CONVERGENCE ANALYSIS OF ASYMPTOTIC PRESERVING SCHEMES FOR STRONGLY MAGNETIZED PLASMAS

FRANCIS FILBET, L. MIGUEL RODRIGUES, AND HAMED ZAKERZADEH

Abstract. The present paper is devoted to the convergence analysis of a class of asymptotic preserving particle schemes *[Filbet & Rodrigues, SIAM J. Numer. Anal., 54(2) (2016):1120–1146]* for the Vlasov equation with a strong external magnetic field. In this regime, classical Particle-In-Cell (PIC) methods are subject to quite restrictive stability constraints on the time and space steps, due to the small Larmor radius and plasma frequency. The asymptotic preserving discretization that we are going to study removes such a constraint while capturing the large-scale dynamics, even when the discretization (in time and space) is too coarse to capture fastest scales. Our error bounds are explicit regarding the discretization, stiffness parameter, initial data and time.

Keywords: Asymptotic preserving schemes; Vlasov–Poisson system; High-order time discretization; Homogeneous magnetic field; Particle methods

2010 MSC: 35Q83, 65M75, 82D10, 65L04, 65M15.

Contents

1. Introduction 1
2. Oscillatory limit of the continuous model 5
3. First-order scheme on the original spatial variable 9
4. First-order scheme on the guiding center variable 17
5. L-stable second-order implicit-explicit scheme 20
6. Numerical experiments 28
7. Conclusion and perspectives 29
Appendix A. Convergence analysis for oscillatory ODEs 29
References 32

1. Introduction

1.1. **Strongly magnetized plasmas.** Magnetized plasmas are encountered in a wide variety of astrophysical situations, but also in magnetic fusion devices such as tokamaks, where a large external magnetic field needs to be applied in order to keep the plasma particles on desired tracks. In numerical simulations of such devices, this large external magnetic field should be taken into account for pushing the particles, in particle methods [1]. However, due to the magnitude of the concerned field, this often adds a new time scale to the simulation, thus possibly a stringent restriction on the time step. In order to handle this additional timescale, one may wish to use numerical schemes whose stability is independent of the restrictive time step, and that compute approximate solutions

Date: January 19, 2021.
FF is supported by the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014–2018 under the grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

Research of LMR has received funding from the city of Rennes.

HZ was supported by LabEx CIMI Toulouse and, partially, by Institut Universitaire de France.

retaining the large-scale behavior implied by the external field, even when time steps are too coarse to capture fast oscillations.

To get some first intuition, one can consider the simplest possible situation and follow the trajectory of a single particle in a constant magnetic field $\mathbf{B}$ subject to no electric field. This trajectory turns out to be a helicoid along the magnetic field lines with a radius proportional to the inverse of the magnitude of $\mathbf{B}$. Hence, when this field becomes very large, the particle gets trapped along the magnetic field lines. A slightly more precise description of the dynamics is that particles spin around a point on the magnetic field line, the *"guiding center"*, whose velocity is smaller than the particle velocities. When electric field effects are taken into account and the magnetic field is not constant, the situation is more complicated, but still, the apparent particle velocity is smaller than the actual one and in many situations the link between the real and the apparent velocity is well-known in terms of electromagnetic fields $\mathbf{B}$ and $\mathbf{E}$; see [3, 11, 14, 29, 32, 35] for instance.

The behavior of a plasma, constituted of a large number of charged particles, is even more complicated and may be described by the Vlasov equation coupled with the Maxwell or Poisson equations to compute the self-consistent fields. The Vlasov equation models, in essence, the evolution of a system of charged particles under the effects of external and self-consistent fields by describing the time-evolution of the unknown $f(t, \boldsymbol{x}, \boldsymbol{v})$, depending on the time t, the position $\boldsymbol{x}$, and the velocity $\boldsymbol{v}$, which represents the distribution of particles in the phase space for each species with $(\boldsymbol{x}, \boldsymbol{v}) \in \mathbb{R}^{d_{\boldsymbol{x}}} \times \mathbb{R}^{d_{\boldsymbol{v}}}$ $(d_{\boldsymbol{x}}, d_{\boldsymbol{v}} = 1, 2, 3)$, as

$$\begin{cases} \dfrac{\partial f}{\partial t} \,+\, \mathrm{div}_{\boldsymbol{x}}(\boldsymbol{v} f) \,+\, \mathrm{div}_{\boldsymbol{v}}(\mathbf{F} f) \,=\, 0\,, \\ \\ f(0, \cdot, \cdot) = f_0\,, \end{cases} \tag{1.1}$$

where the force field $\mathbf{F}(t, \boldsymbol{x}, \boldsymbol{v})$, which is coupled with the distribution function f, makes the equation nonlinear. For instance, for the single-species Vlasov–Poisson model, the force field stems from the electric field $\mathbf{E}(t, \boldsymbol{x})$, *i.e.*, it reads

$$\mathbf{F}(t, \boldsymbol{x}, \boldsymbol{v}) = \frac{q}{m}\, \mathbf{E}(t, \boldsymbol{x})\,, \qquad \mathbf{E}(t, \boldsymbol{x}) = -\nabla_{\boldsymbol{x}} \phi(t, \boldsymbol{x})\,, \qquad -\Delta_{\boldsymbol{x}} \phi = \frac{\rho}{\epsilon_0}\,, \tag{1.2}$$

where (m, q) are the elementary mass and charge (of one particle), $\phi(t, \boldsymbol{x})$ is the electric potential, ϵ_0 is the electric constant, and the charge density $\rho(t, \boldsymbol{x})$ is given in terms of the distribution function as

$$\rho(t, \boldsymbol{x}) := q \int_{\mathbb{R}^{d_{\boldsymbol{v}}}} f(t, \boldsymbol{x}, \boldsymbol{v})\, \mathrm{d}\boldsymbol{v}\,.$$

When, in addition, we take into account a magnetic field $\mathbf{B}(t, \boldsymbol{x})$, the Lorentz force applies, *i.e.*,

$$\mathbf{F}(t, \boldsymbol{x}, \boldsymbol{v}) = \frac{q}{m} \big(\mathbf{E}(t, \boldsymbol{x}) + \boldsymbol{v} \times \mathbf{B}(t, \boldsymbol{x}) \big)\,.$$

Although the framework of our investigation can be adapted easily to the multi-species case, we shall only consider the single-species case, for the sake of simplicity, and scale parameters to $(m, q) \equiv (1, 1)$.

As we are principally interested in the analysis of numerical schemes, we shall limit our focus to the quite simple case, with an external uniform magnetic field (in direction and magnitude), so to illustrate the bottom-line of the analysis with more ease. Note that, making such an assumption, we deprive ourselves of investigating phenomena such as curvature effects while we will be able to decouple dynamics in parallel and normal directions (with respect to the magnetic field), hence, we may concentrate on the particle motion in the perpendicular plane. More explicitly, in the present paper, we set

$$\mathbf{B}(t, \boldsymbol{x}) \;=\; \frac{1}{\varepsilon} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\,,$$

and follow only the evolution of the first two components of the Cartesian space, $\boldsymbol{x} = (x_1, x_2) \in \mathbb{R}^2$. The parameter ε is related to the ratio between the reciprocal Larmor frequency and the advection timescale; see [11, 23] and references therein for more details on such a scaling. We are particularly interested in the regime where $0 < \varepsilon \ll 1$ as it implies that the magnetic field is very strong, which is required to confine the plasma, practically speaking.

Under these assumptions, the *"long-time coherent behavior"* arising in plasmas submitted to a strong external and uniform magnetic field will be obtained by the following Vlasov equation:

$$\begin{cases} \varepsilon \dfrac{\partial f^\varepsilon}{\partial t} + \operatorname{div}_{\boldsymbol{x}}(\boldsymbol{v} f^\varepsilon) + \operatorname{div}_{\boldsymbol{v}} \left((\mathbf{E} - \dfrac{1}{\varepsilon} \boldsymbol{v}^\perp) f^\varepsilon \right) = 0 \,, \\ \\ f^\varepsilon(0, \cdot, \cdot) = f_0 \,, \end{cases} \tag{1.3}$$

where the *orthogonal velocity* $\boldsymbol{v}^\perp := (-v_2, v_1)$ can be seen as the rotation of the original velocity with the rotation matrix $\mathbf{J}$:

$$\boldsymbol{v}^\perp = \mathbf{J}\, \boldsymbol{v} \,, \qquad \mathbf{J} := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} . \tag{1.4}$$

In (1.3) the presence of an epsilon in front of the time derivative encodes that the time scale has been chosen according to the *"slow"* part of the dynamics. On a shorter time scale slow variables would be asymptotically steady at leading-order. Carryint out our analysis with these rescaled times is consistent with the fact that our numerical schemes are designed to capture slow dynamics even with time steps that are large compared to fast oscillation periods.

At the continuous level, considerable efforts have been made on the rigorous derivation of reduced models from kinetic transport equations like (1.3), *i.e.*, in the so-called *oscillatory limit* $\varepsilon \to 0$ in which, there are very fast temporal oscillations in the plasma, of time scale $\mathcal{O}(\varepsilon^2)$, in the orthogonal direction to the magnetic field; see [22, 31, 24, 25, 14] for relatively recent panoramas on the question. In fact, one can obtain this limit system either using the formal Hilbert or Poincaré expansion (as in [12, 13]) or by a rigorous approach (only when the magnetic field is homogeneous in space), *cf.* [19, 20, 34] or more recently using the characteristic curves [33]. Nonetheless, the most remarkable mathematical result is restricted to the two-dimensional setting with a constant magnetic field and with interactions described through the Poisson equation, and yet validates only half of the slow dynamics; see [34], which is built on [20] and recently revisited in [33]. In fact, the reduced limit system for the weak limit $f^\varepsilon \rightharpoonup f$ writes [19]

$$\frac{\partial f}{\partial t} - \mathbf{E}^\perp \cdot \nabla_{\boldsymbol{x}} f - \frac{1}{2} \Delta \phi\, \boldsymbol{v}^\perp \cdot \nabla_{\boldsymbol{v}} f = 0 \,, \qquad (\boldsymbol{x}, \boldsymbol{v}) \in \mathbb{R}^4 \,, \tag{1.5}$$

or equivalently when ϕ satisfies the Poisson equation (1.2),

$$\frac{\partial f}{\partial t} - \mathbf{E}^\perp \cdot \nabla_{\boldsymbol{x}} f + \frac{\rho}{2\,\epsilon_0}\, \boldsymbol{v}^\perp \cdot \nabla_{\boldsymbol{v}} f = 0 \,, \qquad (\boldsymbol{x}, \boldsymbol{v}) \in \mathbb{R}^4 \,,$$

while the limit of charge density converges to a solution of the following [34]:

$$\frac{\partial \rho}{\partial t} - \nabla_{\boldsymbol{x}} \cdot \left(\rho\, \mathbf{E}^\perp \right) = 0 \,, \qquad \boldsymbol{x} \in \mathbb{R}^2 \,. \tag{1.6}$$

Concerning the three dimensional case, we refer to [14] for a recent study where the electric field is either an external applied field or arises from a smoothed Poisson equation.

From the discrete point of view, we are seeking methods which are able to capture this singularly oscillatory limit, so that the numerical method provides a consistent discretization of the limit system as $\varepsilon \to 0$, a concept known as *asymptotic consistency*, with the numerical parameters to be independent of the singular scaling parameter ε. This concept has been widely studied for dissipative systems, since the pioneering works of [26, 28], in the framework of *asymptotic preserving (AP) schemes*; see also the review paper [27]. In the design of well-adapted numerical schemes to capture the slow part of the dynamics with a rather coarse discretization (compared to the fast scales), one could mention the two-scale convergence method [18, 19], the micro-macro decomposition in [6],

lifting with multiple time variables [8, 9, 10, 4], exponential integrators [16, 17], frequency filtering [21], and implicit-explicit time discretizations [12, 13, 36]. The reader is also referred to [5, 7] for some numerical comparisons, including comparisons with more standard methods.

In the present work, we investigate the strategy of a series of work [12, 13, 15], which live in the context of particle methods [30], and hinge upon investigating the characteristics of the system, instead of using directly the PDE. Our goal, indeed, is to provide a complete convergence analysis of the Particle-In-Cell (PIC) methods introduced in [12], solving for the following system of characteristics:

$$\begin{cases} \varepsilon\, \boldsymbol{x}_\varepsilon'(t) = \boldsymbol{v}_\varepsilon(t)\,, \\ \varepsilon\, \boldsymbol{v}_\varepsilon'(t) = \mathbf{E}(t, \boldsymbol{x}_\varepsilon(t)) \,-\, \dfrac{\boldsymbol{v}_\varepsilon^\perp(t)}{\varepsilon}\,, \\ \boldsymbol{x}_\varepsilon(s) = \boldsymbol{x}_\varepsilon^s, \qquad \boldsymbol{v}_\varepsilon(s) = \boldsymbol{v}_\varepsilon^s\,, \end{cases} \tag{1.7}$$

for $t \geq 0$ and for any regime of the scaling parameter ε. So far, and in [12], some well-adapted schemes have been designed and analyzed in the regime where ε^2 is much smaller than the time step of the numerical scheme. Here, we shall perform a complete convergence and asymptotic error analysis for any values of the asymptotic parameter ε and of the time step, denoted Δt in the sequel.

In order to performing such an analysis, we start with estimates for the continuous model in §2, followed by the discrete estimates for two versions of first-order numerical schemes in §3 and §4. Then, we establish the convergence analysis for a second-order L-stable method in §5, and we provide some numerical illustrations in §6.

To carry out a complete and rigorous analysis, details of the system and the schemes obviously come into play but some enlightening insights on the final outcomes may be obtained by considering an abstract system. We conclude by providing the reader with such abstract considerations in Appendix A.

Notational convention 1.1. *Hereinafter, and for the sake of brevity, we use*

- $\lesssim A$ *to denote* $\leq c_0 A$ *for some universal constant* c_0*, and*
- $\underset{\alpha}{\lesssim} A$ *to denote* $\leq c_0(\alpha) A$ *for some constant* c_0 *depending on* α*. introduced in* [12]*, for the two-dimensional system with a homogeneous external magnetic field.*

Notational convention 1.2. *Our estimates shall be expressed in terms of*

$$K_0 \;:=\; \|\mathbf{E}\|_{L^\infty}\,, \quad K_t \;:=\; \|\partial_t \mathbf{E}\|_{L^\infty}\,, \quad K_x \;:=\; \|\mathrm{d}_{\boldsymbol{x}} \mathbf{E}\|_{L^\infty}\,, \quad K_{xx} \;:=\; \|\mathrm{d}^2_{\boldsymbol{x}} \mathbf{E}\|_{L^\infty}\,, \quad \cdots\,,$$

where $\mathrm{d}_{\boldsymbol{x}}$ *denotes the total differential operator in the* $\boldsymbol{x}$ *variable*

$$\mathrm{d}_{\boldsymbol{x}} \mathbf{F}(t, \boldsymbol{x}_0)\boldsymbol{z} := \lim_{h \to 0} \frac{\mathbf{F}(t, \boldsymbol{x}_0 + h\,\boldsymbol{z}) - \mathbf{F}(t, \boldsymbol{x}_0)}{h}\,.$$

In particular, we assume global bounds on the electric field and its derivatives. We expect that a counterpart could be obtained when fields are only locally bounded but the initial density is compactly supported. We omit, however, to state and prove this variant of our results as required adaptations are expected to be quite technical but rather classical.

Remark 1.3. *Part of our motivation to make explicit the dependence on* $\mathbf{E}$ *in our estimates comes from the desire to reduce the gap in extrapolating our results to a more nonlinear context where field equations couple fields to densities. In this respect, it is crucial to note that each time derivative of* $\mathbf{E}$ *leads to an extra* ε*-factor, since in a nonlinear context one expects to prove uniform* L^∞*-bounds only on* $(\varepsilon\partial_t)^\alpha \mathrm{d}^\beta_{\boldsymbol{x}} \mathbf{E}$ *and not on* $\partial_t^\alpha \mathrm{d}^\beta_{\boldsymbol{x}} \mathbf{E}$ *itself.*

2. Oscillatory limit of the continuous model

In this section, we consider the characteristic system (1.7) of the Vlasov equation (1.3) in the oscillatory limit, which corresponds to the limit system (1.6). It is worth mentioning that our presentation of the asymptotic analysis at the continuous level is close to [14, §3], though with a different scaling for the time. Despite this similarity, we prefer to expound this continuous analysis mainly because, later on, and at the discrete level, we have to go along similar lines for the numerical method.

We aim to compare the characteristic system (1.7) with the limit system as $\varepsilon \to 0$, that is, the *guiding-center approximation*, obtained as the solution of the following equation:

$$\begin{cases} \boldsymbol{x}'(t) \,=\, -\mathbf{E}^\perp(t, \boldsymbol{x}(t)), \qquad \forall t \geq 0, \\ \\ \boldsymbol{x}(s) = \boldsymbol{x}^s \,. \end{cases} \tag{2.1}$$

In particular, in this section, we prove $\boldsymbol{x} = \lim_{\varepsilon \to 0} \boldsymbol{x}_\varepsilon$ when $(\boldsymbol{x}_\varepsilon, \boldsymbol{v}_\varepsilon)$ solves (1.7) and $\boldsymbol{x}$ solves (2.1), and we quantify the corresponding error estimate.

To work with a quantity that is slower than $\boldsymbol{x}_\varepsilon$, we introduce $\boldsymbol{y}_\varepsilon$ defined as

$$\boldsymbol{y}_\varepsilon(t) := \boldsymbol{x}_\varepsilon(t) - \varepsilon\, \boldsymbol{v}_\varepsilon^\perp(t)\,, \tag{2.2}$$

where the couple $(\boldsymbol{x}_\varepsilon, \boldsymbol{v}_\varepsilon)$ is the solution to the characteristic curves (1.7) of the kinetic model (1.3). Note that, hereinafter and for later use, we denote by

$$(\mathbf{X}_\varepsilon(t, s, \boldsymbol{x}^s_\varepsilon, \boldsymbol{v}^s_\varepsilon), \mathbf{V}_\varepsilon(t, s, \boldsymbol{x}^s_\varepsilon, \boldsymbol{v}^s_\varepsilon)) := (\boldsymbol{x}_\varepsilon(t), \boldsymbol{v}_\varepsilon(t))$$

the value of the solution to (1.7), at any time t, and accordingly $\mathbf{Y}_\varepsilon(t, s, \boldsymbol{x}^s_\varepsilon, \boldsymbol{v}^s_\varepsilon) := \boldsymbol{y}_\varepsilon(t)$. Likewise, when $\boldsymbol{x}$ solves (2.1), we denote $\mathbf{X}(t, s, \boldsymbol{x}^s) := \boldsymbol{x}(t)$.

As a preliminary remark, we note that solutions to (1.7) are global in time as soon as $\mathbf{E} \in L^\infty$. Also, we fix classical notation $W^{k,p}$ for the Sobolev space with derivatives up to order k, measured in the L^p norm. We shall use, hereinafter, canonical Euclidean ℓ_2-norms on vectors and corresponding operator norms on linear operators and matrices. Then, we have the following result:

Theorem 2.1. *(i) Assume that $\mathbf{E} \in W^{1,\infty}$. Then, for any $\varepsilon > 0$, for the difference between flows of (1.7) and (2.1), we have*

$$\|\mathbf{X}_\varepsilon(t, 0, \boldsymbol{x}^0_\varepsilon, \boldsymbol{v}^0_\varepsilon) - \mathbf{X}(t, 0, \boldsymbol{x}^0_\varepsilon)\| \underset{\mathbf{E}}{\lesssim} \varepsilon\, e^{K_x\, t}(1 + t^2)\, (\|\boldsymbol{v}^0_\varepsilon\| + \varepsilon)\,.$$

(ii) Assume that $\mathbf{E} \in W^{2,\infty}$. Then, for any $\varepsilon > 0$, we have

$$\|\mathbf{Y}_\varepsilon(t, 0, \boldsymbol{x}^0_\varepsilon, \boldsymbol{v}^0_\varepsilon) - \mathbf{X}(t, 0, \boldsymbol{x}^0_\varepsilon - \varepsilon(\boldsymbol{v}^0_\varepsilon)^\perp)\| \underset{\mathbf{E}}{\lesssim} \varepsilon^2\, (1 + t^4)\, e^{2\, K_x\, t}\, \left(1 + \varepsilon^2 + \|\boldsymbol{v}^0_\varepsilon\|^2\right)\,.$$

Remark 2.2. *The reader may rightfully remark that the foregoing estimates do not scale sharply when $t \to 0$ or $t \to \infty$. For instance, as the left-hand side vanishes at time $t = 0$, one could expect that the right-hand vanishes as well. Indeed, one may simply resolve such an issue by changing the $\mathcal{O}(\varepsilon)$ bound (for the first case) into $\mathcal{O}(\min(\varepsilon, t/\varepsilon))$, by taking into account direct bounds on time derivatives. However, we have chosen to disregard this refinement as this provides an improvement only when $t = \mathcal{O}(\varepsilon^2)$. Also, in the reverse direction $t \to \infty$, we have chosen not to optimize constants or power of times. Typically, for the sake of simplicity, we have chosen to use bounds such as*

$$e^{K_x t} \int_0^t e^{-K_x s} \mathrm{d}s \,=\, \frac{e^{K_x t} - 1}{K_x} \,\leq\, t\, e^{K_x\, t}\,.$$

Before detailing the proof of Theorem 2.1, we would like to highlight that, in essence, there are two steps in the proof :

(i) The first step is to prove the boundedness of the solutions of the characteristic system (1.7) with respect to ε, sometimes referred to as ε-boundedness below. We will discuss this kind of ε-uniform estimates in §2.1. Note that rough direct bounds would predict blow up in terms of ε, not because of the skew-symmetric ε^{-2} term but due to existing ε^{-1} terms.

(ii) In the second step, one derives, from system (1.7), that the function to be compared with, either $\boldsymbol{x}_\varepsilon$ or $\boldsymbol{y}_\varepsilon$, satisfies an equation, which is asymptotically close to the expected limiting equation (2.1), the guiding center equation. This step is algebraic in nature and is carried out in §2.2. It builds upon the fact that the velocity equation in the characteristic system (1.7) yields that, formally speaking, $\boldsymbol{v}_\varepsilon$ is the sum of a quantity of size ε and the time derivative of size ε^2. Note that the first step precisely ensures that this formal reasoning is valid.

The conclusion is then obtained, also in §2.2, by combining the two steps with a stability estimate on the expected limiting equation.

2.1. **Uniform estimates on characteristics.** In order to establish uniform estimates with respect to ε, we first define a new variable called $\boldsymbol{z}_\varepsilon$, as follows [1]

$$\boldsymbol{z}_\varepsilon(t) \,:=\, \boldsymbol{v}_\varepsilon(t) \,+\, \varepsilon\, \mathbf{E}^\perp(t, \boldsymbol{x}_\varepsilon(t)), \quad t \geq 0\,, \tag{2.3}$$

whose temporal dynamics is more purely oscillatory than the one of $\boldsymbol{v}_\varepsilon$, in the sense that the influence of non-oscillatory terms is less significant, namely here $\boldsymbol{z}_\varepsilon' + \varepsilon^{-2}\boldsymbol{z}_\varepsilon^\perp = \mathcal{O}(1)$ whereas $\boldsymbol{v}_\varepsilon' + \varepsilon^{-2}\boldsymbol{v}_\varepsilon^\perp = \mathcal{O}(\varepsilon^{-1})$. This can be seen explicitly, since the time evolution of $\boldsymbol{z}_\varepsilon$ (*cf.* [12, eq. (5)]) obeys

$$\begin{aligned} \boldsymbol{z}_\varepsilon'(t) \,&=\, -\frac{1}{\varepsilon^2}\boldsymbol{z}_\varepsilon^\perp(t) \,+\, \varepsilon\, \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{E}^\perp(t, \boldsymbol{x}_\varepsilon(t))\,, \\ &=\, -\frac{1}{\varepsilon^2}\, \boldsymbol{z}_\varepsilon^\perp(t) \,+\, \varepsilon\, \partial_t \mathbf{E}^\perp(t, \boldsymbol{x}_\varepsilon(t)) \,+\, \mathrm{d}_{\boldsymbol{x}}\mathbf{E}^\perp(t, \boldsymbol{x}_\varepsilon(t)) \Big(\boldsymbol{z}_\varepsilon(t) \,-\, \varepsilon\, \mathbf{E}^\perp(t, \boldsymbol{x}_\varepsilon(t))\Big)\,, \end{aligned} \tag{2.4}$$

by using $\varepsilon\, \boldsymbol{x}_\varepsilon'(t) = \boldsymbol{v}_\varepsilon(t)$ from (1.7). We should emphasize that the motivation for introducing $\boldsymbol{z}_\varepsilon$ and looking for a dynamics as purely oscillatory as possible is that the oscillatory part of the evolution preserves the Euclidean norm, hence one obtains readily a good estimate by using $\boldsymbol{z}_\varepsilon$, as in the following lemma.

