left panel: the initial condition in (36) with $\Delta t = 0.625$ for the coarse solver. Right panel: the initial condition in (37) with $\Delta t = 1.25$ for the coarse solver.

interesting case of $N = 80$ (resp. $N = 160$), we achieve convergence after only $k = 9$ iterations. This value of N corresponds to a coarse time step of almost 10 (resp. 5) rapid oscillations and to a ratio $\Delta t/\delta t$ of 800 (resp. 400). Thus, coupled to the parallel computations of the fine solver, this strategy could be very effective in terms of computational costs.

• Finally, we perform longer simulations, by keeping the coarse time step fixed while the final time increases with N. The obtained results (see Fig. 13) are not as good as those reported in the previous section, as a consequence of the accuracy of the approximation of the reduced model (compare Fig. 4 to Fig. 3). We fix $\varepsilon = 0.01$, the coarse time step to $\Delta t = 1.25$ and we use the initial condition in (37); smaller values of Δt do not significantly improve the error of the reduced model. As for the particle in (36), we fix $\Delta t = 0.625$, in order to approximate the slow circular motion with a similar accuracy as for the other particle. When considering the initial condition in (37), the results of the parareal algorithm are not fully satisfactory when the final time is large (see Fig. 13): for example, when $T = 1000$, *i.e.* almost 16000 cyclotron periods, we obtain an error of order 10^{-5} after less than $k = 60$ iterations (recall $N = 800$) but afterwards, the error decays very slowly, a quite large number of parareal iterations being necessary to achieve a much smaller error. The situation is completely different when using the initial condition in (36). We obtain good convergence results of the parareal algorithm in large times: at $T = 1000$, for $N = 1600$ time slices, we achieve a 10^{-10} error after $k = 25$ iterations.

To further understand the rationale behind the slow convergence of the algorithm for this test case, we assess the long-term energy error, which is a major issue in applications. First, we verified that the volume-preserving numerical scheme G^4, used as fine solver, preserves the Hamiltonian of the system (32), at the accuracy of the machine precision. More precisely, the Hamiltonian is $\mathcal{H}(\mathbf{x}, \mathbf{v}) = |\mathbf{v}|^2/2$, since there is no electric term. Following [7], we plot in Fig. 14 the error in the energy

$$\mathcal{H}(\mathbf{x}_n^k, \mathbf{v}_n^k) - \mathcal{H}(\mathbf{x}_0, \mathbf{v}_0) \tag{40}$$

where $(\mathbf{x}_0, \mathbf{v}_0)$ is the initial condition and $(\mathbf{x}_n^k, \mathbf{v}_n^k)$ is the k-th iterate of the parareal algorithm. We display, for both initial conditions given by (36) and (37), the energy error corresponding to the first 6 parareal iterations and then, the $k = 10$-th iterate respectively the $k = 50$-th iterate, taking into account when the parareal convergence is achieved (see Fig. 13). We can see that the energy error of the particle in (36) has no large amplitude oscillations in time and the convergence is quite fast, unlike the particle in (37). This is in accordance with the results in Fig. 13.

CRediT authorship contribution statement

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

Laura Grigori, Van-Thanh Nguyen and Julien Salomon thank the funding by the French National Research Agency (ANR) Contract ANR-15-CE23-0019 (project CINE-PARA). The authors wish to express their gratitude to the anonymous referees for their valuable remarks.

References

[1] G. Ariel, S.J. Kim, R. Tsai, Parareal multiscale methods for highly oscillatory dynamical systems, SIAM J. Sci. Comput. 38 (6) (2016) A3540–A3564.
[2] G. Bal, On the convergence and the stability of the parareal algorithm to solve partial differential equations, in: T.J. Barth, M. Griebel, D.E. Keyes, R.M. Nieminen, D. Roose, T. Schlick, R. Kornhuber, R. Hoppe, J. Périaux, O. Pironneau, O. Widlund, J. Xu (Eds.), Domain Decomposition Methods in Science and Engineering, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005, pp. 425–432.
[3] P.F. Fischer, F. Hecht, Y. Maday, A parareal in time semi-implicit approximation of the Navier-Stokes equations, in: Domain Decomposition Methods in Science and Engineering, Springer, 2005, pp. 433–440.
[4] E. Frénod, Application of the averaging method to the gyrokinetic plasma, Asymptot. Anal. 46 (2006) 1–28.
[5] E. Frénod, F. Watbled, The Vlasov equation with strong magnetic field and oscillating electric field as a model for isotope resonant separation, Electron. J. Differ. Equ. 2002 (6) (2002) 1–20.
[6] M.J. Gander, E. Hairer, Nonlinear convergence analysis for the parareal algorithm, in: U. Langer, M. Discacciati, D.E. Keyes, O.B. Widlund, W. Zulehner (Eds.), Domain Decomposition Methods in Science and Engineering XVII, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 45–56.
[7] M.J. Gander, E. Hairer, Analysis for parareal algorithms applied to Hamiltonian differential equations, J. Comput. Appl. Math. 259 (2014) 2–13.
[8] M.J. Gander, S. Vandewalle, Analysis of the parareal time-parallel time-integration method, SIAM J. Sci. Comput. 29 (2) (2007) 556–578.
[9] T. Haut, B. Wingate, An asymptotic parallel-in-time method for highly oscillatory PDEs, SIAM J. Sci. Comput. 36 (2) (2014) A693–A713.
[10] R.D. Hazeltine, J.D. Meiss, Plasma Confinement, Courier Corporation, 2003.
[11] Y. He, Y. Sun, J. Liu, H. Qin, Volume-preserving algorithms for charged particle dynamics, J. Comput. Phys. 281 (2015) 135–147.
[12] M. Kretzschmar, Particle motion in a Penning trap, Eur. J. Phys. 12 (5) (1991) 240.
[13] F. Legoll, T. Lelievre, G. Samaey, A micro-macro parareal algorithm: application to singularly perturbed ordinary differential equations, SIAM J. Sci. Comput. 35 (4) (2013) A1951–A1986.
[14] J.-L. Lions, Y. Maday, G. Turinici, A “parareal” in time discretization of PDE's, C. R. Acad. Sci., Ser. 1 Math. 332 (2001) 661–668.
[15] Y. Maday, Parareal in time algorithm for kinetic systems based on model reduction, in: High-Dimensional Partial Differential Equations in Science and Engineering, in: CRM Proc. Lecture Notes, vol. 41, 2007, pp. 183–194.
[16] G.A. Staff, E.M. Rønquist, Stability of the parareal algorithm, in: T.J. Barth, M. Griebel, D.E. Keyes, R.M. Nieminen, D. Roose, T. Schlick, R. Kornhuber, R. Hoppe, J. Périaux, O. Pironneau, O. Widlund, J. Xu (Eds.), Domain Decomposition Methods in Science and Engineering, Springer Berlin Heidelberg, Berlin, Heidelberg, 2005, pp. 449–456.