Lemma 2.3. *Let us assume that $\mathbf{E} \in W^{1,\infty}$ and consider the system corresponding to the characteristic curves of the Vlasov equation (1.7). Then, the auxiliary variable $\boldsymbol{z}_\varepsilon$ introduced in (2.4) and the velocity $\boldsymbol{v}_\varepsilon$ are bounded as*

$$\begin{cases} \|\boldsymbol{z}_\varepsilon(t)\| \,\leq\, e^{K_x\, t}\big(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\, K_0\big) + \varepsilon\, t\, e^{K_x\, t}\, (K_t \,+\, K_x\, K_0)\,, \\ \|\boldsymbol{v}_\varepsilon(t)\| \,\leq\, \|\boldsymbol{z}_\varepsilon(t)\| \,+\, K_0\, \varepsilon\,. \end{cases}$$

Proof. By taking the scalar product of (2.4) with $\boldsymbol{z}_\varepsilon(t)$, which cancels out the singular term $-\boldsymbol{z}_\varepsilon^\perp(t)/\varepsilon^2$, we obtain

$$\begin{aligned} \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\, \|\boldsymbol{z}_\varepsilon(t)\|^2 \,&=\, \varepsilon\, \boldsymbol{z}_\varepsilon(t) \cdot \partial_t \mathbf{E}^\perp(t, \boldsymbol{x}_\varepsilon(t)) \\ &\qquad +\, \boldsymbol{z}_\varepsilon(t) \cdot \mathrm{d}_{\boldsymbol{x}}\mathbf{E}^\perp(t, \boldsymbol{x}_\varepsilon(t))\, \boldsymbol{z}_\varepsilon(t) \,-\, \varepsilon\, \boldsymbol{z}_\varepsilon(t) \cdot (\mathrm{d}_{\boldsymbol{x}}\mathbf{E}^\perp \mathbf{E}^\perp)(t, \boldsymbol{x}_\varepsilon(t)) \\ &\leq \varepsilon\, K_t\, \|\boldsymbol{z}_\varepsilon(t)\| \,+\, K_x\, \|\boldsymbol{z}_\varepsilon(t)\|^2 \,+\, \varepsilon\, K_x\, K_0\, \|\boldsymbol{z}_\varepsilon(t)\|\,. \end{aligned}$$

Denoting by t_0 the supremum of times in $[0, t]$ where $\boldsymbol{z}_\varepsilon$ vanishes[2], one may simplify by $\|\boldsymbol{z}_\varepsilon\|$ on (t_0, t) to derive for any $s \in (t_0, t)$

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{z}_\varepsilon\|(s) \leq\, K_x\, \|\boldsymbol{z}_\varepsilon(s)\| + \varepsilon\, (K_t \,+\, K_x\, K_0)\,,$$

which, by integration, it yields

$$\begin{aligned} \|\boldsymbol{z}_\varepsilon(t)\| \,&\leq\, e^{K_x(t-t_0)}\|\boldsymbol{z}_\varepsilon(t_0)\| + \varepsilon\, (t - t_0)\, e^{K_x\,(t-t_0)}\, (K_t \,+\, K_x\, K_0)\,, \\ &\leq\, e^{K_x t}\|\boldsymbol{z}_\varepsilon(0)\| + \varepsilon\, t\, e^{K_x\, t}\, (K_t \,+\, K_x\, K_0)\,. \end{aligned}$$

The second estimate, on $\boldsymbol{v}_\varepsilon(t)$, follows readily from $\boldsymbol{v}_\varepsilon(t) = \boldsymbol{z}_\varepsilon(t) \,-\, \varepsilon \mathbf{E}^\perp(t, \boldsymbol{x}_\varepsilon(t))$. □

[1] Note that the definition of $\boldsymbol{z}_\varepsilon$ differs from the one in [12] only by a scaling of ε.

[2] By definition $t_0 = 0$ if $\boldsymbol{z}_\varepsilon$ does not vanish.

So, one concludes that provided that the electric field is regular enough, *e.g.*, $\mathbf{E} \in W^{1,\infty}$, the norms $\|\boldsymbol{z}_\varepsilon\|$ and $\|\boldsymbol{v}_\varepsilon\|$ are bounded locally in time, uniformly with respect to ε.

2.2. **Proof of Theorem 2.1.** Now, we derive from the second equation of (1.7)

$$\varepsilon\,(\boldsymbol{v}_\varepsilon^\perp)'(t) \,-\, \mathbf{E}^\perp(t,\boldsymbol{x}_\varepsilon(t)) \,=\, \frac{\boldsymbol{v}_\varepsilon(t)}{\varepsilon}\,,$$

which, combined with the first equation of (1.7), yields

$$(\boldsymbol{x}_\varepsilon - \varepsilon \boldsymbol{v}_\varepsilon^\perp)'(t) \,=\, -\mathbf{E}^\perp(t,\boldsymbol{x}_\varepsilon(t))\,. \tag{2.5}$$

This shows that $\boldsymbol{x}_\varepsilon$ satisfies an equation seemingly close to the guiding center equation (2.1).

Then, in order to prove the first estimate of Theorem 2.1, we subtract (2.5) from the limit system in (2.1), integrate over $[0,t]$, and use the Lipschitz bound on $\mathbf{E}$ as well as Lemma 2.3, to obtain

$$\begin{aligned}\|\boldsymbol{x}_\varepsilon(t) - \mathbf{X}(t,0,\boldsymbol{x}_\varepsilon^0)\| &\le \varepsilon(\|\boldsymbol{v}_\varepsilon^0\| + \|\boldsymbol{v}_\varepsilon(t)\|) + K_x \int_0^t \|\boldsymbol{x}_\varepsilon(s) - \mathbf{X}(s,0,\boldsymbol{x}_\varepsilon^0)\|\,\mathrm{d}s \\ &\le K_x \int_0^t \|\boldsymbol{x}_\varepsilon(s) - \mathbf{X}(s,0,\boldsymbol{x}_\varepsilon^0)\|\,\mathrm{d}s \\ &\quad + \varepsilon\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,\big(K_0\varepsilon + e^{K_x\,t}(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,K_0) + \varepsilon\,t\,e^{K_x t}\,(K_t \,+\, K_x\,K_0)\big)\,.\end{aligned} \tag{2.6}$$

At this stage, we are ready to apply *Grönwall's lemma* (in the integral form): Let A, a, and K be non-negative constants such that

$$A(t) \le a(t) + K\int_0^t A(s)\,\mathrm{d}s, \qquad \forall t \ge 0. \tag{2.7a}$$

Then, it holds

$$A(t) \le a(t) + K\int_0^t e^{K(t-s)}a(s)\,\mathrm{d}s\,, \qquad \forall t \ge 0. \tag{2.7b}$$

For applying this lemma to the bound of $\|\boldsymbol{x}_\varepsilon(t) - \mathbf{X}(t,0,\boldsymbol{x}_\varepsilon^0)\|$ in (2.6), we make use of crude estimates for simplicity:

$$\int_0^t e^{K(t-s)}s^a\,e^{Ks}\,\mathrm{d}s \le e^{Kt}\,t^{a+1}\,, \qquad\qquad \int_0^t e^{K(t-s)}s^a\,\mathrm{d}s \le e^{Kt}\,t^{a+1}\,,$$

Thus, one gets

$$\|\boldsymbol{x}_\varepsilon(t) - \mathbf{X}(t,0,\boldsymbol{x}_\varepsilon^0)\| \lesssim \varepsilon\,e^{K_x\,t}(1+t)\Big(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,K_0 + \varepsilon\,t\,(K_t \,+\, K_x\,K_0)\Big)\,.$$

This concludes the proof of the first part of Theorem 2.1 (using that $t \lesssim (1+t^2)$).

Now, in terms of the new variable $\boldsymbol{y}_\varepsilon := \boldsymbol{x}_\varepsilon - \varepsilon\,\boldsymbol{v}_\varepsilon^\perp$, defined in (2.2), equation (2.5) writes

$$\begin{cases}\boldsymbol{y}_\varepsilon'(t) \,=\, -\mathbf{E}^\perp(t,\boldsymbol{y}_\varepsilon(t) + \varepsilon\boldsymbol{v}_\varepsilon^\perp(t))\,, \qquad \forall t \ge 0,\\ \boldsymbol{y}_\varepsilon(0) = \boldsymbol{x}_\varepsilon^0 - \varepsilon(\boldsymbol{v}_\varepsilon^0)^\perp\,.\end{cases} \tag{2.8}$$

By Taylor expansion (with integral remainder) we obtain

$$\boldsymbol{y}_\varepsilon' \,=\, -\mathbf{E}^\perp(t,\boldsymbol{y}_\varepsilon) \,-\, \varepsilon\,\mathrm{d}_{\boldsymbol{x}}\mathbf{E}^\perp(t,\boldsymbol{y}_\varepsilon)\,\boldsymbol{v}_\varepsilon^\perp \,+\, \varepsilon^2\,\Theta_\varepsilon(t,\boldsymbol{y}_\varepsilon,\boldsymbol{v}_\varepsilon)\,, \tag{2.9}$$

where the remainder function Θ_ε is bounded as

$$\|\Theta_\varepsilon(t,\boldsymbol{y}_\varepsilon,\boldsymbol{v}_\varepsilon)\| \,\le\, \frac{1}{2}\,K_{xx}\,\|\boldsymbol{v}_\varepsilon\|^2\,.$$

Recalling that we would like to obtain an evolution equation for $\boldsymbol{y}_\varepsilon$ that would be an $\mathcal{O}(\varepsilon^2)$-perturbation of the guiding center equation in (2.1), the issue, now, is to replace the variable

$\boldsymbol{v}_\varepsilon$ on the right hand side of (2.9). For this purpose, we use, once again, the second equation of (1.7) written as $\boldsymbol{v}_\varepsilon^\perp = -\varepsilon^2\,\boldsymbol{v}_\varepsilon' + \varepsilon\mathbf{E}(t,\boldsymbol{x}_\varepsilon)$, and conclude that

$$\boldsymbol{y}_\varepsilon' = -\mathbf{E}^\perp(t,\boldsymbol{y}_\varepsilon) - \varepsilon^2\left[\mathrm{d}_{\boldsymbol{x}}\mathbf{E}^\perp(t,\boldsymbol{y}_\varepsilon)\,\mathbf{E}(t,\boldsymbol{x}_\varepsilon) + \Theta_\varepsilon(t,\boldsymbol{y}_\varepsilon,\boldsymbol{v}_\varepsilon)\right] + \varepsilon^3\,\mathrm{d}_{\boldsymbol{x}}\mathbf{E}^\perp(t,\boldsymbol{y}_\varepsilon)\,\boldsymbol{v}_\varepsilon'\,.$$

The last term of the foregoing equation will be written with the help of a complete time derivative, *i.e.*,

$$\mathrm{d}_{\boldsymbol{x}}\mathbf{E}^\perp(t,\boldsymbol{y}_\varepsilon)\,\boldsymbol{v}_\varepsilon' = \left(\mathrm{d}_{\boldsymbol{x}}\mathbf{E}^\perp(t,\boldsymbol{y}_\varepsilon)\,\boldsymbol{v}_\varepsilon\right)' - \left(\partial_t\mathrm{d}_{\boldsymbol{x}}\mathbf{E}^\perp(t,\boldsymbol{y}_\varepsilon) + \mathrm{d}^2_{\boldsymbol{x}}\mathbf{E}^\perp(t,\boldsymbol{y}_\varepsilon)\,\boldsymbol{y}_\varepsilon'\right)\boldsymbol{v}_\varepsilon\,.$$

Then, using (2.5), we obtain

$$\begin{aligned}\left(\boldsymbol{y}_\varepsilon - \varepsilon^3\mathrm{d}_{\boldsymbol{x}}\mathbf{E}^\perp(t,\boldsymbol{y}_\varepsilon)\boldsymbol{v}_\varepsilon\right)' &= -\mathbf{E}^\perp(t,\boldsymbol{y}_\varepsilon) - \varepsilon^2\left[\mathrm{d}_{\boldsymbol{x}}\mathbf{E}^\perp(t,\boldsymbol{y}_\varepsilon)\mathbf{E}(t,\boldsymbol{x}_\varepsilon) + \Theta_\varepsilon(t,\boldsymbol{y}_\varepsilon,\boldsymbol{v}_\varepsilon)\right]\\ &\quad -\varepsilon^3\left(\partial_t\mathrm{d}_{\boldsymbol{x}}\mathbf{E}^\perp(t,\boldsymbol{y}_\varepsilon) - \mathrm{d}^2_{\boldsymbol{x}}\mathbf{E}^\perp(t,\boldsymbol{y}_\varepsilon)\mathbf{E}^\perp(t,\boldsymbol{x}_\varepsilon)\right)\boldsymbol{v}_\varepsilon\,.\end{aligned}$$

which is an equation $\mathcal{O}(\varepsilon^2)$-close to the equation in (2.1).

Finally, and similarly as for the first estimate, we subtract the foregoing equation from the equation in (2.1) for a solution emanated from $\boldsymbol{x}_\varepsilon^0 - \varepsilon(\boldsymbol{v}_\varepsilon^0)^\perp$ at $s=0$, integrate over $[0,t]$, and use the Lipschitz bound on $\mathbf{E}$ as well as Lemma 2.3, to get

$$\begin{aligned}\|\boldsymbol{y}_\varepsilon(t) - \mathbf{X}(t,0,\boldsymbol{x}_\varepsilon^0 - \varepsilon(\boldsymbol{v}_\varepsilon^0)^\perp)\| &\le K_x\int_0^t \|\boldsymbol{y}_\varepsilon(s) - \mathbf{X}(s,0,\boldsymbol{x}_\varepsilon^0 - \varepsilon(\boldsymbol{v}_\varepsilon^0)^\perp)\|\,\mathrm{d}s\\ &\quad + \varepsilon^2\frac{K_{xx}}{2}\int_0^t\|\boldsymbol{v}_\varepsilon(s)\|^2\,\mathrm{d}s + \varepsilon^3(K_{tx}+K_{xx}K_0)\int_0^t\|\boldsymbol{v}_\varepsilon(s)\|\,\mathrm{d}s\\ &\quad + \varepsilon^3\,K_x\,(\|\boldsymbol{v}_\varepsilon^0\| + \|\boldsymbol{v}_\varepsilon(t)\|) + \varepsilon^2 K_x K_0\,t.\end{aligned}$$

In the same line of argument as for the first estimate, that is, using Lemma 2.3 and the Grönwall lemma, one obtains after some manipulations

$$\begin{aligned}&\|\boldsymbol{y}_\varepsilon(t) - \mathbf{X}(t,0,\boldsymbol{x}_\varepsilon^0 - \varepsilon(\boldsymbol{v}_\varepsilon^0)^\perp)\|\\ &\quad\lesssim \varepsilon^2\,t\,e^{K_x\,t}\,(1+K_x t)\,\left[e^{K_x\,t}\,K_{xx}\|\boldsymbol{v}_\varepsilon^0\|^2 + K_x\,K_0\right]\\ &\qquad + \varepsilon^3\,e^{K_x\,t}\,(1+K_x t)\,(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,K_0)\,\left[t\,e^{K_x\,t}\,K_0\,K_{xx} + t\,(K_{tx}+K_{xx}K_0) + K_x\right]\\ &\qquad + \varepsilon^4\,t\,e^{K_x\,t}\,(1+K_x t)\,(K_t\,+\,K_x\,K_0)\,\left[K_{xx}\,t^2\,e^{K_x\,t}\,(K_t\,+\,K_x\,K_0) + t\,(K_{tx}+K_{xx}\,K_0) + K_x\right]\,,\end{aligned}$$

which concludes the desired estimate by employing Young inequalities, typically, in the form

$$\varepsilon\,t\,\|\boldsymbol{v}_\varepsilon^0\| \lesssim \varepsilon^2\,t^2 + \|\boldsymbol{v}_\varepsilon^0\|^2\,.$$

2.3. **From characteristics to PDE's: estimates on the density.** As we will explain in this section, thanks to L^∞-bounds on the characteristics system (see Lemma 2.3) and its asymptotic evolution (see Theorem 2.1), it is straightforward to derive bounds on particle distributions, that is, on the densities, in the $W^{-1,1}$ topology. We recall that in the dual space $W^{-1,1} := (W^{1,\infty})^*$, the canonical seminorm is defined by

$$\|\mu\|_{\dot W^{-1,1}} := \sup_{\|\nabla\varphi\|_{L^\infty}\le 1}\left|\int\varphi\,\mathrm{d}\mu\right|\,,$$

for $\mu\in W^{-1,1}$. Incidentally, note that the seminorm on $W^{-1,1}$ defines a distance, equivalent to the 1-Wasserstein distance, on probability measures with a finite first moment.

Let us also recall the classical link between characteristics and solutions to continuity equations. In fact, the solution of an abstract continuity equation

$$\partial_t G + \mathrm{div}_{\boldsymbol{a}}(\mathcal{X}\,G) = 0\,,$$

with initial datum G_0, writes $G(t,\cdot) = \mathbf{A}(t,0,\cdot)_*(G_0)$, where $\mathbf{A}$ is the *flow* associated with the differential equation $\boldsymbol{a}' = \mathcal{X}(t,\boldsymbol{a})$, and $\mathcal{A}_*(\mu)$ denotes the push-forward of μ by $\mathcal{A}$, which is defined by

$$\langle\mathcal{A}_*(\mu),\varphi\rangle := \langle\mu,\varphi\circ\mathcal{A}\rangle\,,$$

for all test-functions φ. Note that the backbone of particle methods is the fact that the push-forward of a Dirac mass is given by $\mathcal{A}_*(\delta_{\boldsymbol{a}_0}) = \delta_{\mathcal{A}(\boldsymbol{a}_0)}$. Note also that when $\mathcal{X}$ is divergence-free, the formula matches the one solving the associated transport equation, namely $G(t,\cdot) = G_0 \circ \mathbf{A}(0,t,\cdot)$, where $\circ$ stands for function composition, but differs otherwise.

In particular, solutions to (1.3) are obtained as

$$f^\varepsilon(t,\cdot) = (\mathbf{X}_\varepsilon, \mathbf{V}_\varepsilon)(t,0,\cdot)_*(f_0),$$

and, consequently, the corresponding charge density reads

$$\rho^\varepsilon(t,\cdot) = \mathbf{X}_\varepsilon(t,0,\cdot)_*(f_0)\,.$$

Having these, one can readily derive the following estimate on the density.

Corollary 2.4. *Assume that the electric field is Lipschitz $\mathbf{E} \in W^{1,\infty}$ and that f_0 is a probability with finite first moment. Then, for any $\varepsilon > 0$, the following estimate holds*

$$\|\rho^\varepsilon(t,\cdot) - \rho(t,\cdot)\|_{\dot{W}^{-1,1}} \underset{\mathbf{E}}{\lesssim} \varepsilon\, e^{K_x\, t}(1+t^2) \int_{\mathbb{R}^2\times\mathbb{R}^2} (\varepsilon + \|\boldsymbol{v}\|)\, \mathrm{d}f_0(\boldsymbol{x},\boldsymbol{v})\,,$$

where ρ^ε is the charge density computed from f^ε solution to (1.3) whereas ρ solves to (1.6) with the same initial datum as ρ^ε.

Proof. For any test function $\varphi \in W^{1,\infty}(\mathbb{R}^2)$, one gets from the formula recalled above

$$\langle \rho^\varepsilon(t,\cdot) - \rho(t,\cdot), \varphi\rangle \;=\; \int_{\mathbb{R}^2\times\mathbb{R}^2} \left(\varphi\left(\mathbf{X}_\varepsilon(t,0,\boldsymbol{x},\boldsymbol{v})\right) - \varphi\left(\mathbf{X}(t,0,\boldsymbol{x}^0)\right)\right) \mathrm{d}f_0(\boldsymbol{x},\boldsymbol{v})\,.$$

so that

$$|\langle \rho^\varepsilon(t,\cdot) - \rho(t,\cdot), \varphi\rangle| \le \|\nabla\varphi\|_{L^\infty(\mathbb{R}^2)} \int_{\mathbb{R}^2\times\mathbb{R}^2} \|\mathbf{X}_\varepsilon(t,0,\boldsymbol{x},\boldsymbol{v}) - \mathbf{X}(t,0,\boldsymbol{x})\|\, \mathrm{d}f_0(\boldsymbol{x},\boldsymbol{v}),$$

and the proof is concluded by applying the first part of Theorem 2.1. □

3. First-order scheme on the original spatial variable

In the present section, we first complete the analysis of the first-order IMEX method introduced in [12]. For a chosen constant time discretization step Δt, we define a discrete time evolution, for $n \ge 0$, by

$$\begin{cases} \dfrac{\boldsymbol{x}_\varepsilon^{n+1} - \boldsymbol{x}_\varepsilon^n}{\Delta t} &= \dfrac{\boldsymbol{v}_\varepsilon^{n+1}}{\varepsilon}\,, \\[2ex] \varepsilon\, \dfrac{\boldsymbol{v}_\varepsilon^{n+1} - \boldsymbol{v}_\varepsilon^n}{\Delta t} &= \mathbf{E}(t_n, \boldsymbol{x}_\varepsilon^n) - \dfrac{(\boldsymbol{v}_\varepsilon^{n+1})^\perp}{\varepsilon}\,, \end{cases} \tag{3.1}$$

where $t_n = n\,\Delta t$. Note that the scheme (3.1) is semi-implicit but the implicit part (the velocity update) is linear. It only requires solving the 2×2 linear system

$$\boldsymbol{v}_\varepsilon^{n+1} + \frac{\Delta t}{\varepsilon^2}(\boldsymbol{v}_\varepsilon^{n+1})^\perp = \boldsymbol{v}_\varepsilon^n + \frac{\Delta t}{\varepsilon}\mathbf{E}(t_n, \boldsymbol{x}_\varepsilon^n)\,.$$

This highlights the role played by the matrix $\mathrm{Id} + \lambda\,\mathbf{J}$, with $\lambda := \Delta t/\varepsilon^2 > 0$ and $\mathbf{J}$ as in (1.4); see Lemma 3.6 for further discussion on this matrix.

As we will see later on, in Theorem 3.1, the scheme (3.1) has a unique solution, which means that it allows defining the sequence $(\boldsymbol{x}_\varepsilon^n, \boldsymbol{v}_\varepsilon^n)_{n\ge 0}$, which is expected to approximate the solution $(\boldsymbol{x}_\varepsilon, \boldsymbol{v}_\varepsilon)$ of (1.7) at times $(t_n)_{n\ge 0}$. The foregoing scheme is designed to capture the evolution of the space variable $\boldsymbol{x}_\varepsilon$ even when $\Delta t \gg \varepsilon^2$, *i.e.*, the asymptotic regime in which the traditional schemes are doomed to fail due to instability. We aim to show that this asymptotic convergence is sufficient for obtaining an ε-uniform error estimate, *i.e.*, to ensure that the convergence of the scheme for the numerical error $\|\boldsymbol{x}_\varepsilon(t_n) - \boldsymbol{x}_\varepsilon^n\|$ will be locally uniform with respect to $t_n \in \mathbb{R}^+$ and $\varepsilon \in \mathbb{R}^+$. Moreover, we will reveal that the convergence rate can be improved on the guiding center variable $\boldsymbol{y}_\varepsilon$ compared to $\boldsymbol{x}_\varepsilon$ itself, when $\varepsilon \ll 1$. Indeed, this variable, introduced in (2.2), may be thought as a

slower version of $\boldsymbol{x}_\varepsilon$, hence it is expectedly more advantageous to be used in the strongly oscillatory regime. To state comparisons for the discrete dynamics, we introduce the discrete counterpart of $\boldsymbol{y}_\varepsilon(t)$ in (2.2), that is,

$$\boldsymbol{y}_\varepsilon^n = \boldsymbol{x}_\varepsilon^n - \varepsilon\,(\boldsymbol{v}_\varepsilon^n)^\perp\,,$$

for which, using the scheme (3.1), one gets the update

$$\frac{\boldsymbol{y}_\varepsilon^{n+1} - \boldsymbol{y}_\varepsilon^n}{\Delta t} = -\mathbf{E}^\perp(t_n, \boldsymbol{y}_\varepsilon^n + \varepsilon\,(\boldsymbol{v}_\varepsilon^n)^\perp)\,. \tag{3.2}$$

By taking formally the limit $\varepsilon \to 0$, we would anticipate that, at the discrete level, the reduced asymptotic model for the limit of $\boldsymbol{x}_\varepsilon^n$ is the explicit Euler scheme, that is

$$\frac{\boldsymbol{x}^{n+1} - \boldsymbol{x}^n}{\Delta t} = -\mathbf{E}^\perp(t_n, \boldsymbol{x}^n)\,, \quad \text{for } n \geq 0\,. \tag{3.3}$$

Now, we can state our main theorem on the scheme (3.1).

Theorem 3.1. *The first-order scheme (3.1) possesses a unique solution. Moreover*

(i) when $\mathbf{E} \in W^{1,\infty}$, the space variable $\boldsymbol{x}_\varepsilon$ satisfies for all $n \geq 0$, $\Delta t > 0$ and $\varepsilon > 0$,

$$\|\boldsymbol{x}_\varepsilon^n - \boldsymbol{x}_\varepsilon(t_n)\| \underset{\mathbf{E}}{\lesssim} (1+t_n^2)\, e^{2K_x t_n}(1 + \|\boldsymbol{v}_\varepsilon^0\| + \varepsilon) \times \min\left(\frac{\Delta t}{\varepsilon^3}(1+\varepsilon^2), \Delta t + \varepsilon\right),$$

(ii) when $\mathbf{E} \in W^{2,\infty}$, the guiding center variable $\boldsymbol{y}_\varepsilon$ satisfies for all $n \geq 0$, $\Delta t > 0$ and $\varepsilon > 0$,

$$\|\boldsymbol{y}_\varepsilon^n - \boldsymbol{y}_\varepsilon(t_n)\| \underset{\mathbf{E}}{\lesssim} (1+t_n^4)\, e^{2K_x t_n}(1 + \|\boldsymbol{v}_\varepsilon^0\|^2 + \varepsilon^2) \times \min\left(\frac{\Delta t}{\varepsilon^3}\,(1+\varepsilon)\,, \Delta t + \varepsilon^2\right).$$

Corollary 3.2. *Theorem 3.1 implies that, for different regimes of Δt and ε,*

$$\|\boldsymbol{x}_\varepsilon^n - \boldsymbol{x}_\varepsilon(t_n)\| \underset{(t_n, \mathbf{E}, \|\boldsymbol{v}_\varepsilon^0\|)}{\lesssim} (1+\varepsilon) \begin{cases} \frac{\Delta t}{\varepsilon^3}(1+\varepsilon^2), & \text{if} \quad \Delta t \lesssim \varepsilon^4\,, \\ \varepsilon, & \text{if} \quad \varepsilon^4 \lesssim \Delta t \lesssim \varepsilon\,, \\ \Delta t, & \text{if} \quad \varepsilon \lesssim \Delta t\,, \end{cases} \tag{3.4}$$

and

$$\|\boldsymbol{y}_\varepsilon^n - \boldsymbol{y}_\varepsilon(t_n)\| \underset{(t_n, \mathbf{E}, \|\boldsymbol{v}_\varepsilon^0\|)}{\lesssim} (1+\varepsilon^2) \begin{cases} \frac{\Delta t}{\varepsilon^3}(1+\varepsilon), & \text{if} \quad \Delta t \lesssim \varepsilon^5\,, \\ \varepsilon^2, & \text{if} \quad \varepsilon^5 \lesssim \Delta t \lesssim \varepsilon^2\,, \\ \Delta t, & \text{if} \quad \varepsilon^2 \lesssim \Delta t\,. \end{cases} \tag{3.5}$$

Remark 3.3. *The estimates of Corollary 3.2 shows that for a fixed Δt, when ε is either sufficiently small or sufficiently large, both estimates boil down to a uniform $\mathcal{O}(\Delta t)$ estimate. Nonetheless, the worst-ε scenarios yield the following uniform rates :*

$$\|\boldsymbol{x}_\varepsilon^n - \boldsymbol{x}_\varepsilon(t_n)\| \underset{(t_n, \mathbf{E}, \|\boldsymbol{v}_\varepsilon^0\|)}{\lesssim} (\Delta t)^{1/4}\,, \qquad \|\boldsymbol{y}_\varepsilon^n - \boldsymbol{y}_\varepsilon(t_n)\| \underset{(t_n, \mathbf{E}, \|\boldsymbol{v}_\varepsilon^0\|)}{\lesssim} (\Delta t)^{2/5}\,,$$

obtained, respectively, when $\Delta t \sim \varepsilon^4$ and when $\Delta t \sim \varepsilon^5$.

Based on the foregoing estimates, it will be painless to obtain an error estimate at the particle density level. Note, however, that the full PIC error estimates would involve errors not due to the time discretizations and, hence, not taken into account here. To state the corresponding corollary, we denote the discrete flow $(\mathbf{X}_\varepsilon^{\Delta t}, \mathbf{V}_\varepsilon^{\Delta t})(t_n, t_s, \boldsymbol{x}^s, \boldsymbol{v}^s)$ as the solution to (3.1) starting from $(\boldsymbol{x}^s, \boldsymbol{v}^s)$ at index s. Note in particular that $n \geq s \geq 0$ and $t_n = n\Delta t$, $t_s = s\Delta t$. For later use, we also set

$$\mathbf{Y}_\varepsilon^{\Delta t}(t_n, t_s, \boldsymbol{x}^s, \boldsymbol{v}^s) := (\mathbf{X}_\varepsilon^{\Delta t} - \varepsilon(\mathbf{V}_\varepsilon^{\Delta t})^\perp)(t_n, t_s, \boldsymbol{x}^s, \boldsymbol{v}^s).$$

Corollary 3.4. *Assume that the electric field is* $\mathbf{E} \in W^{1,\infty}$ *and that* f_0 *is a probability with finite first moment. Then, for any* $\varepsilon > 0$*, the following uniform error estimate holds*

$$\|\rho^\varepsilon_{\Delta t}(t_n,\cdot) - \rho^\varepsilon(t_n,\cdot)\|_{W^{-1,1}} \underset{\mathbf{E}}{\lesssim} \min\left(\frac{\Delta t}{\varepsilon^3}(1+\varepsilon^2), \Delta t + \varepsilon\right) e^{2K_x\, t_n}(1+t_n^2) \int_{\mathbb{R}^2\times\mathbb{R}^2} (1+\varepsilon+\|\boldsymbol{v}\|)\, \mathrm{d}f_0(\boldsymbol{x},\boldsymbol{v})\,,$$

where ρ^ε *is the charge density computed from* f^ε *solution to* (1.3) *whereas* $\rho^\varepsilon_{\Delta t}$ *is defined at times* $(t_n)_{n\in\mathbf{N}} = (n\Delta t)_{n\in\mathbf{N}}$ *by*

$$\rho^\varepsilon_{\Delta t}(t_n,\cdot) \;=\; \mathbf{X}^{\Delta t}_\varepsilon(t_n,0,\cdot,\cdot)_*(f_0)\,.$$

As suggested by the presence of the min function in claimed estimates, the proof of Theorem 3.1, provided in subsequent subsections, combines two kinds of estimates: the $\Delta t/\varepsilon^3$ part which arises from a detailed version of classical convergence estimates (Proposition 3.7 below), and ε-uniform part which stems from combining three estimates through the triangle inequality, namely, asymptotic estimates at continuous (Theorem 2.1) and discrete (Proposition 3.11 below) levels, and a classical convergence estimate for the non-stiff reduced asymptotic model (ε-independent), which will be discussed in Proposition 3.9 below. Thus, eventually, it leads to an error estimate which is $\mathcal{O}(\varepsilon+\Delta t)$, *i.e.*,

$$\|\boldsymbol{x}_\varepsilon(t_n) - \boldsymbol{x}^n_\varepsilon\| \;\le\; \|\boldsymbol{x}_\varepsilon(t_n) - \boldsymbol{x}(t_n)\| + \|\boldsymbol{x}(t_n) - \boldsymbol{x}^n\| + \|\boldsymbol{x}^n - \boldsymbol{x}^n_\varepsilon\| \underset{(t_n,\mathbf{E},\|\boldsymbol{v}^0_\varepsilon\|)}{\lesssim} \varepsilon + \Delta t + \varepsilon\,,$$

with $\boldsymbol{x}(t_n)$ solving (2.1) and $(\boldsymbol{x}^n)_{n\ge0}$ solving (3.3).

3.1. **Direct convergence estimates.** In this section, we discuss direct convergence estimates, *i.e.*, we estimate the difference of the numerical solution and the exact one, for the ε-dependent system (in Proposition 3.7) and the asymptotic model (in Proposition 3.9). The former estimate is called, hereinafter, the *direct estimate* and gives rise to the $\Delta t/\varepsilon^3$ part in Theorem 3.1.

3.1.1. *ε-dependent direct estimate.* We begin with the direct estimate for which it would be more convenient to carry out the analysis in variables $(\boldsymbol{y}_\varepsilon, \boldsymbol{v}_\varepsilon)$ rather than $(\boldsymbol{x}_\varepsilon, \boldsymbol{v}_\varepsilon)$. The first step is the direct consistency analysis; we introduce consistency errors, using the $\boldsymbol{y}_\varepsilon$-update and (3.3), as

$$\boldsymbol{\tau}^n_{\boldsymbol{y}} \;:=\; \frac{\boldsymbol{y}_\varepsilon(t_{n+1}) - \boldsymbol{y}_\varepsilon(t_n)}{\Delta t} \;+\; \mathbf{E}^\perp(t_n, \boldsymbol{y}_\varepsilon(t_n) \;+\; \varepsilon\, \boldsymbol{v}^\perp_\varepsilon(t_n))\,, \tag{3.6a}$$

$$\boldsymbol{\tau}^n_{\boldsymbol{v}} \;:=\; \frac{\boldsymbol{v}_\varepsilon(t_{n+1}) - \boldsymbol{v}_\varepsilon(t_n)}{\Delta t} \;-\; \frac{\mathbf{E}(t_n, \boldsymbol{y}_\varepsilon(t_n) + \varepsilon\, \boldsymbol{v}^\perp_\varepsilon(t_n))}{\varepsilon} \;+\; \frac{\boldsymbol{v}^\perp_\varepsilon(t_{n+1})}{\varepsilon^2}\,. \tag{3.6b}$$

These local truncation errors can be bounded as in the following lemma.

Lemma 3.5. *For every* $\varepsilon > 0$ *and* $\Delta t > 0$*, the truncation errors* $(\boldsymbol{\tau}^n_{\boldsymbol{y}})_{n\ge0}$ *and* $(\boldsymbol{\tau}^n_{\boldsymbol{v}})_{n\ge0}$ *defined in* (3.6a)–(3.6a) *are bounded as*

$$\|\boldsymbol{\tau}^n_{\boldsymbol{y}}\| \underset{\mathbf{E}}{\lesssim} \frac{\Delta t}{\varepsilon}\; e^{K_x\, t_{n+1}}\; \Big(\|\boldsymbol{v}^0_\varepsilon\| + \varepsilon\,(1+t_{n+1})\Big)\,,$$

$$\|\boldsymbol{\tau}^n_{\boldsymbol{v}}\| \underset{\mathbf{E}}{\lesssim} \frac{\Delta t}{\varepsilon^4}\; e^{K_x\, t_{n+1}}\; (1+\varepsilon^2)\,\Big(\|\boldsymbol{v}^0_\varepsilon\| + \varepsilon\,(1+t_{n+1})\Big)\,.$$

Proof. The truncation error $\boldsymbol{\tau}^n_{\boldsymbol{y}}$ can be written as

$$\boldsymbol{\tau}^n_{\boldsymbol{y}} \;=\; \frac{\boldsymbol{y}_\varepsilon(t_{n+1}) - \boldsymbol{y}_\varepsilon(t_n)}{\Delta t} - \boldsymbol{y}'_\varepsilon(t_n)\,.$$

Thus, for any $n \ge 0$, $\|\boldsymbol{\tau}^n_{\boldsymbol{y}}\| \le \frac{1}{2}\Delta t\, \max_{[t_n,t_{n+1}]} \|\boldsymbol{y}''_\varepsilon\|$, where the second derivative reads

$$\boldsymbol{y}''_\varepsilon(t) \;=\; -\partial_t \mathbf{E}^\perp(t, \boldsymbol{x}_\varepsilon(t)) - \frac{1}{\varepsilon}\mathrm{d}_{\boldsymbol{x}}\mathbf{E}^\perp(t, \boldsymbol{x}_\varepsilon(t))\boldsymbol{v}_\varepsilon(t)\,,$$

which, thanks to Lemma 2.3, yields the first bound:

$$\begin{aligned}\|\boldsymbol{\tau}_{\boldsymbol{y}}^n\| &\leq \frac{\Delta t}{2}\left((K_t + K_x K_0)(1 + K_x\, t_{n+1}\, e^{K_x\, t_{n+1}}) + \frac{K_x}{\varepsilon} e^{K_x\, t_{n+1}}\left(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\, K_0\right)\right),\\ &\underset{\mathbf{E}}{\lesssim} \frac{\Delta t}{\varepsilon}\, e^{K_x\, t_{n+1}}\, \left(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,(1 + t_{n+1})\right).\end{aligned}$$

Likewise, for the second estimate, one has

$$\boldsymbol{\tau}_{\boldsymbol{v}}^n = \frac{\boldsymbol{v}_\varepsilon(t_{n+1}) - \boldsymbol{v}_\varepsilon(t_n)}{\Delta t} - \boldsymbol{v}_\varepsilon'(t_n) + \frac{\boldsymbol{v}_\varepsilon^\perp(t_{n+1}) - \boldsymbol{v}_\varepsilon^\perp(t_n)}{\varepsilon^2},$$

which implies the estimate

$$\|\boldsymbol{\tau}_{\boldsymbol{v}}^n\| \leq \frac{\Delta t}{2} \max_{[t_n, t_{n+1}]} \|\boldsymbol{v}_\varepsilon''\| + \frac{\Delta t}{\varepsilon^2} \max_{[t_n, t_{n+1}]} \|\boldsymbol{v}_\varepsilon'\|.$$

Moreover, the second order derivative reads

$$\boldsymbol{v}_\varepsilon''(t) = \frac{1}{\varepsilon}\partial_t \mathbf{E}(t, \boldsymbol{x}_\varepsilon(t)) + \frac{1}{\varepsilon^2}\mathrm{d}_{\boldsymbol{x}}\mathbf{E}(t, \boldsymbol{x}_\varepsilon(t))\boldsymbol{v}_\varepsilon(t) - \frac{1}{\varepsilon^3}(\mathbf{E}(t, \boldsymbol{x}_\varepsilon(t)))^\perp - \frac{1}{\varepsilon^4}\boldsymbol{v}_\varepsilon(t),$$

which concludes the proof combined with Lemma 2.3, *i.e.*,

$$\begin{aligned}\|\boldsymbol{\tau}_{\boldsymbol{v}}^n\| &\lesssim \frac{\Delta t}{\varepsilon}\, e^{K_x\, t_{n+1}} \left((K_t + K_x K_0)\left[1 + \left(K_x + \frac{1}{\varepsilon^2}\right) t_{n+1}\right] + \left(K_0 + \frac{\|\boldsymbol{v}_\varepsilon^0\|\|}{\varepsilon}\right)\left[K_x + \frac{1}{\varepsilon^2}\right]\right)\\ &\underset{\mathbf{E}}{\lesssim} \frac{\Delta t}{\varepsilon^4}\, e^{K_x\, t_{n+1}}\, (1 + \varepsilon^2)\left(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,(1 + t_{n+1})\right).\end{aligned}$$

□

In order to obtain the $\Delta t/\varepsilon^3$ bounds, the missing piece is the stability analysis. So, we aim to investigate the stability of the implicit part, with the corresponding matrix $\mathrm{Id} + \lambda\,\mathbf{J}$ to be inverted. The following stability result is based on the fact that $\mathbf{J}$ is skew-symmetric and $\mathbf{J}^2 = -\mathrm{Id}$. Its appearance in our convergence analysis reveals the L-stability of the implicit part of our first-order IMEX schemes, that removes high frequency oscillations.

Lemma 3.6. *Let* $\mathbf{J}$ *be as in* (1.4). *Then, for any* $\lambda > 0$, $\mathrm{Id} + \lambda\mathbf{J}$ *is invertible and*

$$(\mathrm{Id} + \lambda\,\mathbf{J})^{-1} = \frac{1}{1 + \lambda^2}(\mathrm{Id} - \lambda\,\mathbf{J}),$$

whose norm is bounded like

$$\Lambda_\lambda^{-1} := \left\|(\mathrm{Id} + \lambda\,\mathbf{J})^{-1}\right\| = \frac{1}{\sqrt{1 + \lambda^2}} < \min\left(1, \frac{1}{\lambda}\right). \tag{3.7}$$

Proof. The formula for the inverse is readily deduced from $\mathbf{J}^2 = -\mathrm{Id}$. To compute its norm, note that, for any vector $\boldsymbol{a}$, $\boldsymbol{a}$ and $\mathbf{J}\boldsymbol{a}$ are orthogonal and $\|\mathbf{J}\boldsymbol{a}\| = \|\boldsymbol{a}\|$. From this, follows $\|\mathrm{Id} - \lambda\,\mathbf{J}\| = \sqrt{1 + \lambda^2}$, hence the formula for Λ_λ^{-1}. The final estimate stems from $1 + \lambda^2 > \max(1, \lambda^2)$. □

Gathering consistency estimates from Lemma 3.5 and stability information from Lemma 3.6, we now provide direct error bounds in Proposition 3.7, corresponding to the $\Delta t/\varepsilon^3$ part of estimates in Theorem 3.1.

Proposition 3.7. *Assume that* $\mathbf{E} \in W^{2,\infty}$. *Then, the error of the unique solution of the scheme* (3.1) *denoted by* $(\boldsymbol{x}_\varepsilon^n, \boldsymbol{v}_\varepsilon^n, \boldsymbol{y}_\varepsilon^n)_{n\geq 0}$, *in the convergence to* $(\boldsymbol{x}_\varepsilon(t_n), \boldsymbol{v}_\varepsilon(t_n), \boldsymbol{y}_\varepsilon(t_n))$ *as the exact solution of the system* (1.7), *is bounded as*

$$\begin{aligned}\|\boldsymbol{x}_\varepsilon^n - \boldsymbol{x}_\varepsilon(t_n)\| &\leq \|\boldsymbol{y}_\varepsilon^n - \boldsymbol{y}_\varepsilon(t_n)\| + \varepsilon\|\boldsymbol{v}_\varepsilon^n - \boldsymbol{v}_\varepsilon(t_n)\|\\ &\underset{\mathbf{E}}{\lesssim} \frac{\Delta t}{\varepsilon^3}\, t_n\, e^{2K_x t_n}(1 + \varepsilon^2)\Big(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,(1 + t_n)\Big).\end{aligned} \tag{3.8}$$

Remark 3.8. *As implicit in the foregoing statement, due to the fact that nonlinear terms only depend on* $(\boldsymbol{y},\boldsymbol{v})$ *through* $\boldsymbol{x}=\boldsymbol{y}+\varepsilon\boldsymbol{v}^\perp$*, it is expedient to use the norm* $\|\cdot\|_\varepsilon$ *defined by*

$$\|(\boldsymbol{y},\boldsymbol{v})\|_\varepsilon := \|\boldsymbol{y}\| + \varepsilon\|\boldsymbol{v}\|\,. \tag{3.9}$$

Proof. For the sake of conciseness, we denote numerical errors by

$$\boldsymbol{e}_{\boldsymbol{y}}^n := \boldsymbol{y}_\varepsilon^n - \boldsymbol{y}_\varepsilon(t_n)\,, \qquad \boldsymbol{e}_{\boldsymbol{v}}^n := \boldsymbol{v}_\varepsilon^n - \boldsymbol{v}_\varepsilon(t_n)\,.$$

By reformulating the velocity update in (3.1) as

$$\boldsymbol{v}_\varepsilon^{n+1} = \left(\mathrm{Id} + \frac{\Delta t}{\varepsilon^2}\mathbf{J}\right)^{-1}\left(\boldsymbol{v}_\varepsilon^n + \frac{\Delta t}{\varepsilon}\mathbf{E}(t_n,\boldsymbol{y}_\varepsilon^n + \varepsilon\,(\boldsymbol{v}_\varepsilon^n)^\perp)\right),$$

and using the stability of the implicit operator in Lemma 3.6, we obtain the following bounds for $\|\boldsymbol{e}_{\boldsymbol{y}}^{n+1}\|$ and $\|\boldsymbol{e}_{\boldsymbol{v}}^{n+1}\|$ for all $n\geq 0$:

$$\begin{aligned}
\|\boldsymbol{e}_{\boldsymbol{y}}^{n+1}\| &\leq \|\boldsymbol{e}_{\boldsymbol{y}}^n\| + K_x\Delta t\,\|(\boldsymbol{e}_{\boldsymbol{y}}^n,\boldsymbol{e}_{\boldsymbol{v}}^n)\|_\varepsilon + \Delta t\,\|\boldsymbol{\tau}_{\boldsymbol{y}}^n\|,\\
\|\boldsymbol{e}_{\boldsymbol{v}}^{n+1}\| &\leq \Lambda_\lambda^{-1}\left(\|\boldsymbol{e}_{\boldsymbol{v}}^n\| + K_x\frac{\Delta t}{\varepsilon}\,\|(\boldsymbol{e}_{\boldsymbol{y}}^n,\boldsymbol{e}_{\boldsymbol{v}}^n)\|_\varepsilon + \Delta t\|\boldsymbol{\tau}_{\boldsymbol{v}}^n\|\right),
\end{aligned}$$

which, thanks to Lemma 3.6, yields,

$$\|(\boldsymbol{e}_{\boldsymbol{y}}^{n+1},\boldsymbol{e}_{\boldsymbol{v}}^{n+1})\|_\varepsilon \leq (1+2K_x\Delta t)\|(\boldsymbol{e}_{\boldsymbol{y}}^n,\boldsymbol{e}_{\boldsymbol{v}}^n)\|_\varepsilon + \Delta t\,\|(\boldsymbol{\tau}_{\boldsymbol{y}}^n,\boldsymbol{\tau}_{\boldsymbol{v}}^n)\|_\varepsilon\,, \qquad n\geq 0\,.$$

Note that $\|(\boldsymbol{e}_{\boldsymbol{y}}^0,\boldsymbol{e}_{\boldsymbol{v}}^0)\|_\varepsilon = 0$. Thus, by iteration, we deduce that for any $n>0$,

$$\begin{aligned}
\|(\boldsymbol{e}_{\boldsymbol{y}}^n,\boldsymbol{e}_{\boldsymbol{v}}^n)\|_\varepsilon &\leq \sum_{\ell=0}^{n-1}(1+2K_x\Delta t)^{n-\ell-1}\|(\boldsymbol{\tau}_{\boldsymbol{y}}^\ell,\boldsymbol{\tau}_{\boldsymbol{v}}^\ell)\|_\varepsilon\;\Delta t\,,\\
&\leq \sum_{\ell=0}^{n-1} e^{2\,K_x\,(t_n-t_{\ell+1})}\|(\boldsymbol{\tau}_{\boldsymbol{y}}^\ell,\boldsymbol{\tau}_{\boldsymbol{v}}^\ell)\|_\varepsilon\;\Delta t\,.
\end{aligned}$$

The proof is, then, concluded by applying the consistency result in Lemma 3.5. □

3.1.2. *ε-independent classical estimate for the asymptotic model.* Here, we discuss the classical convergence analysis of the asymptotic numerical model (3.3) to the guiding center equation (2.1).

Proposition 3.9. *Assume that* $\mathbf{E}\in W^{1,\infty}$*. Then, it holds*

$$\|\boldsymbol{x}^n - \boldsymbol{x}(t_n)\| \leq \frac{\Delta t}{2}\,t_n\,e^{K_x\,t_n}\left(K_t + K_xK_0\right), \tag{3.10}$$

when $(\boldsymbol{x}^n)_{n\in\mathbf{N}}$ *and* $\boldsymbol{x}(t_n)$ *solve respectively* (3.3) *and* (2.1)*, with the same initial data.*

Proof. The proof is omitted as, thanks to the ε-independence of the asymptotic model, it boils down to a simpler version of the proof of Proposition 3.7, indeed, the quite classical error analysis of forward Euler integration. □

3.2. **Asymptotic estimates.** To obtain a bound on the asymptotic error $\|\boldsymbol{x}^n-\boldsymbol{x}_\varepsilon^n\|$, thus a discrete counterpart of Theorem 2.1, we refine below the analysis of [12, Section 4.1].

The first step is to obtain ε-uniform, local-in-time estimates, thus a discrete counterpart of Lemma 2.3. To do so, it is convenient to introduce $(\boldsymbol{z}_\varepsilon^n)_{n\geq 1}$ as the discrete analogue of the auxiliary variable $\boldsymbol{z}_\varepsilon$, *i.e.*,

$$\boldsymbol{z}_\varepsilon^n := \boldsymbol{v}_\varepsilon^n + \varepsilon\,\mathbf{E}^\perp(t_{n-1},\boldsymbol{x}_\varepsilon^{n-1}), \qquad n\geq 1\,. \tag{3.11}$$

Then, from (3.1), it follows

$$\frac{\boldsymbol{z}_\varepsilon^{n+1}-\boldsymbol{z}_\varepsilon^n}{\Delta t} = -\frac{1}{\varepsilon^2}(\boldsymbol{z}_\varepsilon^{n+1})^\perp + \varepsilon\,\frac{\mathbf{E}^\perp(t_n,\boldsymbol{x}_\varepsilon^n)-\mathbf{E}^\perp(t_{n-1},\boldsymbol{x}_\varepsilon^{n-1})}{\Delta t}, \quad \text{for } n\geq 1\,. \tag{3.12}$$

Thus, the following lemma provides the required estimates.

Lemma 3.10. *Assume that* $\mathbf{E} \in W^{1,\infty}$ *and consider* (3.1). *Then, the variable* $(\boldsymbol{z}_\varepsilon^n)_{n\geq 1}$ *defined in* (3.11) *and the velocity computed by the scheme are bounded, for* $n \geq 1$, *as*

$$\begin{cases} \|\boldsymbol{z}_\varepsilon^n\| \;\leq\; e^{K_x\, t_n}\big(\|\boldsymbol{v}_\varepsilon^0\| + 2\,\varepsilon\, K_0\big) + \varepsilon\, t_n\, e^{K_x\, t_n}\,(K_t \;+\; K_x\, K_0)\;, \\ \|\boldsymbol{v}_\varepsilon^n\| \;\leq\; \|\boldsymbol{z}_\varepsilon^n\| \;+\; K_0\,\varepsilon\,. \end{cases}$$

Proof. First, we observe that for any $n \geq 1$,

$$\begin{aligned} \|\mathbf{E}^\perp(t_n, \boldsymbol{x}_\varepsilon^n) - \mathbf{E}^\perp(t_{n-1}, \boldsymbol{x}_\varepsilon^{n-1})\| &\leq \left(K_t \;+\; K_x \frac{\|\boldsymbol{v}_\varepsilon^n\|}{\varepsilon}\right)\Delta t\,, \\ &\leq \left(K_t + K_x\left(\frac{\|\boldsymbol{z}_\varepsilon^n\|}{\varepsilon} + K_0\right)\right)\Delta t\,, \end{aligned}$$

which, combined with the $\boldsymbol{z}_\varepsilon$-update (3.12) and Lemma 3.6, yields

$$\|\boldsymbol{z}_\varepsilon^{n+1}\| \leq\; (1 + K_x\,\Delta t)\;\|\boldsymbol{z}_\varepsilon^n\| \;+\; \varepsilon\,\Delta t\;(K_t \;+\; K_x\, K_0)\;, \qquad \text{for } n \geq 1.$$

Iterating on the foregoing estimate implies

$$\|\boldsymbol{z}_\varepsilon^n\| \leq\; e^{K_x(t_n - t_1)}\,\|\boldsymbol{z}_\varepsilon^1\| \;+\; \varepsilon\,(t_n - t_1)e^{K_x(t_n-t_1)}\,(K_t \;+\; K_x\, K_0)\;, \qquad \text{for } n \geq 1\,, \tag{3.13}$$

which gives the final estimate, since $\|\boldsymbol{z}_\varepsilon^1\| \leq \|\boldsymbol{v}_\varepsilon^1\| + \varepsilon\, K_0$ and $\boldsymbol{v}_\varepsilon^1$ is bounded from (3.1) and Lemma 3.6 as

$$\|\boldsymbol{v}_\varepsilon^1\| \leq \|\boldsymbol{v}_\varepsilon^0\| \;+\; \varepsilon\, K_0\,.$$

□

With ε-uniform bounds in hands, we are in a position to prove asymptotic estimates.

Proposition 3.11. *(i) Assume* $\mathbf{E} \in W^{1,\infty}$. *The difference between flows of* (3.1) *and* (3.3) *satisfies, for any* $\varepsilon > 0$,

$$\|\mathbf{X}_\varepsilon^{\Delta t}(t_n, 0, \boldsymbol{x}_\varepsilon^0, \boldsymbol{v}_\varepsilon^0) - \mathbf{X}^{\Delta t}(t_n, 0, \boldsymbol{x}_\varepsilon^0)\| \underset{\mathbf{E}}{\lesssim} \varepsilon\, e^{K_x t_n}\,(1 + \;t_n^2)\;\big(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\big)\;.$$

(ii) Assume $\mathbf{E} \in W^{2,\infty}$. *Then, for any* $\varepsilon > 0$,

$$\|\mathbf{Y}_\varepsilon^{\Delta t}(t_n, 0, \boldsymbol{x}_\varepsilon^0, \boldsymbol{v}_\varepsilon^0) \;-\; \mathbf{X}^{\Delta t}(t_n, 0, \boldsymbol{x}_\varepsilon^0 - \varepsilon(\boldsymbol{v}_\varepsilon^0)^\perp)\| \underset{\mathbf{E}}{\lesssim} e^{2K_x t_n}\,(1 + t_n^4)\,\big(\varepsilon\,\Delta t\,\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon^2\,(1 + \|\boldsymbol{v}_\varepsilon^0\|^2 + \varepsilon^2)\big)\,.$$

Proof. To obtain the first estimate, we subtract (3.3) from (3.2) and make a summation from 0 to $(n-1)$ to get

$$\boldsymbol{y}_\varepsilon^n - \boldsymbol{x}^n = \boldsymbol{y}_\varepsilon^0 - \boldsymbol{x}^0 \;-\; \Delta t \sum_{\ell=0}^{n-1}(\mathbf{E}^\perp(t_\ell, \boldsymbol{x}_\varepsilon^\ell) - \mathbf{E}^\perp(t_\ell, \boldsymbol{x}^\ell))\,, \qquad \text{for } n \geq 0,$$

with $(\boldsymbol{x}^n)_{n\geq 0} := (\mathbf{X}^{\Delta t}(t_n, 0, \boldsymbol{x}_\varepsilon^0))_{n\geq 0}$. Thus,

$$\|\boldsymbol{x}_\varepsilon^n - \boldsymbol{x}^n\| \;\leq\; \varepsilon\,(\|\boldsymbol{v}_\varepsilon^n\| + \|\boldsymbol{v}_\varepsilon^0\|) + K_x \Delta t \sum_{\ell=0}^{n-1}\|\boldsymbol{x}_\varepsilon^\ell - \boldsymbol{x}^\ell\|\,, \qquad \text{for } n \geq 0\,.$$

Now we use the discrete Grönwall lemma, in the form that

$$A_n \leq a_n + K\Delta t \sum_{\ell=0}^{n-1} A_\ell, \qquad \text{for } n \geq 0, \tag{3.14a}$$

(A, a and K being non-negative) implies

$$A_n \leq a_n + K\Delta t \sum_{\ell=0}^{n-1}(1 + K\Delta t)^{n-(\ell+1)} a_\ell \leq a_n + K\Delta t \sum_{\ell=0}^{n-1} e^{K(t_n - t_{\ell+1})} a_\ell\,, \qquad \text{for } n \geq 0. \tag{3.14b}$$

Following similar steps as of the continuous case, this leads to the estimate

$$\|\boldsymbol{x}_\varepsilon^n - \boldsymbol{x}^n\| \le \varepsilon\, e^{K_x t_n}\,(1+K_x\,t_n)\,\left(\|\boldsymbol{v}_\varepsilon^0\| + \max_{0\le\ell\le n}(e^{-K_x t_\ell}\|\boldsymbol{v}_\varepsilon^\ell\|)\right), \qquad n\ge 0\,,$$

which proves the first inequality, thanks to the velocity estimate in Lemma 3.10.

As in continuous case, the proof of the second estimate requires significantly more algebraic manipulations. One starts with the Taylor expansion of the right hand side of (3.2), that is,

$$\frac{\boldsymbol{y}_\varepsilon^{n+1}-\boldsymbol{y}_\varepsilon^n}{\Delta t} = -\,\mathbf{E}^\perp(t_n,\boldsymbol{y}_\varepsilon^n) \,-\, \varepsilon\,\mathrm{d}_x\mathbf{E}^\perp(t_n,\boldsymbol{y}_\varepsilon^n)(\boldsymbol{v}_\varepsilon^n)^\perp + \varepsilon^2\Theta_\varepsilon(t_n,\boldsymbol{y}_\varepsilon^n,\boldsymbol{v}_\varepsilon^n)\,, \qquad n\ge 0\,, \tag{3.15}$$

with the remainder term Θ_ε such that $\|\Theta_\varepsilon(t,\boldsymbol{y},\boldsymbol{v})\| \le \frac{1}{2}K_{xx}\|\boldsymbol{v}\|^2$. So as to rewrite the linear term $\boldsymbol{v}_\varepsilon^n$ in (3.15) using again (3.1), we observe that, for $n\ge 1$,

$$\begin{aligned}
\mathrm{d}_x\mathbf{E}(t_n,\boldsymbol{y}_\varepsilon^n)(\boldsymbol{v}_\varepsilon^n)^\perp &= -\,\varepsilon^2\,\mathrm{d}_x\mathbf{E}(t_n,\boldsymbol{y}_\varepsilon^n)\left(\frac{\boldsymbol{v}_\varepsilon^n-\boldsymbol{v}_\varepsilon^{n-1}}{\Delta t} \,-\, \frac{1}{\varepsilon}\mathbf{E}(t_{n-1},\boldsymbol{x}_\varepsilon^{n-1})\right),\\
&= -\,\frac{\varepsilon^2}{\Delta t}\left(\mathrm{d}_x\mathbf{E}(t_n,\boldsymbol{y}_\varepsilon^n)\boldsymbol{v}_\varepsilon^n - \mathrm{d}_x\mathbf{E}(t_{n-1},\boldsymbol{y}_\varepsilon^{n-1})\boldsymbol{v}_\varepsilon^{n-1}\right) + \varepsilon\,\mathrm{d}_{\boldsymbol{x}}\mathbf{E}(t_n,\boldsymbol{y}_\varepsilon^n)\,\mathbf{E}(t_{n-1},\boldsymbol{x}_\varepsilon^{n-1})\\
&\quad + \frac{\varepsilon^2}{\Delta t}\left(\mathrm{d}_x\mathbf{E}(t_n,\boldsymbol{y}_\varepsilon^n) - \mathrm{d}_x\mathbf{E}(t_{n-1},\boldsymbol{y}_\varepsilon^{n-1})\right)\,\boldsymbol{v}_\varepsilon^{n-1}\,,
\end{aligned}$$

whose last term can be bounded, using (3.2), as

$$\|\mathrm{d}_x\mathbf{E}(t_n,\boldsymbol{y}_\varepsilon^n) - \mathrm{d}_x\mathbf{E}(t_{n-1},\boldsymbol{y}_\varepsilon^{n-1})\| \le \Delta t\,\left(K_{tx}+K_{xx}\,K_0\right)\,, \qquad \text{for } n\ge 1.$$

Then, for all $n\ge 0$, we denote $\boldsymbol{y}^n := \mathbf{X}^{\Delta t}(t_n,0,\boldsymbol{x}_\varepsilon^0-\varepsilon(\boldsymbol{v}_\varepsilon^0)^\perp)$, and subtract (3.15) from the equation satisfied by $(\boldsymbol{y}^n)_{n\ge 0}$, given by (3.3). Inserting the latter reformulation in (3.15) and summing give

$$\|\boldsymbol{y}_\varepsilon^n - \boldsymbol{y}^n\| \;\le\; \|\boldsymbol{y}_\varepsilon^1-\boldsymbol{y}^1\| + \varepsilon^2 r_n \;+\; K_x\Delta t\sum_{\ell=1}^{n-1}\|\boldsymbol{y}_\varepsilon^\ell-\boldsymbol{y}^\ell\|\,, \qquad \text{for } n\ge 1,$$

with r_n defined as

$$\begin{aligned}
r_n := {}& \varepsilon\,K_x\,(\|\boldsymbol{v}_\varepsilon^{n-1}\| + \|\boldsymbol{v}_\varepsilon^1\|) \;+\; K_x\,K_0\,(t^n-t^1)\\
&+\;\varepsilon\,\Delta t\,(K_{tx}+K_{xx}\,K_0)\sum_{\ell=1}^{n-1}\|\boldsymbol{v}_\varepsilon^{\ell-1}\| \;+\; \Delta t\,\frac{K_{xx}}{2}\sum_{\ell=1}^{n-1}\|\boldsymbol{v}_\varepsilon^\ell\|^2\,.
\end{aligned}$$

So, the discrete Grönwall lemma yields

$$\|\boldsymbol{y}_\varepsilon^n-\boldsymbol{y}^n\| \le e^{K_x(t_n-t_1)}\,(1+K_x\,(t_n-t_1))\,\left(\|\boldsymbol{y}_\varepsilon^1-\boldsymbol{y}^1\| + \varepsilon^2\max_{2\le\ell\le n}(e^{-K_x(t_\ell-t_1)}r_\ell)\right), \qquad n\ge 1\,.$$

Moreover, the $\boldsymbol{y}_\varepsilon$-update (3.2) implies that, for the first step,

$$\|\boldsymbol{y}_\varepsilon^1-\boldsymbol{y}^1\| \;\le\; K_x\Delta t\,\varepsilon\,\|\boldsymbol{v}_\varepsilon^0\|\,.$$

The proof is, then, concluded by combining this estimate with the bound for the velocity in Lemma 3.10. □

3.3. **Proofs of Theorem 3.1 and Corollary 3.4.** Lemma 3.6 confirms the unique solvability of the scheme, *i.e.*, the matrix $\mathrm{Id}+\lambda\mathbf{J}$ is invertible, so the implicit part provides a unique velocity update. As mentioned hereinbefore, the error estimates in Theorem 3.1 is obtained by taking the minimum of the estimates from Proposition 3.7 and the sum of estimates in Theorem 2.1 and Propositions 3.9 and 3.11. Indeed since $\varepsilon\,\Delta t \lesssim \varepsilon^2+\Delta t^2 \le \varepsilon^2 + t_n\Delta t$ for $n\ge 1$, the presence in Proposition 3.11 of an $\mathcal{O}(\varepsilon\,\Delta t)$-term beside the expected $\mathcal{O}(\varepsilon^2)$-term does not deteriorate the final error estimate.

Corollary 3.4 is then deduced from Theorem 3.1 exactly as Corollary 2.4 was concluded from Theorem 2.1. For a more abstract version, see *e.g.* [14, Proposition 2.1].

Remark 3.12. *The $\mathcal{O}(\varepsilon\,\Delta t)$-term in Proposition 3.11 may be tracked down to the initial step of the scheme (3.1). As already pointed out in* [13, Remark 2.6]*, similar issues in the asymptotic error rates occur, in general, for higher-order schemes and they do impact the numerical convergence rates. Hence the need to understand how to fix this — here harmless — issue. To some extent, the problem may be cured by modifying the initial step of the scheme, e.g., with either a small time step or a fully-implicit treatment. An inspection of the proof of Proposition 3.11 shows that, here, setting the initial time step of size $(\Delta t)_0 \lesssim \varepsilon$ would be sufficient and resolves the issue. The analysis of the effect of an implicit treatment requires more work and is discussed in the next section.*

3.4. **Variant with a fully-implicit first step.** In this section, we discuss the gain in the $\boldsymbol{y}$-estimate of Proposition 3.11 obtained by modifying the first step of the scheme (3.1) into a fully-implicit version. We consider the scheme obtained by combining

$$\begin{cases} \dfrac{\boldsymbol{x}_\varepsilon^1 - \boldsymbol{x}_\varepsilon^0}{\Delta t} &= \dfrac{\boldsymbol{v}_\varepsilon^1}{\varepsilon}, \\ \varepsilon\,\dfrac{\boldsymbol{v}_\varepsilon^1 - \boldsymbol{v}_\varepsilon^0}{\Delta t} &= \mathbf{E}(t_1, \boldsymbol{x}_\varepsilon^1) \,-\, \dfrac{(\boldsymbol{v}_\varepsilon^1)^\perp}{\varepsilon}. \end{cases} \tag{3.16}$$

with (3.1) for $n \geq 1$. Accordingly, the expected asymptotic scheme is the combination of one implicit Euler step followed by explicit Euler steps, that is,

$$\begin{cases} \dfrac{\boldsymbol{x}^1 - \boldsymbol{x}^0}{\Delta t} &= -\mathbf{E}^\perp(t_1, \boldsymbol{x}^1)\,, \\ \dfrac{\boldsymbol{x}^{n+1} - \boldsymbol{x}^n}{\Delta t} &= -\mathbf{E}^\perp(t_n, \boldsymbol{x}^n)\,, \qquad \text{for } n \geq 1\,. \end{cases} \tag{3.17}$$

A rash inspection may lead to the deceptive conclusion that a time step $\Delta t = \mathcal{O}(\varepsilon)$ is required to solve (3.16) (for instance applying a Newton's method). However, this first step (3.16) can be equivalently written in terms of $(\boldsymbol{y}, \boldsymbol{v})$ as

$$(\boldsymbol{y}_\varepsilon^1, \boldsymbol{v}_\varepsilon^1) = \mathbf{F}(\boldsymbol{y}_\varepsilon^1, \boldsymbol{v}_\varepsilon^1),$$

where $\mathbf{F} = (\mathbf{F}_{\boldsymbol{y}}, \mathbf{F}_{\boldsymbol{v}})$ defined as

$$\begin{cases} \mathbf{F}_{\boldsymbol{y}}(\boldsymbol{y}, \boldsymbol{v}) &= \boldsymbol{y}_\varepsilon^0 - \Delta t\, \mathbf{E}^\perp(t_1, \boldsymbol{y} \,+\, \varepsilon\,(\boldsymbol{v})^\perp)\,, \\ \mathbf{F}_{\boldsymbol{v}}(\boldsymbol{y}, \boldsymbol{v}) &= \left(\mathrm{Id} + \dfrac{\Delta t}{\varepsilon^2}\mathbf{J}\right)^{-1} \left(\boldsymbol{v}_\varepsilon^0 + \dfrac{\Delta t}{\varepsilon}\mathbf{E}(t_1, \boldsymbol{y} \,+\, \varepsilon\,(\boldsymbol{v})^\perp)\right). \end{cases}$$

So, it can be seen that the map $\mathbf{F}$ has a Lipschitz constant which is not bigger than $2K_x\Delta t$ for the norm $\|\cdot\|_\varepsilon$ introduced in (3.9). Therefore, thanks to the contraction mapping theorem, the unique solvability of the first step (3.16) is assured for a small enough time step, explicitly when $K_x\Delta t < 1/2$. For this alternate scheme, we prove the following modification of the second asymptotic estimate of Proposition 3.11.

Proposition 3.13. *Assume $\mathbf{E} \in W^{2,\infty}$ and $K_x\Delta t < 1/2$. Then,*

$$\|\mathbf{Y}_\varepsilon^{\Delta t}(t_n, 0, \boldsymbol{x}_\varepsilon^0, \boldsymbol{v}_\varepsilon^0) \,-\, \mathbf{X}^{\Delta t}(t_n, 0, \boldsymbol{x}_\varepsilon^0 - \varepsilon(\boldsymbol{v}_\varepsilon^0)^\perp)\| \underset{\mathbf{E}}{\lesssim} \varepsilon^2\, \frac{e^{2K_x t_n}\,(1 + t_n^4)}{1 - K_x\Delta t}\,(1 + \|\boldsymbol{v}_\varepsilon^0\|^2 + \varepsilon^2)\,,$$

where $\mathbf{Y}_\varepsilon^{\Delta t} := \mathbf{X}_\varepsilon^{\Delta t} - \varepsilon(\mathbf{V}_\varepsilon^{\Delta t})^\perp$, and $\mathbf{X}^{\Delta t}$ and $(\mathbf{X}_\varepsilon^{\Delta t}, \mathbf{V}_\varepsilon^{\Delta t})$ denote, respectively, the discrete asymptotic flow (3.17) and the discrete flow obtained from (3.16)–(3.1).

Remark 3.14. *Note that in Proposition 3.13, one could, instead, impose the constraint $K_x(\Delta t)_0 < 1/2$ only on the first step.*

Proof. As in Proposition 3.11, an ε-uniform bound on the computed velocity is a prerequisite for the asymptotic estimate. The proof of such a bound is identical to the one of Lemma 3.10 except for the first step, where we need to re-define $\boldsymbol{z}_\varepsilon^1$ as

$$\boldsymbol{z}_\varepsilon^1 := \boldsymbol{v}_\varepsilon^1 \,+\, \varepsilon\, \mathbf{E}^\perp(t_1, \boldsymbol{x}_\varepsilon^1).$$

Indeed, with this modification (3.12) is unaltered and, despite the modification of the first step, Lemma 3.6 still implies

$$\|\boldsymbol{v}_\varepsilon^1\| \le \|\boldsymbol{v}_\varepsilon^0\| \,+\, \varepsilon\, K_0\,.$$

so that the conclusion of Lemma 3.10 still holds for the new scheme.

Thus, as in the proof of Proposition 3.11 and with the same r_n thereof, one gets the following estimate for all $n \ge 1$

$$\|\boldsymbol{y}_\varepsilon^n - \boldsymbol{y}^n\| \le\, e^{K_x(t_n - t_1)}\,(1 + K_x\,(t_n - t_1))\,\left(\|\boldsymbol{y}_\varepsilon^1 - \boldsymbol{y}^1\| + \varepsilon^2 \max_{2\le \ell\le n}(e^{-K_x(t_\ell - t_1)} r_\ell)\right).$$

As for the first step, with the same Θ_ε as before,

$$\frac{\boldsymbol{y}_\varepsilon^1 - \boldsymbol{y}_\varepsilon^0}{\Delta t} = -\,\mathbf{E}^\perp(t_1, \boldsymbol{y}_\varepsilon^1) \,-\, \varepsilon\, \mathrm{d}_x\mathbf{E}^\perp(t_1, \boldsymbol{y}_\varepsilon^1)(\boldsymbol{v}_\varepsilon^1)^\perp + \varepsilon^2 \Theta_\varepsilon(t_1, \boldsymbol{y}_\varepsilon^1, \boldsymbol{v}_\varepsilon^1)\,,$$

The reformulation of $\mathrm{d}_x\mathbf{E}(t_1, \boldsymbol{y}_\varepsilon^1)(\boldsymbol{v}_\varepsilon^1)^\perp$ writes

$$\begin{aligned}
\mathrm{d}_x\mathbf{E}(t_1, \boldsymbol{y}_\varepsilon^1)(\boldsymbol{v}_\varepsilon^1)^\perp \;&=\; -\,\varepsilon^2\, \mathrm{d}_x\mathbf{E}(t_1, \boldsymbol{y}_\varepsilon^1)\left(\frac{\boldsymbol{v}_\varepsilon^1 - \boldsymbol{v}_\varepsilon^0}{\Delta t} \,-\, \frac{1}{\varepsilon}\mathbf{E}(t_1, \boldsymbol{x}_\varepsilon^1)\right)\\
&=\; -\,\frac{\varepsilon^2}{\Delta t}\left(\mathrm{d}_x\mathbf{E}(t_1, \boldsymbol{y}_\varepsilon^1)\boldsymbol{v}_\varepsilon^1 - \mathrm{d}_x\mathbf{E}(t_0, \boldsymbol{y}_\varepsilon^0)\boldsymbol{v}_\varepsilon^0\right) + \varepsilon\, \mathrm{d}_{\boldsymbol{x}}\mathbf{E}(t_1, \boldsymbol{y}_\varepsilon^1)\,\mathbf{E}(t_1, \boldsymbol{x}_\varepsilon^1)\\
&\quad +\, \frac{\varepsilon^2}{\Delta t}\left(\mathrm{d}_x\mathbf{E}(t_1, \boldsymbol{y}_\varepsilon^1) - \mathrm{d}_x\mathbf{E}(t_0, \boldsymbol{y}_\varepsilon^0)\right)\,\boldsymbol{v}_\varepsilon^0\,,
\end{aligned}$$

which yields the estimate

$$\begin{aligned}
\|\boldsymbol{y}_\varepsilon^1 - \boldsymbol{y}^1\| \;\le\; \frac{\varepsilon^2}{1 - K_x\Delta t}\Bigg(&\varepsilon\, K_x\,(\|\boldsymbol{v}_\varepsilon^1\| + \|\boldsymbol{v}_\varepsilon^0\|) \,+\, \Delta t\, K_x\, K_0\\
&+\, \varepsilon\,\Delta t\,(K_{tx} + K_{xx}\, K_0)\,\|\boldsymbol{v}_\varepsilon^0\| \,+\, \Delta t\,\frac{K_{xx}}{2}\,\|\boldsymbol{v}_\varepsilon^1\|^2\Bigg)\,.
\end{aligned}$$

From here the proof is completed as the one in Proposition 3.11 by collecting all the foregoing estimates. □

It is essential to emphasize that numerical schemes relying on a distinct treatment of the initial step are not really desirable as they are not well-adapted to the case when the magnetic field is not uniformly strong. Namely, such schemes are inefficient when effectively the scaling parameter starts of size $\varepsilon \sim 1$ and later gets small, $\varepsilon \ll 1$. For this reason, we will consider a different type of remedy.

4. First-order scheme on the guiding center variable

We propose in this section a modification of the scheme (3.1), which is simply based on the approximation of $(\boldsymbol{y}_\varepsilon, \boldsymbol{v}_\varepsilon)$ instead of $(\boldsymbol{x}_\varepsilon, \boldsymbol{v}_\varepsilon)$. We define a discrete time evolution by a first-order IMEX method for $n \ge 0$,

$$\left\{\begin{aligned}
\frac{\boldsymbol{v}_\varepsilon^{n+1} - \boldsymbol{v}_\varepsilon^n}{\Delta t} \;&=\; \frac{1}{\varepsilon}\,\mathbf{E}(t_n, \boldsymbol{y}_\varepsilon^n + \varepsilon(\boldsymbol{v}_\varepsilon^n)^\perp)) \,-\, \frac{1}{\varepsilon^2}(\boldsymbol{v}_\varepsilon^{n+1})^\perp,\\
\frac{\boldsymbol{y}_\varepsilon^{n+1} - \boldsymbol{y}_\varepsilon^n}{\Delta t} \;&=\; -\mathbf{E}^\perp\left(t_n, \boldsymbol{y}_\varepsilon^n + \varepsilon(\boldsymbol{v}_\varepsilon^{n+1})^\perp\right).
\end{aligned}\right. \tag{4.1}$$

The aim is the improvement of the asymptotic error, $\boldsymbol{y}_\varepsilon^n - \boldsymbol{y}^n$ where $(\boldsymbol{y}^n)_n$ solves (3.3), the explicit Euler scheme for the guiding center equation (2.1), with initial datum $\boldsymbol{y}^0 = \boldsymbol{y}_\varepsilon^0$. Note that the $\boldsymbol{y}_\varepsilon$-update in (4.1) differs from (3.2), only by the presence of $\boldsymbol{v}_\varepsilon^{n+1}$ instead of $\boldsymbol{v}_\varepsilon^n$. We will see (in Proposition 4.4 below) that this simple difference will remove the unwanted term of Proposition 3.11.

In terms of the spatial variable $\boldsymbol{x}_\varepsilon^n := \boldsymbol{y}_\varepsilon^n + \varepsilon(\boldsymbol{v}_\varepsilon^n)^\perp$, for $n \geq 0$, the scheme (4.1) is equivalently written as

$$\text{(4.2)} \qquad \begin{cases} \dfrac{\boldsymbol{v}_\varepsilon^{n+1} - \boldsymbol{v}_\varepsilon^n}{\Delta t} &= \dfrac{1}{\varepsilon}\,\mathbf{E}(t_n, \boldsymbol{x}_\varepsilon^n) \,-\, \dfrac{1}{\varepsilon^2}(\boldsymbol{v}_\varepsilon^{n+1})^\perp, \\ \dfrac{\boldsymbol{x}_\varepsilon^{n+1} - \boldsymbol{x}_\varepsilon^n}{\Delta t} &= \dfrac{\boldsymbol{v}_\varepsilon^{n+1}}{\varepsilon} + \left[\mathbf{E}^\perp(t_n, \boldsymbol{x}_\varepsilon^n) - \mathbf{E}^\perp\left(t_n, \boldsymbol{x}_\varepsilon^n + \varepsilon\left(\boldsymbol{v}_\varepsilon^{n+1} - \boldsymbol{v}_\varepsilon^n\right)^\perp\right)\right]. \end{cases}$$

Compared to (3.1), the modification to $(\boldsymbol{x}_\varepsilon^n)_{n\geq 1}$ is expected to be $\mathcal{O}(\varepsilon\,\Delta t)$, hence immaterial for asymptotic and numerical convergences of the variable $\boldsymbol{x}_\varepsilon^n$. The following theorem provides the main error estimate concerning the modified scheme (4.1).

Theorem 4.1. *The first-order scheme (4.1) possesses a unique solution. Moreover*

(i) when $\mathbf{E} \in W^{1,\infty}$, the space variable $\boldsymbol{x}_\varepsilon$ satisfies for all $n \geq 0$, $\Delta t > 0$ and $\varepsilon > 0$,

$$\|\boldsymbol{x}_\varepsilon^n - \boldsymbol{x}_\varepsilon(t_n)\| \underset{\mathbf{E}}{\lesssim} (1 + t_n^3)\, e^{(2+K_x\Delta t)K_x t_n}(1 + \|\boldsymbol{v}_\varepsilon^0\| + \varepsilon) \times \min\left(\frac{\Delta t}{\varepsilon^3}(1+\varepsilon^2), \Delta t + \varepsilon\right),$$

(ii) when $\mathbf{E} \in W^{2,\infty}$, the guiding center variable $\boldsymbol{y}_\varepsilon$ satisfies for all $n \geq 0$, $\Delta t > 0$ and $\varepsilon > 0$,

$$\|\boldsymbol{y}_\varepsilon^n - \boldsymbol{y}_\varepsilon(t_n)\| \underset{\mathbf{E}}{\lesssim} (1 + t_n^4)\, e^{(2+K_x\Delta t)K_x t_n}(1 + \|\boldsymbol{v}_\varepsilon^0\|^2 + \varepsilon^2) \times \min\left(\frac{\Delta t}{\varepsilon^3}\,(1+\varepsilon)\,, \Delta t + \varepsilon^2\right).$$

Our strategy to prove Theorem 4.1 follows closely the one to prove Theorem 3.1 in §3.

4.1. **Direct convergence estimates.** To perform a direct numerical convergence analysis, we, first, introduce corresponding consistency errors,

$$\text{(4.3)} \qquad \boldsymbol{\tau}_{\boldsymbol{y}}^n := \frac{\boldsymbol{y}_\varepsilon(t_{n+1}) - \boldsymbol{y}_\varepsilon(t_n)}{\Delta t} + \mathbf{E}^\perp(t_n, \boldsymbol{y}_\varepsilon(t_n) + \varepsilon\,\boldsymbol{v}_\varepsilon^\perp(t_{n+1}))\,,$$

$$\text{(4.4)} \qquad \boldsymbol{\tau}_{\boldsymbol{v}}^n := \frac{\boldsymbol{v}_\varepsilon(t_{n+1}) - \boldsymbol{v}_\varepsilon(t_n)}{\Delta t} - \frac{\mathbf{E}(t_n, \boldsymbol{y}_\varepsilon(t_n) + \varepsilon\,\boldsymbol{v}_\varepsilon^\perp(t_n))}{\varepsilon} + \frac{\boldsymbol{v}_\varepsilon^\perp(t_{n+1})}{\varepsilon^2}\,,$$

which can be estimated as in Lemma 3.5. Note that $\boldsymbol{\tau}_{\boldsymbol{v}}^n$ in (4.3) is identical to (3.6b) whereas $\boldsymbol{\tau}_{\boldsymbol{y}}^n$ differs from (3.6a) by the presence of $\boldsymbol{v}_\varepsilon^{n+1}$ instead of $\boldsymbol{v}_\varepsilon^n$. The corresponding modification to the proof of Lemma 3.5 leads to a statement essentially identical to Lemma 3.5, hence omitted here. From this we derive the following proposition.

Proposition 4.2. *Assume that $\mathbf{E} \in W^{2,\infty}$. The error from the unique solution of the scheme (4.1) to the exact solution of the system (1.7) is bounded as*

$$\text{(4.5)} \qquad \begin{aligned} \|\boldsymbol{x}_\varepsilon^n - \boldsymbol{x}_\varepsilon(t_n)\| &\leq \|\boldsymbol{y}_\varepsilon^n - \boldsymbol{y}_\varepsilon(t_n)\| + \varepsilon\|\boldsymbol{v}_\varepsilon^n - \boldsymbol{v}_\varepsilon(t_n)\| \\ &\underset{\mathbf{E}}{\lesssim} \frac{\Delta t}{\varepsilon^3}\, t_n\, e^{2K_x t_n}(1+\varepsilon^2)(1+\Delta t)\Big(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,(1+t_n)\Big)\,. \end{aligned}$$

Proof. As in the proof of Proposition 3.7, we consider numerical errors

$$\boldsymbol{e}_{\boldsymbol{y}}^n := \boldsymbol{y}_\varepsilon^n - \boldsymbol{y}_\varepsilon(t_n)\,, \qquad \boldsymbol{e}_{\boldsymbol{v}}^n := \boldsymbol{v}_\varepsilon^n - \boldsymbol{v}_\varepsilon(t_n)\,,$$

and use the norm $\|\cdot\|_\varepsilon$ introduced in (3.9). By reformulating the second equation of (4.1) and using Lemma 3.6, we obtain

$$\|\boldsymbol{e}_{\boldsymbol{y}}^{n+1}\| \leq \|\boldsymbol{e}_{\boldsymbol{y}}^n\| + K_x\Delta t\,(\|\boldsymbol{e}_{\boldsymbol{y}}^n\| + \varepsilon\|\boldsymbol{e}_{\boldsymbol{v}}^{n+1})\|) + \Delta t\,\|\boldsymbol{\tau}_{\boldsymbol{y}}^n\|\,,$$

$$\varepsilon\|\boldsymbol{e}_{\boldsymbol{v}}^{n+1}\| \leq \varepsilon\|\boldsymbol{e}_{\boldsymbol{v}}^n\| + K_x\Delta t\,\|(\boldsymbol{e}_{\boldsymbol{y}}^n, \boldsymbol{e}_{\boldsymbol{v}}^n)\|_\varepsilon + \Delta t\,\varepsilon\,\|\boldsymbol{\tau}_{\boldsymbol{v}}^n\|\,,$$

that may be combined to yield

$$\|(\boldsymbol{e}_{\boldsymbol{y}}^{n+1}, \boldsymbol{e}_{\boldsymbol{v}}^{n+1})\|_\varepsilon \,\leq\, (1+K_x\Delta t)^2 \|(\boldsymbol{e}_{\boldsymbol{y}}^{n}, \boldsymbol{e}_{\boldsymbol{v}}^{n})\|_\varepsilon + \Delta t\,(1+K_x\Delta t)\,\|(\boldsymbol{\tau}_{\boldsymbol{y}}^{n}, \boldsymbol{\tau}_{\boldsymbol{v}}^{n})\|_\varepsilon\,.$$

Iterating and applying our new version of Lemma 3.5 conclude the proof. Let us observe that when doing so we use a slightly different form of the discrete Grönwall where $(1+K_x\Delta t)^2$ leads to the same exponential factor as $(1+2K_x\Delta t)$ was, both being bounded by $e^{2K_x\Delta t}$. □

4.2. **Asymptotic estimates.** To prove asymptotic estimates, we first obtain uniform bounds on solutions to (4.1). To do so, as in §3.2, we consider a discrete analogue of the auxiliary variable $\boldsymbol{z}_\varepsilon$, denoted by $(\boldsymbol{z}_\varepsilon^n)_{n\geq 0}$, defined exactly as in (3.11) and also satisfying (3.12).

Lemma 4.3. *Assume that* $\mathbf{E}\in W^{1,\infty}$ *and consider* (4.1). *Then, the variable* $(\boldsymbol{z}_\varepsilon^n)_{n\geq 0}$ *defined in* (3.11) *and the velocity computed by the scheme are bounded, for* $n\geq 1$, *as*

$$\begin{cases} \|\boldsymbol{z}_\varepsilon^n\| \,\leq\, e^{K_x\,t_n}\big(\|\boldsymbol{v}_\varepsilon^0\| + 2\varepsilon\,K_0\big) + \varepsilon\,t_n\,e^{K_x\,t_n}\,(K_t \,+\, K_x\,K_0)\,, \\ \|\boldsymbol{v}_\varepsilon^n\| \,\leq\, \|\boldsymbol{z}_\varepsilon^n\| \,+\, K_0\,\varepsilon\,. \end{cases}$$

Proof. The proof goes along the same lines as the proof of Lemma 3.10. The only difference is that now one has directly, for $n\geq 1$,

$$\|\mathbf{E}^\perp(t_n,\boldsymbol{x}_\varepsilon^n) - \mathbf{E}^\perp(t_{n-1},\boldsymbol{x}_\varepsilon^{n-1})\| \leq \left(K_t + K_x\left(\frac{\|\boldsymbol{z}_\varepsilon^n\|}{\varepsilon} + K_0\right)\right)\Delta t\,,$$

from the $\boldsymbol{x}_\varepsilon$-update

$$\frac{\boldsymbol{x}_\varepsilon^n - \boldsymbol{x}_\varepsilon^{n-1}}{\Delta t} = \frac{\boldsymbol{z}_\varepsilon^n}{\varepsilon} - \mathbf{E}^\perp\left(t_{n-1}, \boldsymbol{x}_\varepsilon^{n-1} + \varepsilon\left(\boldsymbol{v}_\varepsilon^n - \boldsymbol{v}_\varepsilon^{n-1}\right)^\perp\right)\,.$$

□

To state asymptotic estimates, we denote the discrete flow for the scheme (4.1) by $(\mathbf{X}_\varepsilon^{\Delta t}, \mathbf{V}_\varepsilon^{\Delta t})$, and the one for (3.3) by $\mathbf{X}^{\Delta t}$. As before we also define $\mathbf{Y}_\varepsilon^{\Delta t} := \mathbf{X}_\varepsilon^{\Delta t} - \varepsilon(\mathbf{V}_\varepsilon^{\Delta t})^\perp$.

Proposition 4.4. *(i) Assume* $\mathbf{E}\in W^{1,\infty}$. *Then,*

$$\|\mathbf{X}_\varepsilon^{\Delta t}(t_n,0,\boldsymbol{x}_\varepsilon^0,\boldsymbol{v}_\varepsilon^0) - \mathbf{X}^{\Delta t}(t_n,0,\boldsymbol{x}_\varepsilon^0)\| \underset{\mathbf{E}}{\lesssim} \varepsilon\,e^{K_x t_n}\,(1+\,t_n^3)\,\left(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\right)\,.$$

(ii) Assume $\mathbf{E}\in W^{2,\infty}$. *Then,*

$$\|\mathbf{Y}_\varepsilon^{\Delta t}(t_n,0,\boldsymbol{x}_\varepsilon^0,\boldsymbol{v}_\varepsilon^0) \,-\, \mathbf{X}^{\Delta t}(t_n,0,\boldsymbol{x}_\varepsilon^0 - \varepsilon(\boldsymbol{v}_\varepsilon^0)^\perp)\| \underset{\mathbf{E}}{\lesssim} \varepsilon^2\,e^{2K_x t_n}\,(1+t_n^4)\,\left(1+\|\boldsymbol{v}_\varepsilon^0\|^2+\varepsilon^2\right).$$

Proof. To obtain the first estimate, we subtract the $\boldsymbol{y}_\varepsilon$-update in (4.1) from the limit scheme (3.3), and make a summation from 0 to $(n-1)$ to obtain, for $n\geq 0$,

$$\boldsymbol{y}_\varepsilon^n - \boldsymbol{x}^n = \boldsymbol{y}_\varepsilon^0 - \boldsymbol{x}^0 \,-\, \Delta t\sum_{\ell=0}^{n-1}\left(\mathbf{E}^\perp\left(t_\ell, \boldsymbol{x}_\varepsilon^\ell + \varepsilon\left(\boldsymbol{v}_\varepsilon^{\ell+1} - \boldsymbol{v}_\varepsilon^\ell\right)^\perp\right) - \mathbf{E}^\perp(t_\ell,\boldsymbol{x}^\ell)\right)\,,$$

which, thanks to the Lipschitz continuity of the electric field, implies

$$\|\boldsymbol{x}_\varepsilon^n - \boldsymbol{x}^n\| \,\leq\, \varepsilon\,(\|\boldsymbol{v}_\varepsilon^n\| + \|\boldsymbol{v}_\varepsilon^0\|) + K_x\,\varepsilon\Delta t\sum_{\ell=0}^{n-1}(\|\boldsymbol{v}_\varepsilon^\ell\| + \|\boldsymbol{v}_\varepsilon^{\ell+1}\|) + K_x\Delta t\sum_{\ell=0}^{n-1}\|\boldsymbol{x}_\varepsilon^\ell - \boldsymbol{x}^\ell\|\,.$$

Applying the discrete Grönwall lemma, this leads to

$$\|\boldsymbol{x}_\varepsilon^n - \boldsymbol{x}^n\| \underset{\mathbf{E}}{\lesssim} \varepsilon\,e^{K_x t_n}\,(1+t_n^2)\,\left(\|\boldsymbol{v}_\varepsilon^0\| + \max_{0\leq\ell\leq n}(e^{-K_x t_\ell}\|\boldsymbol{v}_\varepsilon^\ell\|)\right), \qquad n\geq 0\,.$$

Finally, applying Lemma 4.3 concludes the proof of the first estimate.

The proof of the second estimate starts with a Taylor expansion of the $\boldsymbol{y}_\varepsilon$-update,

$$\frac{\boldsymbol{y}_\varepsilon^{n+1} - \boldsymbol{y}_\varepsilon^n}{\Delta t} = -\,\mathbf{E}^\perp(t_n,\boldsymbol{y}_\varepsilon^n) \,-\, \varepsilon\,\mathrm{d}_x\mathbf{E}^\perp(t_n,\boldsymbol{y}_\varepsilon^n)(\boldsymbol{v}_\varepsilon^{n+1})^\perp + \varepsilon^2\Theta_\varepsilon(t_n,\boldsymbol{y}_\varepsilon^n,\boldsymbol{v}_\varepsilon^{n+1})\,, \qquad n\geq 0\,,$$

with Θ_ε such that $\|\Theta_\varepsilon(t,\boldsymbol{y},\boldsymbol{v})\| \le \frac12 K_{xx}\|\boldsymbol{v}\|^2$. Now, so as to rewrite the linear term, we observe that, when $n \ge 0$,

$$\begin{aligned}
\mathrm{d}_x\mathbf{E}(t_n,\boldsymbol{y}_\varepsilon^n)(\boldsymbol{v}_\varepsilon^{n+1})^\perp &= -\,\varepsilon^2\,\mathrm{d}_x\mathbf{E}(t_n,\boldsymbol{y}_\varepsilon^n)\left(\frac{\boldsymbol{v}_\varepsilon^{n+1}-\boldsymbol{v}_\varepsilon^n}{\Delta t} - \frac{1}{\varepsilon}\mathbf{E}(t_n,\boldsymbol{x}_\varepsilon^n)\right)\\
&= -\,\frac{\varepsilon^2}{\Delta t}\left(\mathrm{d}_x\mathbf{E}(t_{n+1},\boldsymbol{y}_\varepsilon^{n+1})\boldsymbol{v}_\varepsilon^{n+1} - \mathrm{d}_x\mathbf{E}(t_n,\boldsymbol{y}_\varepsilon^n)\boldsymbol{v}_\varepsilon^n\right) + \varepsilon\,\mathrm{d}_{\boldsymbol{x}}\mathbf{E}(t_n,\boldsymbol{y}_\varepsilon^n)\,\mathbf{E}(t_n,\boldsymbol{x}_\varepsilon^n)\\
&\quad + \frac{\varepsilon^2}{\Delta t}\left(\mathrm{d}_x\mathbf{E}(t_{n+1},\boldsymbol{y}_\varepsilon^{n+1}) - \mathrm{d}_x\mathbf{E}(t_n,\boldsymbol{y}_\varepsilon^n)\right)\,\boldsymbol{v}_\varepsilon^n,
\end{aligned}$$

whose last term is bounded as

$$\|\mathrm{d}_x\mathbf{E}(t_{n+1},\boldsymbol{y}_\varepsilon^{n+1}) - \mathrm{d}_x\mathbf{E}(t_n,\boldsymbol{y}_\varepsilon^n)\| \le \Delta t\,\left(K_{tx} + K_{xx}\,K_0\right)\,.$$

Then, with $\boldsymbol{y}^n := \mathbf{X}^{\Delta t}(t_n,0,\boldsymbol{x}_\varepsilon^0 - \varepsilon(\boldsymbol{v}_\varepsilon^0)^\perp)$ for $n \ge 0$, and arguing as in the proof of Proposition 3.11, we deduce

$$\|\boldsymbol{y}_\varepsilon^n - \boldsymbol{y}^n\| \le\ \varepsilon^2\,e^{K_x t_n}\,(1+K_x\,t_n)\,\max_{1\le\ell\le n}\,(e^{-K_x t_\ell} r_\ell)\,,\qquad n \ge 0\,,$$

where, for $n \ge 0$,

$$r_n := \varepsilon\,K_x\,(\|\boldsymbol{v}_\varepsilon^n\| + \|\boldsymbol{v}_\varepsilon^0\|)\ +\ K_x\,K_0\,t^n + \varepsilon\,\Delta t\,(K_{tx}+K_{xx}\,K_0)\sum_{\ell=0}^{n-1}\|\boldsymbol{v}_\varepsilon^{\ell-1}\|\ +\ \Delta t\,\frac{K_{xx}}{2}\sum_{\ell=1}^{n}\|\boldsymbol{v}_\varepsilon^\ell\|^2\,.$$

The proof is concluded with Lemma 4.3. □

4.3. **Proof of Theorem 4.1.** The proof of Theorem 4.1 goes along the same lines as of the proof of Theorem 3.1. It consists, on the one hand, in applying Proposition 4.2 and, on the other hand, in combining Theorem 2.1 and Propositions 3.9 and 4.4.

5. L-stable second-order implicit-explicit scheme

In this section, we discuss the uniform convergence analysis of a second-order extension of the scheme presented in §4. The analysis we are going to perform is conceptually similar to the ones in §3 and §4 but technically much more involved.

For concision's sake we restrict to the analysis of a single scheme, a scheme written on the guiding center variable. However, a similar analysis could be performed on the second-order version of the scheme of §3 (*cf.* [12]). We stress that here also the deterioration of asymptotic estimates on the guiding variable $\boldsymbol{y}$ by an $\mathcal{O}(\varepsilon\,\Delta t)$-term has no impact on numerical convergence errors since $\varepsilon\,\Delta t \lesssim \varepsilon^2 + \Delta t^2$.

The semi-implicit second-order method we consider is a combination of a Runge–Kutta method (for the explicit part) and an L-stable second-order SDIRK method (for the implicit part), with the parameter γ chosen as the smallest root of the polynomial $\gamma^2 - 2\gamma + 1/2 = 0$, that is, $\gamma = 1 - \frac{1}{\sqrt{2}}$; see [2].

To shed some light on the structure of the scheme, we write the characteristic system (1.7) in the slightly more abstract form,

$$\begin{cases}
\boldsymbol{y}_\varepsilon'(t) \ =\ \mathbf{F}_{\boldsymbol{y}}(t,\boldsymbol{y}_\varepsilon(t),\varepsilon\,\boldsymbol{v}_\varepsilon(t))\,,\\[2mm]
\varepsilon\,\boldsymbol{v}_\varepsilon'(t) \ =\ \mathbf{F}_{\boldsymbol{v}}^\varepsilon(t,\boldsymbol{y}_\varepsilon(t),\varepsilon\,\boldsymbol{v}_\varepsilon(t),\varepsilon\,\boldsymbol{v}_\varepsilon(t))\,,\\[2mm]
\boldsymbol{y}_\varepsilon(s) \ =\ \boldsymbol{y}_\varepsilon^s,\qquad \boldsymbol{v}_\varepsilon(s) \ =\ \boldsymbol{v}_\varepsilon^s,
\end{cases}\tag{5.1}$$

with

$$\mathbf{F}_{\boldsymbol{y}}(t,\widehat{\boldsymbol{y}},\widetilde{\boldsymbol{w}}) := -\mathbf{E}^\perp(t,\widehat{\boldsymbol{y}}+\widetilde{\boldsymbol{w}}^\perp)\,,\qquad\qquad \mathbf{F}_{\boldsymbol{v}}^\varepsilon(t,\widehat{\boldsymbol{y}},\widehat{\boldsymbol{w}},\widetilde{\boldsymbol{w}}) := \mathbf{E}(t,\widehat{\boldsymbol{y}}+\widehat{\boldsymbol{w}}^\perp) - \frac{\widetilde{\boldsymbol{w}}^\perp}{\varepsilon^2}\,,$$

where a variable with a tilde $\widetilde{b}$ is used in stiff parts of the system while a hatted variables $\widehat{b}$ are to be used in non-stiff parts. The identification of stiff and non-stiff parts in (5.1) prepares duplication

at discrete level of the velocity variable $\boldsymbol{v}$ as $\widetilde{\boldsymbol{v}}$ and $\widehat{\boldsymbol{v}}$ to be treated, respectively, implicitly and explicitly. Such a duplication is essential to obtain a semi-implicit scheme that avoids nonlinear iterations; see [2].

The first stage of the scheme is a linearly-implicit update which provides an approximation of the velocity after a time step of size $\gamma\Delta t$,

$$\text{(5.2a)}\qquad \begin{cases} \widetilde{t}_{n+1} \,:=\, t_n \,+\, \gamma\Delta t\,, \\ \varepsilon\dfrac{\widetilde{\boldsymbol{v}}_\varepsilon^{n+1}-\boldsymbol{v}_\varepsilon^n}{\gamma\Delta t} \,=\, \mathbf{F}_{\boldsymbol{v}}^\varepsilon(t_n,\boldsymbol{y}_\varepsilon^n,\varepsilon\,\boldsymbol{v}_\varepsilon^n,\varepsilon\,\widetilde{\boldsymbol{v}}_\varepsilon^{n+1})\,, \end{cases}$$

where $\widetilde{\boldsymbol{v}}_\varepsilon^{n+1}$ approximates $\boldsymbol{v}_\varepsilon(\widetilde{t}_{n+1})$, and is to be used in stiff parts.

Then, the second stage provides an explicit approximation $(\widehat{\boldsymbol{y}}_\varepsilon^{n+1},\widehat{\boldsymbol{v}}_\varepsilon^{n+1})$ of $(\boldsymbol{y}_\varepsilon,\boldsymbol{v}_\varepsilon)(\widehat{t}_{n+1})$ to be used in non stiff parts. It reads, for $n\ge 0$,

$$\text{(5.2b)}\qquad \begin{cases} \widehat{t}_{n+1} \,:=\, t_n \,+\, \dfrac{\Delta t}{2\gamma}\,, \\ \dfrac{\widehat{\boldsymbol{y}}_\varepsilon^{n+1}-\boldsymbol{y}_\varepsilon^n}{\Delta t/(2\gamma)} \,=\, \mathbf{F}_{\boldsymbol{y}}(t_n,\boldsymbol{y}_\varepsilon^n,\varepsilon\,\widetilde{\boldsymbol{v}}_\varepsilon^{n+1})\,, \\ \varepsilon\dfrac{\widehat{\boldsymbol{v}}_\varepsilon^{n+1}-\boldsymbol{v}_\varepsilon^n}{\Delta t/(2\gamma)} \,=\, \mathbf{F}_{\boldsymbol{v}}^\varepsilon(t_n,\boldsymbol{y}_\varepsilon^n,\varepsilon\,\boldsymbol{v}_\varepsilon^n,\varepsilon\,\widetilde{\boldsymbol{v}}_\varepsilon^{n+1})\,. \end{cases}$$

The last stage, which provides the final update, is linearly-implicit and writes

$$\text{(5.2c)}\qquad \begin{cases} \dfrac{\boldsymbol{y}_\varepsilon^{n+1}-\boldsymbol{y}_\varepsilon^n}{\Delta t} &=\, (1-\gamma)\,\mathbf{F}_{\boldsymbol{y}}(t_n,\boldsymbol{y}_\varepsilon^n,\varepsilon\,\widetilde{\boldsymbol{v}}_\varepsilon^{n+1})+\gamma\,\mathbf{F}_{\boldsymbol{y}}(\widehat{t}_{n+1},\widehat{\boldsymbol{y}}_\varepsilon^{n+1},\varepsilon\,\boldsymbol{v}_\varepsilon^{n+1})\,, \\ \varepsilon\dfrac{\boldsymbol{v}_\varepsilon^{n+1}-\boldsymbol{v}_\varepsilon^n}{\Delta t} &=\, (1-\gamma)\,\mathbf{F}_{\boldsymbol{v}}^\varepsilon(t_n,\boldsymbol{y}_\varepsilon^n,\varepsilon\,\boldsymbol{v}_\varepsilon^n,\varepsilon\,\widetilde{\boldsymbol{v}}_\varepsilon^{n+1})+\gamma\,\mathbf{F}_{\boldsymbol{v}}^\varepsilon(\widehat{t}_{n+1},\widehat{\boldsymbol{y}}_\varepsilon^{n+1},\varepsilon\,\widehat{\boldsymbol{v}}_\varepsilon^{n+1},\varepsilon\,\boldsymbol{v}_\varepsilon^{n+1})\,. \end{cases}$$

Since $\mathbf{F}_{\boldsymbol{y}}(t,\boldsymbol{y},0)=-\mathbf{E}^\perp(t,\boldsymbol{y})$, taking formally the limit $\varepsilon\to 0$ suggests that the above discretization tends to the discretization of the guiding center equation by a second-order fully explicit Runge–Kutta scheme, where the first stage is

$$\text{(5.3a)}\qquad \begin{cases} \widehat{t}_{n+1} \,:=\, t_n \,+\, \dfrac{\Delta t}{2\gamma}\,, \\ \dfrac{\widehat{\boldsymbol{x}}^{n+1}-\boldsymbol{x}^n}{\Delta t/(2\gamma)} \,=\, -\mathbf{E}^\perp(t_n,\boldsymbol{x}^n)\,, \end{cases}$$

and the second one reads

$$\text{(5.3b)}\qquad \frac{\boldsymbol{x}^{n+1}-\boldsymbol{x}^n}{\Delta t} \,=\, -(1-\gamma)\,\mathbf{E}^\perp(t_n,\boldsymbol{x}^n)-\gamma\,\mathbf{E}^\perp(\widehat{t}_{n+1},\widehat{\boldsymbol{x}}^{n+1})\,.$$

On this second-order scheme (5.2), our main result is the following theorem.

Theorem 5.1. *The second-order scheme (5.2) possesses a unique solution. Moreover*

(i) there exists C_0 such that when $\mathbf{E}\in W^{2,\infty}$, the space variable $\boldsymbol{x}_\varepsilon$ satisfies for all $n\ge 0$, $\Delta t>0$ and $\varepsilon>0$,

$$\|\boldsymbol{x}_\varepsilon^n-\boldsymbol{x}_\varepsilon(t_n)\| \underset{\mathbf{E}}{\lesssim} e^{C_0\,K_x\,t_n}\,(1+t_n)^3\,(1+\varepsilon^2+\|\boldsymbol{v}_\varepsilon^0\|^2)\,\min\left(\frac{\Delta t^2}{\varepsilon^5}(1+\varepsilon^3),\varepsilon+\Delta t^2\right),$$

(ii) there exists C_0 such that when $\mathbf{E}\in W^{2,\infty}$, the guiding center variable $\boldsymbol{y}_\varepsilon$ satisfies for all $n\ge 0$, $\Delta t>0$ and $\varepsilon>0$,

$$\|\boldsymbol{y}_\varepsilon^n-\boldsymbol{y}_\varepsilon(t_n)\| \underset{\mathbf{E}}{\lesssim} e^{C_0\,K_x\,t_n}\,(1+t_n)^4\,(1+\varepsilon^2+\|\boldsymbol{v}_\varepsilon^0\|^2)\,\min\left(\frac{\Delta t^2}{\varepsilon^5}(1+\varepsilon^3),\varepsilon^2+\Delta t^2\right).$$

The proof of this theorem follows the same strategy as the one of Theorems 3.1 and 4.1.

5.1. **Direct convergence estimates.** To carry out a direct convergence analysis, we introduce consistency errors for each stage of the scheme. We define for the first stage (5.2a)

$$\widetilde{\boldsymbol{\tau}}_{\boldsymbol{v}}^n := \frac{\boldsymbol{v}_\varepsilon(\widetilde{t}_{n+1}) - \boldsymbol{v}_\varepsilon(t_n)}{\gamma\Delta t} - \frac{\mathbf{E}(t_n, \boldsymbol{y}_\varepsilon(t_n) + \varepsilon\, \boldsymbol{v}_\varepsilon^\perp(t_n))}{\varepsilon} + \frac{\boldsymbol{v}_\varepsilon^\perp(\widetilde{t}_{n+1})}{\varepsilon^2}, \tag{5.4}$$

for the intermediate stage (5.2b)

$$\begin{cases} \widehat{\boldsymbol{\tau}}_{\boldsymbol{y}}^n := \dfrac{\boldsymbol{y}_\varepsilon(\widehat{t}_{n+1}) - \boldsymbol{y}_\varepsilon(t_n)}{\Delta t/(2\gamma)} + \mathbf{E}^\perp(t_n, \boldsymbol{y}_\varepsilon(t_n) + \varepsilon\, \boldsymbol{v}_\varepsilon^\perp(\widetilde{t}_{n+1})), \\ \widehat{\boldsymbol{\tau}}_{\boldsymbol{v}}^n := \dfrac{\boldsymbol{v}_\varepsilon(\widehat{t}_{n+1}) - \boldsymbol{v}_\varepsilon(t_n)}{\Delta t/(2\gamma)} - \dfrac{\mathbf{E}(t_n, \boldsymbol{y}_\varepsilon(t_n) + \varepsilon\, \boldsymbol{v}_\varepsilon^\perp(t_n))}{\varepsilon} + \dfrac{\boldsymbol{v}_\varepsilon^\perp(\widetilde{t}_{n+1})}{\varepsilon^2}, \end{cases} \tag{5.5}$$

and for the final stage (5.2c)

$$\begin{cases} \boldsymbol{\tau}_{\boldsymbol{y}}^n := \dfrac{\boldsymbol{y}_\varepsilon(t_{n+1}) - \boldsymbol{y}_\varepsilon(t_n)}{\Delta t} + (1-\gamma)\mathbf{E}^\perp(t_n, \boldsymbol{y}_\varepsilon(t_n) + \varepsilon\, \boldsymbol{v}_\varepsilon^\perp(\widetilde{t}_{n+1})) \\ \qquad\qquad + \gamma\mathbf{E}^\perp(\widehat{t}_{n+1}, \boldsymbol{y}_\varepsilon(\widehat{t}_{n+1}) + \varepsilon\, \boldsymbol{v}_\varepsilon^\perp(t_{n+1})), \\ \boldsymbol{\tau}_{\boldsymbol{v}}^n := \dfrac{\boldsymbol{v}_\varepsilon(t_{n+1}) - \boldsymbol{v}_\varepsilon(t_n)}{\Delta t} - (1-\gamma)\dfrac{\mathbf{E}(t_n, \boldsymbol{y}_\varepsilon(t_n) + \varepsilon\, \boldsymbol{v}_\varepsilon^\perp(t_n))}{\varepsilon} + (1-\gamma)\dfrac{\boldsymbol{v}_\varepsilon^\perp(\widetilde{t}_{n+1})}{\varepsilon^2} \\ \qquad\qquad - \gamma\dfrac{\mathbf{E}(\widehat{t}_{n+1}, \boldsymbol{y}_\varepsilon(\widehat{t}_{n+1}) + \varepsilon\, \boldsymbol{v}_\varepsilon^\perp(\widehat{t}_{n+1}))}{\varepsilon} + \gamma\dfrac{\boldsymbol{v}_\varepsilon^\perp(t_{n+1})}{\varepsilon^2}. \end{cases} \tag{5.6}$$

The following lemma provides bounds for these local truncation errors.

Lemma 5.2. *Assume $\mathbf{E} \in W^{2,\infty}$. Then, for any $n \geq 0$, $\varepsilon > 0$ and $\Delta t > 0$, the consistency errors of stages (5.2a)–(5.2b) satisfy*

$$\begin{cases} \|\widetilde{\boldsymbol{\tau}}_{\boldsymbol{v}}^n\| \underset{\mathbf{E}}{\lesssim} \dfrac{\Delta t}{\varepsilon^4}\, e^{K_x \widetilde{t}_{n+1}}\, (1+\varepsilon^2) \left(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,(1+\widetilde{t}_{n+1})\right), \\ \|\widehat{\boldsymbol{\tau}}_{\boldsymbol{y}}^n\| \underset{\mathbf{E}}{\lesssim} \dfrac{\Delta t}{\varepsilon}\, e^{K_x \widehat{t}_{n+1}}\, \left(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,(1+\widehat{t}_{n+1})\right), \\ \|\widehat{\boldsymbol{\tau}}_{\boldsymbol{v}}^n\| \underset{\mathbf{E}}{\lesssim} \dfrac{\Delta t}{\varepsilon^4}\, e^{K_x \widehat{t}_{n+1}}\, (1+\varepsilon^2) \left(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,(1+\widehat{t}_{n+1})\right), \end{cases}$$

and the consistency errors of the final stage (5.2c) satisfy

$$\begin{cases} \|\boldsymbol{\tau}_{\boldsymbol{y}}^n\| \underset{\mathbf{E}}{\lesssim} \dfrac{(\Delta t)^2}{\varepsilon^3} e^{K_x \widehat{t}_{n+1}} \left(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,(1+\widehat{t}_{n+1})\right) \left(1 + \varepsilon\, e^{K_x \widehat{t}_{n+1}} \left(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,(1+\widehat{t}_{n+1})\right)\right), \\ \|\boldsymbol{\tau}_{\boldsymbol{v}}^n\| \underset{\mathbf{E}}{\lesssim} \dfrac{(\Delta t)^2}{\varepsilon^6} e^{K_x \widehat{t}_{n+1}} \left(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,(1+\widehat{t}_{n+1})\right) \left(1 + \varepsilon^3\, e^{K_x \widehat{t}_{n+1}}\, \left(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,(1+\widehat{t}_{n+1})\right)\right). \end{cases}$$

Proof. We skip the proofs of the estimates on $(\widetilde{\boldsymbol{\tau}}_{\boldsymbol{v}}^n, \widehat{\boldsymbol{\tau}}_{\boldsymbol{y}}^n, \widehat{\boldsymbol{\tau}}_{\boldsymbol{v}}^n)$ as almost identical to those in Lemma 3.5. Concerning $(\boldsymbol{\tau}_{\boldsymbol{y}}^n, \boldsymbol{\tau}_{\boldsymbol{y}}^n)$, though the approach we adopt is also conceptually similar, it is obviously more

technicality-laden. Firstly, from Taylor expansions stem

$$\left\|\frac{\boldsymbol{y}_\varepsilon(t_{n+1})-\boldsymbol{y}_\varepsilon(t_n)}{\Delta t}-\boldsymbol{y}'_\varepsilon(t_n)-\frac{\Delta t}{2}\boldsymbol{y}''_\varepsilon(t_n)\right\| \leq \frac{(\Delta t)^2}{6}\max_{[t_n,t_{n+1}]}\|\boldsymbol{y}'''_\varepsilon\|$$
$$\lesssim_{\mathbf{E}} \frac{(\Delta t)^2}{\varepsilon^3}(1+\varepsilon^2)\left(\varepsilon+\max_{[t_n,t_{n+1}]}\|\boldsymbol{v}_\varepsilon\|\right)+\frac{(\Delta t)^2}{\varepsilon^2}\max_{[t_n,t_{n+1}]}\|\boldsymbol{v}_\varepsilon\|^2,$$

$$\left\|\frac{\boldsymbol{v}_\varepsilon(t_{n+1})-\boldsymbol{v}_\varepsilon(t_n)}{\Delta t}-\boldsymbol{v}'_\varepsilon(t_n)-\frac{\Delta t}{2}\boldsymbol{v}''_\varepsilon(t_n)\right\| \leq \frac{(\Delta t)^2}{6}\max_{[t_n,t_{n+1}]}\|\boldsymbol{v}'''_\varepsilon\|\,,$$
$$\lesssim_{\mathbf{E}} \frac{(\Delta t)^2}{\varepsilon^6}(1+\varepsilon^4)\left(\varepsilon+\max_{[t_n,t_{n+1}]}\|\boldsymbol{v}_\varepsilon\|\right)+\frac{(\Delta t)^2}{\varepsilon^3}\max_{[t_n,t_{n+1}]}\|\boldsymbol{v}_\varepsilon\|^2\,,$$

where the bounds on $\boldsymbol{y}'''_\varepsilon$ and $\boldsymbol{v}'''_\varepsilon$ have been obtained from (1.7).

Likewise

$$\left\|(1-\gamma)\boldsymbol{v}_\varepsilon(\widetilde{t}_{n+1})+\gamma\boldsymbol{v}_\varepsilon(t_{n+1})-\boldsymbol{v}_\varepsilon(t_n)-\frac{\Delta t}{2}\boldsymbol{v}'_\varepsilon(t_n)\right\| \lesssim_{\mathbf{E}} (\Delta t)^2\max_{[t_n,t_{n+1}]}\|\boldsymbol{v}''_\varepsilon\|,$$
$$\lesssim_{\mathbf{E}} \frac{(\Delta t)^2}{\varepsilon^4}(1+\varepsilon^2)\left(\varepsilon+\max_{[t_n,t_{n+1}]}\|\boldsymbol{v}_\varepsilon\|\right).$$

With a bit more manipulations, we also derive on one hand

$$\left\|\mathbf{E}(t_n,\boldsymbol{y}_\varepsilon(t_n)+\varepsilon\,\boldsymbol{v}_\varepsilon^\perp(\widetilde{t}_{n+1}))-\mathbf{E}(t_n,\boldsymbol{x}_\varepsilon(t_n))-\gamma\varepsilon\Delta t\,\mathrm{d}_{\boldsymbol{x}}\mathbf{E}(t_n,\boldsymbol{x}_\varepsilon(t_n))\,(\boldsymbol{v}_\varepsilon^\perp)'(t_n)\right\|$$
$$\lesssim_{\mathbf{E}} \varepsilon^2(\Delta t)^2\left(\max_{[t_n,\widetilde{t}_{n+1}]}\|\boldsymbol{v}'_\varepsilon\|\right)^2+\varepsilon(\Delta t)^2\max_{[t_n,\widetilde{t}_{n+1}]}\|\boldsymbol{v}''_\varepsilon\|$$
$$\lesssim_{\mathbf{E}} \frac{(\Delta t)^2}{\varepsilon^2}\left(\max_{[t_n,\widetilde{t}_{n+1}]}\|\boldsymbol{v}_\varepsilon\|\right)^2+\frac{(\Delta t)^2}{\varepsilon^3}(1+\varepsilon^2)\left(\varepsilon+\max_{[t_n,\widetilde{t}_{n+1}]}\|\boldsymbol{v}_\varepsilon\|\right),$$

on the other hand

$$\left\|\mathbf{E}(\widehat{t}_{n+1},\boldsymbol{y}_\varepsilon(\widehat{t}_{n+1})+\varepsilon\,\boldsymbol{v}_\varepsilon^\perp(t_{n+1}))-\mathbf{E}(\widehat{t}_{n+1},\boldsymbol{x}_\varepsilon(\widehat{t}_{n+1}))-\left(1-\frac{1}{2\gamma}\right)\varepsilon\Delta t\,\mathrm{d}_{\boldsymbol{x}}\mathbf{E}(t_n,\boldsymbol{x}_\varepsilon(t_n))\,(\boldsymbol{v}_\varepsilon^\perp)'(t_n)\right\|$$
$$\lesssim_{\mathbf{E}} \frac{(\Delta t)^2}{\varepsilon^2}\left(\max_{[t_n,\widehat{t}_{n+1}]}\|\boldsymbol{v}_\varepsilon\|\right)^2+\frac{(\Delta t)^2}{\varepsilon^3}(1+\varepsilon^2)\left(\varepsilon+\max_{[t_n,\widehat{t}_{n+1}]}\|\boldsymbol{v}_\varepsilon\|\right),$$

since

$$\left\|\mathrm{d}_{\boldsymbol{x}}\mathbf{E}(\widehat{t}_{n+1},\boldsymbol{x}_\varepsilon(\widehat{t}_{n+1}))-\mathrm{d}_{\boldsymbol{x}}\mathbf{E}(t_n,\boldsymbol{x}_\varepsilon(t_n))\right\| \lesssim_{\mathbf{E}} \frac{\Delta t}{\varepsilon}\left(\varepsilon+\max_{[t_n,\widehat{t}_{n+1}]}\|\boldsymbol{v}_\varepsilon\|\right),$$

and lastly

$$\left\|\mathbf{E}(\widehat{t}_{n+1},\boldsymbol{x}_\varepsilon(\widehat{t}_{n+1}))-\mathbf{E}(t_n,\boldsymbol{x}_\varepsilon(t_n))-\frac{\Delta t}{2\gamma}\partial_t\mathbf{E}(t_n,\boldsymbol{x}_\varepsilon(t_n))-\frac{\Delta t}{2\gamma}\mathrm{d}_{\boldsymbol{x}}\mathbf{E}(t_n,\boldsymbol{x}_\varepsilon(t_n))\boldsymbol{x}'_\varepsilon(t_n)\right\|$$
$$\lesssim_{\mathbf{E}} (\Delta t)^2\left[1+\left(\max_{[t_n,\widehat{t}_{n+1}]}\|\boldsymbol{x}'_\varepsilon\|\right)^2+\max_{[t_n,\widehat{t}_{n+1}]}\|\boldsymbol{x}''_\varepsilon\|\right]$$
$$\lesssim_{\mathbf{E}} \frac{(\Delta t)^2}{\varepsilon^2}\left(\max_{[t_n,\widehat{t}_{n+1}]}\|\boldsymbol{v}_\varepsilon\|\right)^2+\frac{(\Delta t)^2}{\varepsilon^3}(1+\varepsilon^2)\left(\varepsilon+\max_{[t_n,\widehat{t}_{n+1}]}\|\boldsymbol{v}_\varepsilon\|\right).$$

Finally, by using the definition of γ to derive the following identities

$$(1-\gamma)\,\gamma+\gamma\left(1-\frac{1}{2\gamma}\right)=0\,,\qquad\qquad (1-\gamma)\,\gamma+\gamma=\frac{1}{2}\,,$$

one may combine all the estimates with (1.7) to obtain

$$\begin{cases} \|\boldsymbol{\tau}_{\boldsymbol{y}}^n\| & \underset{\mathbf{E}}{\lesssim} \dfrac{(\Delta t)^2}{\varepsilon^3}(1+\varepsilon^2)\left(\varepsilon + \max\limits_{[t_n,\widehat{t}_{n+1}]} \|\boldsymbol{v}_\varepsilon\|\right) + \dfrac{(\Delta t)^2}{\varepsilon^2}\left(\max\limits_{[t_n,\widehat{t}_{n+1}]} \|\boldsymbol{v}_\varepsilon\|\right)^2, \\ \|\boldsymbol{\tau}_{\boldsymbol{v}}^n\| & \underset{\mathbf{E}}{\lesssim} \dfrac{(\Delta t)^2}{\varepsilon^6}(1+\varepsilon^4)\left(\varepsilon + \max\limits_{[t_n,\widehat{t}_{n+1}]} \|\boldsymbol{v}_\varepsilon\|\right) + \dfrac{(\Delta t)^2}{\varepsilon^3}\left(\max\limits_{[t_n,\widehat{t}_{n+1}]} \|\boldsymbol{v}_\varepsilon\|\right)^2, \end{cases}$$

and the proof is concluded by applying Lemma 2.3. □

As we discussed in §3.1, in addition to the consistency estimates of the foregoing lemma, one needs a stability analysis. To investigate the stability of the implicit part, we observe that combining velocity updates of the scheme (5.2), employing the explicit expression of $\mathbf{F}_{\boldsymbol{v}}^{\varepsilon}$, and manipulating the terms, one derives

$$\boldsymbol{v}_\varepsilon^{n+1} = (\mathrm{Id} + \gamma\,\lambda\,\mathbf{J})^{-1}\left(\mathrm{Id} - (1-\gamma)\mathbf{J}\,(\mathrm{Id} + \gamma\,\lambda\,\mathbf{J})^{-1}\right)\boldsymbol{v}_\varepsilon^n + \cdots,$$

with $\lambda = \Delta t/\varepsilon^2$. Hence the need to investigate the stability of the matrix $\mathbf{A}_\lambda$ defined as

$$\mathbf{A}_\lambda := (\mathrm{Id} + \gamma\,\lambda\,\mathbf{J})^{-1}(\mathrm{Id} - (1-\gamma)\lambda\mathbf{J}\,(\mathrm{Id} + \gamma\,\lambda\,\mathbf{J})^{-1}) = (\mathrm{Id} + \gamma\,\lambda\,\mathbf{J})^{-2}(\mathrm{Id} + (2\gamma-1)\lambda\mathbf{J}).$$

The following lemma encodes the L-stability of the implicit part of the IMEX scheme (5.2), damping high frequency oscillations.

Lemma 5.3. *Let* $\mathbf{J}$ *be as in* (1.4). *Then, for any* $\lambda > 0$, *the matrix* $\mathbf{A}_\lambda$ *satisfies*

$$\|\mathbf{A}_\lambda\| = \frac{1}{\sqrt{1 + \frac{\gamma^4\lambda^4}{1+2\gamma^2\lambda^2}}} < 1\,.$$

Proof. Proceeding as in the proof of Lemma 3.6 yields

$$\left\|(\mathrm{Id} + \gamma\,\lambda\,\mathbf{J})^{-2}(\mathrm{Id} + (2\gamma-1)\lambda\mathbf{J})\right\| = \frac{\sqrt{1 + (2\gamma-1)^2\lambda^2}}{1+\gamma^2\lambda^2}\,.$$

Since from the equation defining γ stems $(2\gamma-1)^2 = 2\gamma^2$, this achieves the proof. □

Proposition 5.4. *There exists a constant* $C_0 > 0$ *such that, when* $\mathbf{E} \in W^{2,\infty}$, *the error from the unique solution of the scheme* (5.2) *to the exact solution of the system* (1.7) *is such that for any* $\Delta t > 0$, $n \geq 0$ *and* $\varepsilon > 0$

$$\begin{aligned} \|\boldsymbol{x}_\varepsilon^n - \boldsymbol{x}_\varepsilon(t_n)\| &\leq \|\boldsymbol{y}_\varepsilon^n - \boldsymbol{y}_\varepsilon(t_n)\| + \varepsilon\|\boldsymbol{v}_\varepsilon^n - \boldsymbol{v}_\varepsilon(t_n)\| \\ &\underset{\mathbf{E}}{\lesssim} \frac{(\Delta t)^2}{\varepsilon^5}\,t_n\,e^{C_0K_xt_n}\Big(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,(1+t_n)\Big)\Big(1 + \varepsilon^3\,\left(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,(1+t_n)\right)\Big)\,. \end{aligned}$$

Proof. As in the proof of Proposition 3.7, we introduce numerical errors as

$$\boldsymbol{e}_{\boldsymbol{y}}^n := \boldsymbol{y}_\varepsilon(t_n) - \boldsymbol{y}_\varepsilon^n\,, \qquad \boldsymbol{e}_{\boldsymbol{v}}^n := \boldsymbol{v}_\varepsilon(t_n) - \boldsymbol{v}_\varepsilon^n\,,$$

and likewise for the intermediate stages

$$\boldsymbol{e}_{\widetilde{\boldsymbol{v}}}^n := \boldsymbol{v}_\varepsilon(\widetilde{t}_n) - \widetilde{\boldsymbol{v}}_\varepsilon^n\,, \qquad \boldsymbol{e}_{\widehat{\boldsymbol{y}}}^n := \boldsymbol{y}_\varepsilon(\widehat{t}_n) - \widehat{\boldsymbol{y}}_\varepsilon^n\,, \qquad \boldsymbol{e}_{\widehat{\boldsymbol{v}}}^n := \boldsymbol{v}_\varepsilon(\widehat{t}_n) - \widehat{\boldsymbol{v}}_\varepsilon^n\,.$$

We also use the norm $\|\cdot\|_\varepsilon$ from (3.9) for our estimates and set $\lambda = \Delta t/\varepsilon^2$.

On the one hand, by applying Lemma 3.6, we obtain from the first stage (5.2a)

$$\varepsilon\left\|\boldsymbol{e}_{\widetilde{\boldsymbol{v}}}^{n+1} - (\mathrm{Id} + \gamma\lambda\mathbf{J})^{-1}\boldsymbol{e}_{\boldsymbol{v}}^n\right\| \lesssim \min\left(1, \frac{1}{\lambda}\right)\,\left(K_x\Delta t\,\|(\boldsymbol{e}_{\boldsymbol{y}}^n, \boldsymbol{e}_{\boldsymbol{v}}^n)\|_\varepsilon + \Delta t\,\varepsilon\,\|\widetilde{\boldsymbol{\tau}}_{\boldsymbol{v}}^n\|\right)\,,$$

and from the intermediate stage (5.2b)

$$\|(\boldsymbol{e}^{n+1}_{\widehat{\boldsymbol{y}}}, \boldsymbol{e}^{n+1}_{\widehat{\boldsymbol{v}}})\|_\varepsilon \lesssim (1 + K_x\Delta t)\, \|(\boldsymbol{e}^n_{\boldsymbol{y}}, \boldsymbol{e}^n_{\boldsymbol{v}})\|_\varepsilon + \Delta t\, \|(\widehat{\boldsymbol{\tau}}^n_{\boldsymbol{y}}, \widehat{\boldsymbol{\tau}}^n_{\boldsymbol{v}})\|_\varepsilon \\ + (\lambda + K_x\Delta t)\, \varepsilon \|\boldsymbol{e}^{n+1}_{\widetilde{\boldsymbol{v}}} - (\mathrm{Id} + \gamma\lambda\mathbf{J})^{-1}\boldsymbol{e}^n_{\boldsymbol{v}}\|\,.$$

$$\|(\boldsymbol{e}^{n+1}_{\widehat{\boldsymbol{y}}}, \boldsymbol{e}^{n+1}_{\widehat{\boldsymbol{v}}})\|_\varepsilon \lesssim (1 + K_x\Delta t)^2\, \|(\boldsymbol{e}^n_{\boldsymbol{y}}, \boldsymbol{e}^n_{\boldsymbol{v}})\|_\varepsilon + \Delta t\, \|(\widehat{\boldsymbol{\tau}}^n_{\boldsymbol{y}}, \widehat{\boldsymbol{\tau}}^n_{\boldsymbol{v}})\|_\varepsilon + \Delta t\, (1 + K_x\Delta t)\, \varepsilon\, \|\widetilde{\boldsymbol{\tau}}^n_{\boldsymbol{v}}\|\,.$$

On the other hand, by using Lemmas 3.6 again, the last stage (5.2c) yields both

$$\|\boldsymbol{e}^{n+1}_{\boldsymbol{y}}\| - \|\boldsymbol{e}^n_{\boldsymbol{y}}\| \lesssim K_x\Delta t\, \|(\boldsymbol{e}^n_{\boldsymbol{y}}, \boldsymbol{e}^n_{\boldsymbol{v}})\|_\varepsilon + K_x\Delta t \|(\boldsymbol{e}^{n+1}_{\widehat{\boldsymbol{y}}}, \boldsymbol{e}^{n+1}_{\widehat{\boldsymbol{v}}})\|_\varepsilon + \Delta t\, \|\boldsymbol{\tau}^n_{\boldsymbol{y}}\| \\ + K_x\Delta t\, \varepsilon \|\boldsymbol{e}^{n+1}_{\widetilde{\boldsymbol{v}}} - (\mathrm{Id} + \gamma\lambda\mathbf{J})^{-1}\boldsymbol{e}^n_{\boldsymbol{v}}\|$$

and

$$\varepsilon\|\boldsymbol{e}^{n+1}_{\boldsymbol{v}} - \mathbf{A}_\lambda\, \boldsymbol{e}^n_{\boldsymbol{v}}\| \lesssim K_x\Delta t\, \|(\boldsymbol{e}^n_{\boldsymbol{y}}, \boldsymbol{e}^n_{\boldsymbol{v}})\|_\varepsilon + K_x\Delta t \|(\boldsymbol{e}^{n+1}_{\widehat{\boldsymbol{y}}}, \boldsymbol{e}^{n+1}_{\widehat{\boldsymbol{v}}})\|_\varepsilon + \Delta t\, \varepsilon\|\boldsymbol{\tau}^n_{\boldsymbol{v}}\| \\ + \Delta t \left(\frac{1}{\varepsilon^2} + K_x\right) \varepsilon \|\boldsymbol{e}^{n+1}_{\widetilde{\boldsymbol{v}}} - (\mathrm{Id} + \gamma\lambda\mathbf{J})^{-1}\boldsymbol{e}^n_{\boldsymbol{v}}\|\,.$$

By gathering the foregoing estimates and using Lemma 5.3 we deduce for $n \geq 0$

$$\|(\boldsymbol{e}^{n+1}_{\boldsymbol{y}}, \boldsymbol{e}^{n+1}_{\boldsymbol{v}})\|_\varepsilon - \|(\boldsymbol{e}^n_{\boldsymbol{y}}, \boldsymbol{e}^n_{\boldsymbol{v}})\|_\varepsilon \lesssim K_x\Delta t\, (1 + K_x\Delta t)^2\, \|(\boldsymbol{e}^n_{\boldsymbol{y}}, \boldsymbol{e}^n_{\boldsymbol{v}})\|_\varepsilon + \Delta t\, \|(\boldsymbol{\tau}^n_{\boldsymbol{y}}, \boldsymbol{\tau}^n_{\boldsymbol{v}})\|_\varepsilon \\ + K_x(\Delta t)^2\, \|(\widehat{\boldsymbol{\tau}}^n_{\boldsymbol{y}}, \widehat{\boldsymbol{\tau}}^n_{\boldsymbol{v}})\|_\varepsilon + (\Delta t)^2 \frac{1 + K_x\varepsilon^2}{\varepsilon^2}\, (1 + K_x\Delta t)\, \varepsilon\, \|\widetilde{\boldsymbol{\tau}}^n_{\boldsymbol{v}}\|.$$

The proof is then concluded as in the proof of Proposition 4.2, by application of the discrete Grönwall lemma and Lemma 5.2. □

We, then, continue with the direct convergence analysis of (5.3a)–(5.3b) to (2.1), which is the counterpart of Proposition 5.4 for the asymptotic model.

Proposition 5.5. *There exists a constant $C_0 > 0$ such that when $\mathbf{E} \in W^{2,\infty}$, we have*

$$\|\boldsymbol{x}^n - \boldsymbol{x}(t_n)\| \underset{\mathbf{E}}{\lesssim} (\Delta t)^2\, t_n\, e^{C_0 K_x\, t_n}\,,$$

when $(\boldsymbol{x}^n)_{n\in\mathbf{N}}$ and $\boldsymbol{x}$ solve respectively (5.3a)–(5.3b) and (2.1), with the same initial datum.

Proof. The proof is omitted as essentially a simpler version of the proof of Proposition 5.4. □

5.2. **Asymptotic estimates.** Regarding the asymptotic part of the convergence analysis, and to prepare the comparisons between solutions of (5.2a)–(5.2c) and (5.3a)–(5.3b), we now examine ε-uniform boundedness of solution of the scheme (5.2a)–(5.2c). To do so, as we have done in previous sections, we work with auxiliary variables that are small corrections to stiff velocity variables. They are defined as

$$\widetilde{\boldsymbol{z}}^n_\varepsilon := \widetilde{\boldsymbol{v}}^n_\varepsilon + \varepsilon\, \mathbf{E}^\perp(t_{n-1}, \boldsymbol{y}^{n-1}_\varepsilon + \varepsilon(\boldsymbol{v}^{n-1}_\varepsilon)^\perp) = \varepsilon \mathbf{J}\mathbf{F}^\varepsilon_{\boldsymbol{v}}(t_{n-1}, \boldsymbol{y}^{n-1}_\varepsilon, \varepsilon\, \boldsymbol{v}^{n-1}_\varepsilon, \varepsilon\, \widetilde{\boldsymbol{v}}^n_\varepsilon), \qquad n \geq 1,$$

for the correction to the (intermediate) updated velocity $\widetilde{\boldsymbol{v}}^n_\varepsilon$, and

$$\boldsymbol{z}^n_\varepsilon := \boldsymbol{v}^n_\varepsilon + \varepsilon\, \mathbf{E}^\perp(\widehat{t}_n, \widehat{\boldsymbol{y}}^n_\varepsilon + \varepsilon(\widehat{\boldsymbol{v}}^n_\varepsilon)^\perp) = \varepsilon \mathbf{J}\mathbf{F}^\varepsilon_{\boldsymbol{v}}(\widehat{t}_n, \widehat{\boldsymbol{y}}^n_\varepsilon, \varepsilon\, \widehat{\boldsymbol{v}}^n_\varepsilon, \varepsilon\, \boldsymbol{v}^n_\varepsilon)\,, \qquad n \geq 1\,,$$

for the correction to the (final) updated velocity $\boldsymbol{v}^n_\varepsilon$. By employing the scheme (5.2a)–(5.2c), with $\lambda = \Delta t/\varepsilon^2$, these definitions imply the following updates for $n \geq 1$

$$\widetilde{\boldsymbol{z}}^{n+1}_\varepsilon - (\mathrm{Id} + \gamma\lambda\mathbf{J})^{-1}\boldsymbol{z}^n_\varepsilon = \varepsilon\, (\mathrm{Id} + \gamma\lambda\mathbf{J})^{-1} \left(\mathbf{E}^\perp(t_n, \boldsymbol{y}^n_\varepsilon + \varepsilon(\boldsymbol{v}^n_\varepsilon)^\perp) - \mathbf{E}^\perp(\widehat{t}_n, \widehat{\boldsymbol{y}}^n_\varepsilon + \varepsilon(\widehat{\boldsymbol{v}}^n_\varepsilon)^\perp)\right)$$

$$\boldsymbol{z}^{n+1}_\varepsilon - \mathbf{A}_\lambda \boldsymbol{z}^n_\varepsilon = -(1-\gamma)\lambda\mathbf{J}\, (\mathrm{Id} + \gamma\lambda\mathbf{J})^{-1} \left(\widetilde{\boldsymbol{z}}^{n+1}_\varepsilon - (\mathrm{Id} + \gamma\lambda\mathbf{J})^{-1}\boldsymbol{z}^n_\varepsilon\right) \\ + \varepsilon\, (\mathrm{Id} + \gamma\lambda\mathbf{J})^{-1} \left(\mathbf{E}^\perp(\widehat{t}_{n+1}, \widehat{\boldsymbol{y}}^{n+1}_\varepsilon + \varepsilon(\widehat{\boldsymbol{v}}^{n+1}_\varepsilon)^\perp) - \mathbf{E}^\perp(\widehat{t}_n, \widehat{\boldsymbol{y}}^n_\varepsilon + \varepsilon(\widehat{\boldsymbol{v}}^n_\varepsilon)^\perp)\right)$$

and the initial values

$$\widetilde{\boldsymbol{z}}_\varepsilon^1 - (\mathrm{Id}+\gamma\lambda\mathbf{J})^{-1}\boldsymbol{v}_\varepsilon^0 = \varepsilon\,(\mathrm{Id}+\gamma\lambda\mathbf{J})^{-1}\,\mathbf{E}^\perp(t_0,\boldsymbol{y}_\varepsilon^0+\varepsilon(\boldsymbol{v}_\varepsilon^0)^\perp)$$

$$\boldsymbol{z}_\varepsilon^1 - \mathbf{A}_\lambda\boldsymbol{v}_\varepsilon^0 = -(1-\gamma)\lambda\mathbf{J}\,(\mathrm{Id}+\gamma\lambda\mathbf{J})^{-1}\,\left(\widetilde{\boldsymbol{z}}_\varepsilon^1 - (\mathrm{Id}+\gamma\lambda\mathbf{J})^{-1}\boldsymbol{v}_\varepsilon^0\right) + \varepsilon\,(\mathrm{Id}+\gamma\lambda\mathbf{J})^{-1}\,\mathbf{E}^\perp(\widehat{t}_1,\widehat{\boldsymbol{y}}_\varepsilon^1+\varepsilon(\widehat{\boldsymbol{v}}_\varepsilon^1)^\perp)\,.$$

Lemma 5.6. *There exists a constant $C_0>0$ such that, when $\mathbf{E}\in W^{1,\infty}$, for $n\geq 1$,*

$$\|\widetilde{\boldsymbol{v}}_\varepsilon^n\| \lesssim \|\widetilde{\boldsymbol{z}}_\varepsilon^n\| + K_0\,\varepsilon\,, \qquad \|\boldsymbol{v}_\varepsilon^n\| \lesssim \|\boldsymbol{z}_\varepsilon^n\| + K_0\,\varepsilon\,,$$

$$\|\widetilde{\boldsymbol{z}}_\varepsilon^n\| \lesssim \min\left(1,\frac{\varepsilon^2}{\Delta t}\right)\left(\|\boldsymbol{v}_\varepsilon^{n-1}\| + K_0\,\varepsilon\right)\,, \qquad \|\widehat{\boldsymbol{v}}_\varepsilon^n\| \lesssim \|\boldsymbol{v}_\varepsilon^{n-1}\| + K_0\,\varepsilon\,,$$

and

$$\|\boldsymbol{z}_\varepsilon^n\| \lesssim e^{C_0K_x\,t_n}\left(\|\boldsymbol{v}_\varepsilon^0\| + \varepsilon\,K_0 + \varepsilon\,t^n(K_t+K_xK_0)\right)\,.$$

Proof. It follows from (5.2b)–(5.2c) that for $n\geq 1$,

$$\|\boldsymbol{y}_\varepsilon^n - \widehat{\boldsymbol{y}}_\varepsilon^n\| \lesssim K_0\Delta t\,, \qquad \|\widehat{\boldsymbol{y}}_\varepsilon^{n+1} - \widehat{\boldsymbol{y}}_\varepsilon^n\| \lesssim K_0\Delta t\,, \qquad \|\boldsymbol{v}_\varepsilon^n - \widehat{\boldsymbol{v}}_\varepsilon^n\| \lesssim \frac{\Delta t}{\varepsilon^2}\left(\|\boldsymbol{z}_\varepsilon^n\| + \|\widetilde{\boldsymbol{z}}_\varepsilon^n\|\right)\,,$$

and

$$\|\widehat{\boldsymbol{v}}_\varepsilon^{n+1} - \widehat{\boldsymbol{v}}_\varepsilon^n\| \lesssim \frac{\Delta t}{\varepsilon^2}\left(\|\boldsymbol{z}_\varepsilon^n\| + \|\widetilde{\boldsymbol{z}}_\varepsilon^n\| + \|\widetilde{\boldsymbol{z}}_\varepsilon^{n+1}\|\right) \lesssim \frac{\Delta t}{\varepsilon^2}\left(\|\boldsymbol{z}_\varepsilon^n\| + \|\widetilde{\boldsymbol{z}}_\varepsilon^n\| + K_0\varepsilon\right)\,.$$

Thus for some $c_0>0$, for $n\geq 1$,

$$\begin{aligned}\|\widetilde{\boldsymbol{z}}_\varepsilon^{n+1}\| - \|\boldsymbol{z}_\varepsilon^n\| &\lesssim K_x\Delta t\left(\|\boldsymbol{z}_\varepsilon^n\| + \|\widetilde{\boldsymbol{z}}_\varepsilon^n\|\right) + \varepsilon\Delta t(K_t+K_xK_0)\,,\\ \|\boldsymbol{z}_\varepsilon^{n+1}\| - \|\boldsymbol{z}_\varepsilon^n\| - c_0\,K_x\Delta t\|\widetilde{\boldsymbol{z}}_\varepsilon^n\| &\lesssim K_x\Delta t\,\|\boldsymbol{z}_\varepsilon^n\| + \varepsilon\Delta t(K_t+K_xK_0)\,,\end{aligned}$$

that may be combined to give for $n\geq 1$,

$$\begin{aligned}&(\|\boldsymbol{z}_\varepsilon^{n+1}\| + c_0K_x\Delta t\|\widetilde{\boldsymbol{z}}_\varepsilon^{n+1}\|) - (\|\boldsymbol{z}_\varepsilon^n\| + c_0K_x\Delta t\|\widetilde{\boldsymbol{z}}_\varepsilon^n\|)\\ &\qquad\lesssim K_x\Delta t(1+K_x\Delta t)(\|\boldsymbol{z}_\varepsilon^n\| + c_0K_x\Delta t\|\widetilde{\boldsymbol{z}}_\varepsilon^n\|) + \varepsilon\Delta t(1+K_x\Delta t)(K_t+K_xK_0)\,.\end{aligned}$$

At this stage completing an application of the discrete Grönwall lemma with the initial bounds

$$\|\widetilde{\boldsymbol{z}}_\varepsilon^1\| \lesssim \|\boldsymbol{v}_\varepsilon^0\| + K_0\varepsilon\,, \qquad \|\boldsymbol{z}_\varepsilon^1\| \lesssim \|\boldsymbol{v}_\varepsilon^0\| + K_0\varepsilon\,.$$

achieves the proofs of the bound on $\boldsymbol{z}_\varepsilon^n$.

Bounds on $\boldsymbol{v}_\varepsilon^n$ and $\widetilde{\boldsymbol{v}}_\varepsilon^n$ are obvious from the definitions, bounds on $\widetilde{\boldsymbol{z}}_\varepsilon^n$ and $\widehat{\boldsymbol{v}}_\varepsilon^n$ follow from (5.2a)-(5.2b) and Lemma 3.6. □

Now, to state a comparison result we introduce notation $\mathbf{X}^{\Delta t}$ and $(\mathbf{X}_\varepsilon^{\Delta t},\mathbf{V}_\varepsilon^{\Delta t})$ to denote discrete flows for (5.3a)-(5.3b) and (5.2a)–(5.2c). We also set $\mathbf{Y}_\varepsilon^{\Delta t} := \mathbf{X}_\varepsilon^{\Delta t} - \varepsilon\,(\mathbf{V}_\varepsilon^{\Delta t})^\perp$.

Proposition 5.7. *(i) There exists a constant $C_0>0$ such that when $\mathbf{E}\in W^{1,\infty}$*

$$\|\mathbf{X}_\varepsilon^{\Delta t}(t_n,0,\boldsymbol{x}_\varepsilon^0,\boldsymbol{v}_\varepsilon^0) - \mathbf{X}^{\Delta t}(t_n,0,\boldsymbol{x}_\varepsilon^0)\| \underset{\mathbf{E}}{\lesssim} \varepsilon\,e^{C_0K_x\,t_n}(1+t^n)\left(\|\boldsymbol{v}_\varepsilon^0\|+\varepsilon\right)\,.$$

(ii) There exists a constant $C_0>0$ such that when $\mathbf{E}\in W^{2,\infty}$

$$\|\mathbf{Y}_\varepsilon^{\Delta t}(t_n,0,\boldsymbol{x}_\varepsilon^0,\boldsymbol{v}_\varepsilon^0) - \mathbf{X}^{\Delta t}(t_n,0,\boldsymbol{x}_\varepsilon^0-\varepsilon(\boldsymbol{v}_\varepsilon^0)^\perp)\| \underset{\mathbf{E}}{\lesssim} \varepsilon^2\,e^{C_0K_x\,t_n}\,(1+t_n)^3\,\left(1+\|\boldsymbol{v}_\varepsilon^0\|^2+\varepsilon^2\right)\,.$$

Proof. Along the proof we use the notational conventions introduced in the proof of Proposition 3.11 and variations thereof.

As before, concerning the first estimate, we sum differences between respective equations and observe that for $n\geq 0$

$$\|\boldsymbol{x}_\varepsilon^n - \boldsymbol{x}^n\| \lesssim \varepsilon\,[\|\boldsymbol{v}_\varepsilon^n\| + \|\boldsymbol{v}_\varepsilon^0\|] + K_x\Delta t\sum_{\ell=0}^{n-1}\|\boldsymbol{x}_\varepsilon^\ell - \boldsymbol{x}^\ell\| + K_x\Delta t\sum_{\ell=1}^{n}\|\widehat{\boldsymbol{x}}_\varepsilon^\ell - \widehat{\boldsymbol{x}}^\ell\|$$

where for $n\geq 1$, $\widehat{\boldsymbol{x}}_\varepsilon^n := \widehat{\boldsymbol{y}}_\varepsilon^n + \varepsilon(\widehat{\boldsymbol{v}}_\varepsilon^n)^\perp$. Now, from (5.2a) and (5.3a) follows for $n\geq 1$

$$\|\widehat{\boldsymbol{x}}_\varepsilon^n - \widehat{\boldsymbol{x}}^n\| \lesssim \|\boldsymbol{x}_\varepsilon^{n-1} - \boldsymbol{x}^{n-1}\|\,(1+K_x\Delta t) + \varepsilon\,[\|\widehat{\boldsymbol{v}}_\varepsilon^n\| + \|\boldsymbol{v}_\varepsilon^{n-1}\|] + K_x\varepsilon\Delta t[\|\widetilde{\boldsymbol{v}}_\varepsilon^n\| + \|\boldsymbol{v}_\varepsilon^{n-1}\|]\,.$$

Hence for $n \geq 0$

$$\|\boldsymbol{x}_\varepsilon^n - \boldsymbol{x}^n\| \lesssim K_x \Delta t\,(1 + K_x \Delta t) \sum_{\ell=0}^{n-1} \|\boldsymbol{x}_\varepsilon^\ell - \boldsymbol{x}^\ell\| \\ + \varepsilon\,(\|\boldsymbol{v}_\varepsilon^n\| + \|\boldsymbol{v}_\varepsilon^0\|) + \varepsilon\, K_x \Delta t\,(1 + K_x \Delta t) \sum_{\ell=1}^{n} (\|\widehat{\boldsymbol{v}}_\varepsilon^\ell\| + \|\widetilde{\boldsymbol{v}}_\varepsilon^\ell\| + \|\boldsymbol{v}_\varepsilon^{\ell-1}\|)\,.$$

Finally, applying the discrete Grönwall lemma (3.14a)–(3.14b), combined with the velocity bound of Lemma 5.6 concludes the proof of the first inequality.

For the second estimate, one begins with the Taylor expansion of the $\boldsymbol{y}_\varepsilon$-update (5.2c), that is for $n \geq 0$,

$$\begin{aligned} \frac{\boldsymbol{y}_\varepsilon^{n+1} - \boldsymbol{y}_\varepsilon^n}{\Delta t} &+ (1-\gamma)\mathbf{E}^\perp(t_n, \boldsymbol{y}_\varepsilon^n) + \gamma\, \mathbf{E}^\perp(\widehat{t}_{n+1}, \widehat{\boldsymbol{y}}_\varepsilon^{n+1}) \\ &= -\varepsilon\, \mathrm{d}_x \mathbf{E}^\perp(\widehat{t}_{n+1}, \widehat{\boldsymbol{y}}_\varepsilon^{n+1}) \left(\gamma(\boldsymbol{v}_\varepsilon^{n+1})^\perp + (1-\gamma)(\widetilde{\boldsymbol{v}}_\varepsilon^{n+1})^\perp\right) \\ &\quad - \varepsilon\,(1-\gamma)\, \left(\mathrm{d}_x \mathbf{E}^\perp(t_n, \boldsymbol{y}_\varepsilon^n) - \mathrm{d}_x \mathbf{E}^\perp(\widehat{t}_{n+1}, \widehat{\boldsymbol{y}}_\varepsilon^{n+1})\right) (\widetilde{\boldsymbol{v}}_\varepsilon^{n+1})^\perp \\ &\quad + \varepsilon^2 \left((1-\gamma)\Theta_\varepsilon(t_n, \boldsymbol{y}_\varepsilon^n, \widetilde{\boldsymbol{v}}_\varepsilon^{n+1}) + \gamma\, \Theta_\varepsilon(\widehat{t}_{n+1}, \widehat{\boldsymbol{y}}_\varepsilon^{n+1}, \boldsymbol{v}_\varepsilon^{n+1})\right)\,, \end{aligned}$$

with Θ_ε such that $\|\Theta_\varepsilon(t, \boldsymbol{y}, \boldsymbol{v})\| \leq \frac{1}{2} K_{xx} \|\boldsymbol{v}\|^2$. To see that the third line also possesses an ε^2-bound, on may combine Lemma 3.6 that gives for any $n \geq 0$,

$$\|\widetilde{\boldsymbol{v}}_\varepsilon^{n+1}\| \lesssim K_0 \varepsilon + \frac{\varepsilon^2}{\Delta t} \|\boldsymbol{v}_\varepsilon^n\|$$

with (5.2b) that yields

$$\left\|\mathrm{d}_x \mathbf{E}(t_n, \boldsymbol{y}_\varepsilon^n) - \mathrm{d}_x \mathbf{E}(\widehat{t}_{n+1}, \widehat{\boldsymbol{y}}_\varepsilon^{n+1})\right\| \lesssim \min\left(K_x, \Delta t\ (K_{tx} + K_{xx} K_0)\right)\,.$$

Therefore, one may focus on the second line. We observe that, when $n \geq 1$,

$$\begin{aligned} \mathrm{d}_x \mathbf{E}(\widehat{t}_n, \widehat{\boldsymbol{y}}_\varepsilon^n) \left(\gamma(\boldsymbol{v}_\varepsilon^n)^\perp + (1-\gamma)(\widetilde{\boldsymbol{v}}_\varepsilon^n)^\perp\right) \;=\; &- \frac{\varepsilon^2}{\Delta t} \left(\mathrm{d}_x \mathbf{E}(\widehat{t}_{n+1}, \widehat{\boldsymbol{y}}_\varepsilon^{n+1}) \boldsymbol{v}_\varepsilon^n - \mathrm{d}_x \mathbf{E}(\widehat{t}_n, \widehat{\boldsymbol{y}}_\varepsilon^n) \boldsymbol{v}_\varepsilon^{n-1}\right) \\ &+ \varepsilon\, \mathrm{d}_{\boldsymbol{x}} \mathbf{E}(\widehat{t}_n, \widehat{\boldsymbol{y}}_\varepsilon^n)\ \left(\gamma \mathbf{E}(\widehat{t}_n, \widehat{\boldsymbol{y}}_\varepsilon^n) + (1-\gamma)\mathbf{E}(t_{n-1}, \boldsymbol{x}_\varepsilon^{n-1})\right) \\ &+ \frac{\varepsilon^2}{\Delta t} \left(\mathrm{d}_x \mathbf{E}(\widehat{t}_{n+1}, \widehat{\boldsymbol{y}}_\varepsilon^{n+1}) - \mathrm{d}_x \mathbf{E}(\widehat{t}_n, \widehat{\boldsymbol{y}}_\varepsilon^n)\right)\ \boldsymbol{v}_\varepsilon^n\,, \end{aligned}$$

with the following estimate concerning the last term

$$\left\|\mathrm{d}_x \mathbf{E}(\widehat{t}_{n+1}, \widehat{\boldsymbol{y}}_\varepsilon^{n+1}) - \mathrm{d}_x \mathbf{E}(\widehat{t}_n, \widehat{\boldsymbol{y}}_\varepsilon^n)\right\| \lesssim \Delta t\ (K_{tx} + K_{xx} K_0)\,.$$

As a consequence summing yields, for $n \geq 0$,

$$\|\boldsymbol{y}_\varepsilon^n - \boldsymbol{y}^n\| \lesssim K_x \Delta t \sum_{\ell=0}^{n-1} \|\boldsymbol{y}_\varepsilon^\ell - \boldsymbol{y}^\ell\| + K_x \Delta t \sum_{\ell=1}^{n} \|\widehat{\boldsymbol{y}}_\varepsilon^\ell - \widehat{\boldsymbol{y}}^\ell\| \\ + \varepsilon^3\, K_x \left(\|\boldsymbol{v}_\varepsilon^n\| + \|\boldsymbol{v}_\varepsilon^0\|\right) + \varepsilon^2\, t^n\, K_x\, K_0 + \varepsilon^3 \Delta t\ (K_{tx} + K_{xx} K_0) \sum_{\ell=0}^{n} \|\boldsymbol{v}_\varepsilon^\ell\| \\ + \varepsilon^2 \Delta t\, K_{xx} \left(\sum_{\ell=0}^{n} \|\boldsymbol{v}_\varepsilon^\ell\|^2 + \sum_{\ell=1}^{n} \|\widetilde{\boldsymbol{v}}_\varepsilon^\ell\|^2\right)\,.$$

Finally, we note that for $n \geq 1$

$$\|\widehat{\boldsymbol{y}}_\varepsilon^n - \widehat{\boldsymbol{y}}^n\| \lesssim \|\boldsymbol{y}_\varepsilon^{n-1} - \boldsymbol{y}^{n-1}\|\,(1 + K_x \Delta t) + \varepsilon^2 \Delta t\, K_x\, K_0 + \varepsilon^3\, K_x\, \|\boldsymbol{v}_\varepsilon^{n-1}\|\,.$$

Inserting the latter in the former leaves, for $n \geq 0$,

$$\begin{aligned}\|\boldsymbol{y}_\varepsilon^n - \boldsymbol{y}^n\| \lesssim K_x \Delta t (1 + K_x \Delta t) \sum_{\ell=0}^{n-1} \|\boldsymbol{y}_\varepsilon^\ell - \boldsymbol{y}^\ell\| + \varepsilon^3 K_x \left(\|\boldsymbol{v}_\varepsilon^n\| + \|\boldsymbol{v}_\varepsilon^0\|\right) \\ + \varepsilon^2 t^n K_x K_0 (1 + K_x \Delta t) + \varepsilon^3 \Delta t \left(K_x^2 + K_{tx} + K_{xx} K_0\right) \sum_{\ell=0}^{n} \|\boldsymbol{v}_\varepsilon^\ell\| \\ + \varepsilon^2 \Delta t K_{xx} \left(\sum_{\ell=0}^{n} \|\boldsymbol{v}_\varepsilon^\ell\|^2 + \sum_{\ell=1}^{n} \|\widetilde{\boldsymbol{v}}_\varepsilon^\ell\|^2 \right) .\end{aligned}$$

The proof is again achieved by combining the discrete Grönwall lemma with Lemma 5.6. □

5.3. **Proof of Theorem 5.1.** The unique solvability of the scheme is again a direct consequence of Lemma 3.6. Regarding error estimates, one gathers direct estimates in Proposition 5.4 with the combination of Propositions 5.5 and 5.7 and Theorem 2.1 to conclude the proof.

6. Numerical experiments

In this section, we provide an illustration of the error estimates proved in Theorems 4.1 and 5.1 on the simple example of the motion of a single particle subject to an electric field $\mathbf{E} = -\nabla_{\boldsymbol{x}} \phi$ deriving from the potential

$$\phi(\boldsymbol{x}) = \frac{1}{2} \left(\|\boldsymbol{x}\|^2 + \frac{1}{10\pi} \cos^2(2\pi x_2) \right), \qquad \boldsymbol{x} = (x_1, x_2)\,.$$

Initial conditions are chosen as

$$\boldsymbol{x}_\varepsilon^0 = (1,1), \qquad \boldsymbol{v}_\varepsilon^0 = (3,3).$$

Note that for this electric potential, the electric potential does not fit exactly in the framework of our theorems since it is unbounded (though its derivatives from order one and onward are bounded). Yet our observations fit well with our theoretical conclusions.

We observe two error indicators for the variable $\boldsymbol{y}_\varepsilon$,

$$\begin{cases} \mathcal{E}_{\boldsymbol{y}}(\Delta t, \varepsilon) := \displaystyle\sum_{n=1}^{N_T} \Delta t \, \|\boldsymbol{y}_\varepsilon^n - \boldsymbol{y}_\varepsilon(t_n)\|, \\ \mathcal{E}_{\boldsymbol{y},\,\mathrm{gc}}(\Delta t, \varepsilon) := \displaystyle\sum_{n=1}^{N_T} \Delta t \, \|\boldsymbol{y}_\varepsilon^n - \mathbf{X}(t_n, 0, \boldsymbol{x}^0 - \varepsilon(\boldsymbol{v}_\varepsilon^0)^\perp)\|, \end{cases}$$

where $\boldsymbol{y}_\varepsilon := \boldsymbol{x}_\varepsilon - \varepsilon(\boldsymbol{v}_\varepsilon)^\perp$ stems from the solution $(\boldsymbol{x}_\varepsilon, \boldsymbol{v}_\varepsilon)$ of the system of characteristics (1.7) (with $s = 0$) and $\mathbf{X}$ is the flow for the guiding center equation (2.1), whereas $\boldsymbol{y}_\varepsilon^n$ is our numerical approximation of $\boldsymbol{y}_\varepsilon(t_n)$. So, in other words, $\mathcal{E}_{\boldsymbol{y}}$ and $\mathcal{E}_{\boldsymbol{y},\,\mathrm{gc}}$ measure the difference from the numerical approximation to, respectively, the exact ε-dependent and asymptotic solutions, hence quantify respectively numerical convergence and asymptotic convergence. Note that $\mathcal{E}_{\boldsymbol{y}}$ and $\mathcal{E}_{\boldsymbol{y},\,\mathrm{gc}}$ are averaged errors in time so that they take into account the possibly-large errors which may originate from the initial layer. Similarly, we define errors for the velocity variable $\boldsymbol{v}_\varepsilon$ as

$$\begin{cases} \mathcal{E}_{\boldsymbol{v}}(\Delta t, \varepsilon) := \displaystyle\sum_{n=1}^{N_T} \Delta t \, \|\boldsymbol{v}_\varepsilon^n - \boldsymbol{v}_\varepsilon(t_n)\|, \\ \mathcal{E}_{\boldsymbol{v},\,\mathrm{gc}}(\Delta t, \varepsilon) := \displaystyle\sum_{n=1}^{N_T} \Delta t \, \|\varepsilon^{-1}\boldsymbol{v}_\varepsilon^n - \boldsymbol{v}_{\mathrm{gc}}(t_n)\|, \end{cases}$$

using the guiding center velocity $\boldsymbol{v}_{\mathrm{gc}}(t) := -\mathbf{E}^\perp(t, \boldsymbol{x}_\varepsilon(t))$ as the asymptotic velocity.

Since the exact solution of this example is not available, we perform a very accurate numerical simulation by a fourth-order explicit Runge–Kutta scheme as the reference solution. The time step

for this reference solution is chosen small enough, namely of order ε^2 when $\varepsilon \ll 1$, so as to capture the very fast oscillations.

We, then, perform some numerical experiments with the first-order scheme (4.1) to illustrate the results stated in Theorems 4.1.

In Fig. 1(A), we illustrate the computed errors for the first-order scheme (4.1), which match the estimate of Theorem 4.1; one can identify in the figure, roughly speaking, the *min* function of the estimate, *cf.* [27, Fig. 1]: for $\varepsilon \sim 1$, the error grows as one decreases ε up to some turning point in the curve after which the error decreases with ε getting closer to zero. More precisely, one can observe that, when Δt is much smaller than ε (right part of Fig. 1(A)), the error is $\mathcal{O}(\Delta t)$, *i.e.*, the scheme is first-order accurate with respect to Δt, as the classical analysis may suggest. Note that in this regime, the other error estimate in Theorem 4.1, which behaves like $\varepsilon^2 + \Delta t$, is quite large. On the other hand, for smaller values of ε, it is the latter bound which saturates the numerical error due to the blow-up of the classical error estimate, so the error is dominated by $\mathcal{O}(\varepsilon^2)$ as the slope of the error curve suggests. Indeed, we see this saturation in the intermediate region for $\varepsilon \in [10^{-3}, 10^{-1}]$. Of course, when we refine the time step Δt, the intermediate region moves to the left. Finally, when ε^2 is very small compared to Δt, the error in Δt dominates and the error does not decrease with ε any longer.

Moreover, in Fig. 1(B), we compare our numerical approximation with the resolved reference solution of the guiding center model $\mathbf{X}(\cdot, 0, \boldsymbol{x}^0 - \varepsilon(\boldsymbol{v}_\varepsilon^0)^\perp$. This, in fact, confirms that when $\varepsilon \ll 1$, the scheme is first-order accurate and corresponds to the left part of Fig. 1(A), as one expects that the reference solution converges to its asymptotic limit.

We also present the numerical error on the velocity variable, in terms of $\mathcal{E}_{\boldsymbol{v}}$ and $\mathcal{E}_{\boldsymbol{v},\,\mathrm{gc}}$. Fig. 1(C) confirms the point that the scheme should be first-order with respect to Δt, when $\varepsilon \sim 1$. However, for $\varepsilon \ll 1$ and a with large time step, the numerical scheme does not capture fast oscillations; so, no convergence to the reference velocity can be observed. In this regime, the computed velocity is only able to compute slow scale dynamics represented by the guiding center velocity $\boldsymbol{v}_{\mathrm{gc}}$, as Fig. 1(D) suggests.

Furthermore, we perform numerical experiments with the second-order scheme (5.2) to illustrate Theorem 5.1. The numerical results are shown in Fig. 2. We observe the same behavior of the numerical error, but, of course, with a smaller error owing to the higher order of accuracy with respect to Δt, though, again, there is an intermediate region where the dominant error term is $\mathcal{O}(\varepsilon^2)$. These numerical tests, with a smooth solution, underline the expected but clear advantage of the second-order scheme compared to the first-order approximation.

7. Conclusion and perspectives

In this paper, we have presented a complete convergence analysis of particle-in-cell methods for the two-dimensional Vlasov equation, with a given electric field, submitted to an external magnetic field, which is homogeneous in space and time and very strong, of order $1/\varepsilon$ with $\varepsilon \ll 1$. In fact, we have estimated the error for semi-implicit first- and second-order IMEX schemes, and confirmed the stability, accuracy and convergence of these schemes, for any possible values of the time step and of the scaling parameter ε. These theoretical results have been supported by numerical experiments.

An immediate extension is to investigate, in practice and analysis, the applicability of the presented framework for more complicated cases, *e.g.*, for the three-dimensional system or with an inhomogeneous magnetic field. Another interesting extension could be to derive and analyze higher-order asymptotic models, to improve the error estimate in terms of ε. This would be rather involved as higher order terms are coupled with the evolution of the energy; see [13, 15] for instance.

Appendix A. Convergence analysis for oscillatory ODEs

As announced in the introduction, we conclude with abstract considerations on the numerical analysis of oscillatory ODEs. Though insufficient to prove the relevant results, these considerations provide enlightening insights supporting correct educated guesses on the final outcomes.

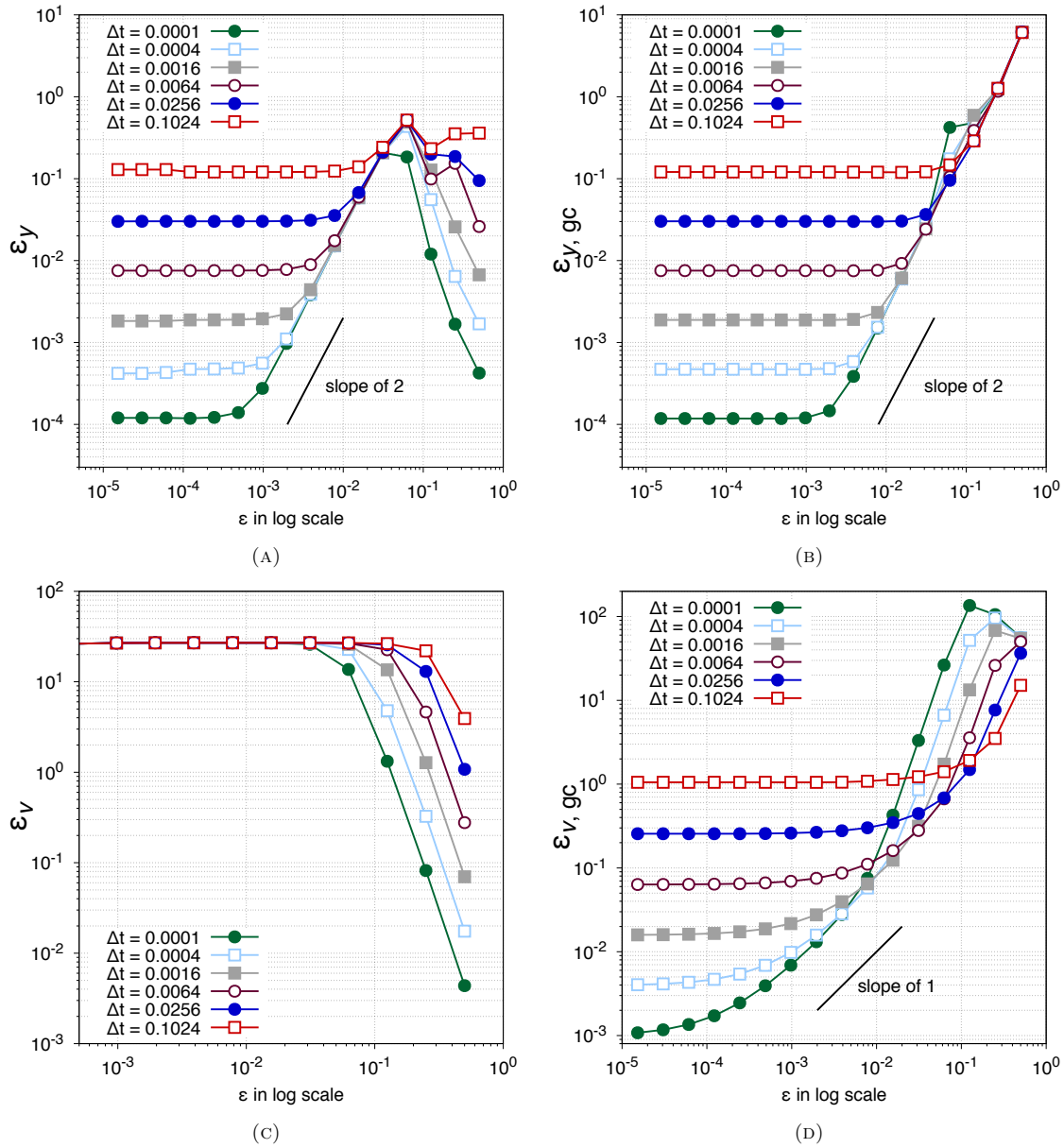

FIGURE 1. **First-order scheme (4.1):** (A) Error between the approximation $\boldsymbol{y}_\varepsilon^n$ and (reference) solution $\boldsymbol{y}_\varepsilon(t_n)$ of (1.7) denoted by $\mathcal{E}_{\boldsymbol{y}}$. (B) Error between the approximation $\boldsymbol{y}_\varepsilon^n$ and the asymptotic (reference) solution $\mathbf{X}(t_n, 0, \boldsymbol{x}^0 - \varepsilon(\boldsymbol{v}^0)^\perp)$ of (2.1) denoted by $\mathcal{E}_{\boldsymbol{y},\,\mathrm{gc}}$. (C) Error between the approximation $\boldsymbol{v}_\varepsilon^n$ and the (reference) solution $\boldsymbol{v}_\varepsilon(t_n)$ of (1.7) denoted by $\mathcal{E}_{\boldsymbol{v}}$. (D) Error between the approximation $\varepsilon^{-1}\boldsymbol{v}_\varepsilon^n$ and the asymptotic (reference) velocity $\boldsymbol{v}_{\mathrm{gc}}(t_n)$ denoted by $\mathcal{E}_{\boldsymbol{v},\,\mathrm{gc}}$.

Let us discuss a system of the form

$$\text{(A.1)}\qquad \begin{cases} (\boldsymbol{a}_\varepsilon + \varepsilon^{r_t}\,\mathbf{G}_t^\varepsilon(\cdot, \boldsymbol{a}_\varepsilon, \boldsymbol{b}_\varepsilon))'(t) = \mathbf{F}_{\boldsymbol{a}}^\varepsilon(t, \boldsymbol{a}_\varepsilon(t)) + \varepsilon^{r_x}\,\mathbf{G}_x^\varepsilon(t, \boldsymbol{a}_\varepsilon(t), \boldsymbol{b}_\varepsilon(t))\,, \\ \boldsymbol{b}_\varepsilon'(t) = -\dfrac{1}{\varepsilon^2}\mathbf{J}\,\boldsymbol{b}_\varepsilon(t) + \mathbf{F}_{\boldsymbol{b}}^\varepsilon(t, \boldsymbol{a}_\varepsilon(t), \boldsymbol{b}_\varepsilon(t))\,, \end{cases}$$

(with $\mathbf{F}_{\boldsymbol{a}}^\varepsilon$, $\mathbf{F}_{\boldsymbol{b}}^\varepsilon$, $\mathbf{G}_t^\varepsilon$, $\mathbf{G}_x^\epsilon$ uniformly smooth) and try to guess what may be expected on the numerical computation of the slow variable $\boldsymbol{a}_\varepsilon$. Expanding the first equation of the system with the second suggests that, with such a goal in mind, a direct convergence analysis of a discretization of (A.1) could be carried out by working with the vector $(\boldsymbol{a}_\varepsilon, \varepsilon^{\min(r_t-2, r_x)}\boldsymbol{b}_\varepsilon)$, and, arguing recursively, that its $(m+1)$th derivative is bounded by a multiple of $\max(\varepsilon^{-(2m-\min(r_t-2,r_x))_+}, \varepsilon^{-(2(m+1)-\min(r_t-2,r_x))})$.

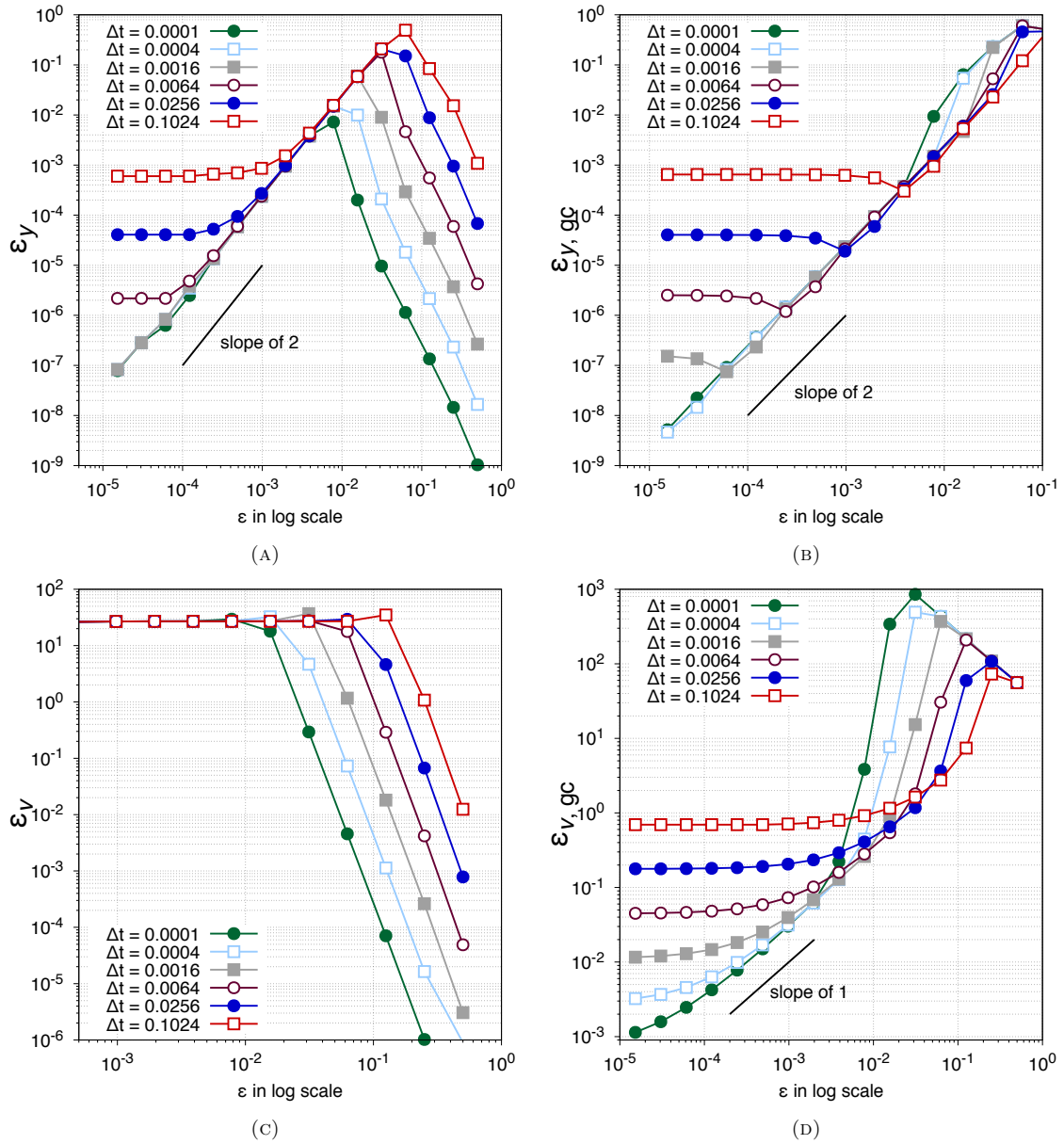

FIGURE 2. **Second-order scheme** (5.2)**:** (A) Error between the approximation $\boldsymbol{y}_\varepsilon^n$ and (reference) solution $\boldsymbol{y}_\varepsilon(t_n)$ of (1.7) denoted by $\mathcal{E}_{\boldsymbol{y}}$. (B) Error between the approximation $\boldsymbol{y}_\varepsilon^n$ and the asymptotic (reference) solution $\mathbf{X}(t_n, 0, \boldsymbol{x}^0 - \varepsilon(\boldsymbol{v}^0)^\perp)$ of (2.1) denoted by $\mathcal{E}_{\boldsymbol{y},\,\mathrm{gc}}$. (C) Error between the approximation $\boldsymbol{v}_\varepsilon^n$ and the (reference) solution $\boldsymbol{v}_\varepsilon(t_n)$ of (1.7) denoted by $\mathcal{E}_{\boldsymbol{v}}$. (D) Error between the approximation $\varepsilon^{-1}\boldsymbol{v}_\varepsilon^n$ and the asymptotic (reference) velocity $\boldsymbol{v}_{\mathrm{gc}}(t_n)$ denoted by $\mathcal{E}_{\boldsymbol{v},\,\mathrm{gc}}$..

As a consequence, a direct convergence analysis of a scheme of order m that would be unconditionally stable would result for the numerical approximation of $(\boldsymbol{a}_\varepsilon, \varepsilon^{\min(r_t-2,r_x)}\boldsymbol{b}_\varepsilon)$, thus also of $\boldsymbol{a}_\varepsilon$, into a bound on numerical error by a multiple of

$$(\Delta t)^m \times \max\left(\frac{1}{\varepsilon^{(2m-\min(r_t-2,r_x))_+}}, \frac{1}{\varepsilon^{(2(m+1)-\min(r_t-2,r_x))}}\right).$$

For concreteness note that when analyzing the computation of the guiding center, $r_t = 3$ and $r_x = 2$ so that the bound is $\Delta t^m/\varepsilon^{(2m+1)}$. The bound is somewhat optimal in the prediction of the computational error for $\varepsilon^{\min(r_t-2,r_x)}\boldsymbol{b}_\varepsilon$. With this respect note that even if by a particularly clever method, for instance through stroboscopic averaging, one is able to improve the computation of a fast variable at particularly well-chosen set of discrete times, this extra precision will be lost

when recovering by interpolation from these discrete times an approximation of $\boldsymbol{b}_\varepsilon$ on the whole continuous time interval.

System (A.1) suggests that the variable $\boldsymbol{a}_\varepsilon$ is actually $\mathcal{O}(\varepsilon^{\min(r_t,r_x)})$-close as $\varepsilon \to 0$ to a solution $\boldsymbol{a}$ of the uncoupled non-stiff equation

$$\boldsymbol{a}'_\varepsilon(t) = \mathbf{F}^\varepsilon_{\boldsymbol{a}}(t, \boldsymbol{a}(t))\,. \tag{A.2}$$

For a scheme of order m consistent with the foregoing asymptotic and unconditionally stable this suggests a bound of the numerical error in the approximation of $\boldsymbol{a}_\varepsilon$ by a multiple of

$$\varepsilon^{\min(r_t,r_x)} + \Delta t^m\,.$$

Note that to conclude to the latter bound it is sufficient to know that the $\boldsymbol{a}$-part of the solution of the discrete scheme for (A.1) converge as $\varepsilon \to 0$ to a solution of a scheme of order m for (A.2) with rate $\mathcal{O}(\varepsilon^{\min(r_t,r_x)} + \varepsilon\,\Delta t^{m\,\frac{\min(r_t,r_x)-1}{\min(r_t,r_x)}})$, leaving room for some depreciation of the continuous rate $\mathcal{O}(\varepsilon^{\min(r_t,r_x)})$.

This provides a final bound of the numerical error for the variable $\boldsymbol{a}_\varepsilon$ by a multiple of

$$\min\left(\varepsilon^{\min(r_t,r_x)} + \Delta t^m, (\Delta t)^m \times \max\left(\frac{1}{\varepsilon^{(2m-\min(r_t-2,r_x))_+}}, \frac{1}{\varepsilon^{(2(m+1)-\min(r_t-2,r_x))}}\right)\right)\,.$$

In the present paper, we have turned the foregoing formal discussion into rigorous convergence analysis for some of the schemes introduced in [12] and a few extensions.

References

[1] C. K. Birdsall and A. B. Langdon, *Plasma physics via computer simulation*, Series in plasma physics, Taylor & Francis, New York, 2005.

[2] S. Boscarino, F. Filbet, and G. Russo, *High order semi-implicit schemes for time dependent partial differential equations*, J. Sci. Comput., 68 (2016), pp. 975–1001.

[3] A. J. Brizard and T. S. Hahm, *Foundations of nonlinear gyrokinetic theory*, Rev. Mod. Phys., 79 (2007), pp. 421–468.

[4] P. Chartier, N. Crouseilles, M. Lemou, F. Méhats, and X. Zhao, *Uniformly accurate methods for Vlasov equations with non-homogeneous strong magnetic field*, Math. Comp., 88 (2019), pp. 2697–2736.

[5] P. Chartier, N. Crouseilles, and X. Zhao, *Numerical methods for the two-dimensional Vlasov–Poisson equation in the finite Larmor radius approximation regime*, J. Comput. Phys., 375 (2018), pp. 619–640.

[6] N. Crouseilles, E. Frénod, S. A. Hirstoaga, and A. Mouton, *Two-scale macro-micro decomposition of the Vlasov equation with a strong magnetic field*, Math. Models Methods Appl. Sci., 23 (2013), pp. 1527–1559.

[7] N. Crouseilles, S. A. Hirstoaga, and X. Zhao, *Multiscale particle-in-cell methods and comparisons for the long-time two-dimensional Vlasov-Poisson equation with strong magnetic field*, Comput. Phys. Commun., 222 (2018), pp. 136–151.

[8] N. Crouseilles, M. Lemou, and F. Méhats, *Asymptotic preserving schemes for highly oscillatory Vlasov–Poisson equations*, J. Comput. Phys., 248 (2013), pp. 287–308.

[9] N. Crouseilles, M. Lemou, F. Méhats, and X. Zhao, *Uniformly accurate forward semi-Lagrangian methods for highly oscillatory Vlasov-Poisson equations*, Multiscale Model. Simul., 15 (2017), pp. 723–744.

[10] ——, *Uniformly accurate particle-in-cell method for the long time solution of the two-dimensional Vlasov-Poisson equation with uniform strong magnetic field*, J. Comput. Phys., 346 (2017), pp. 172–190.

[11] P. Degond and F. Filbet, *On the asymptotic limit of the three dimensional Vlasov–Poisson system for large magnetic field: Formal Derivation*, J. Stat. Phys., 165 (2016), pp. 765–784.

[12] F. Filbet and L. M. Rodrigues, *Asymptotically stable particle-in-cell methods for the Vlasov–Poisson system with a strong external magnetic field*, SIAM J. Numer. Anal., 54 (2016), pp. 1120–1146.

[13] ——, *Asymptotically preserving particle-in-cell methods for inhomogeneous strongly magnetized plasmas*, SIAM J. Numer. Anal., 55 (2017), pp. 2416–2443.

[14] F. Filbet and L. M. Rodrigues, *Asymptotics of the three-dimensional Vlasov equation in the large magnetic field limit*, J. Éc. polytech. Math., 7 (2020), pp. 1009–1067.

[15] F. Filbet and C. Yang, *Numerical simulations to the Vlasov–Poisson system with a strong magnetic field*, arXiv preprint arXiv:1805.10888, (2018).

[16] E. Frénod, S. A. Hirstoaga, M. Lutz, and E. Sonnendrücker, *Long time behaviour of an exponential integrator for a Vlasov–Poisson system with strong magnetic field*, Commun. Comput. Phys., 18 (2015), pp. 263–296.

[17] E. Frénod, S. A. Hirstoaga, and E. Sonnendrücker, *An exponential integrator for a highly oscillatory Vlasov equation*, Discrete Contin. Dyn. Syst. Ser. S, 8 (2015), pp. 169–183.
[18] E. Frénod and É. Sonnendrücker, *Homogenization of the Vlasov equation and of the Vlasov–Poisson system with a strong external magnetic field*, Asymptotic Anal., 18 (1998), pp. 193–213.
[19] ———, *Long time behavior of the two-dimensional Vlasov equation with a strong external magnetic field*, Math. Mod. Meth. Appl. S., 10 (2000), pp. 539–553.
[20] F. Golse and L. Saint-Raymond, *The Vlasov–Poisson system with strong magnetic field*, J. Math. Pure. Appl., 78 (1999), pp. 791–817.
[21] E. Hairer, C. Lubich, and B. Wang, *A filtered Boris algorithm for charged-particle dynamics in a strong magnetic field*, arXiv preprint arXiv:1907.07452, (2019).
[22] D. Han-Kwan, *Contribution à l'étude mathématique des plasmas fortement magnétisés*, PhD thesis, Université Pierre et Marie Curie-Paris VI, 2011.
[23] R. D. Hazeltine and J. D. Meiss, *Plasma Confinement*, Dover Publications, Mineola, New York, 2005.
[24] M. Herda, *Analyse asymptotique et numérique de quelques modèles pour le transport de particules chargées*, PhD thesis, Université Claude Bernard Lyon 1, 2017.
[25] M. Herda and L. M. Rodrigues, *Anisotropic Boltzmann-Gibbs dynamics of strongly magnetized Vlasov-Fokker-Planck equations*, Kinet. Relat. Models, 12 (2019), pp. 593–636.
[26] S. Jin, *Efficient asymptotic-preserving (AP) schemes for some multiscale kinetic equations*, SIAM J. Sci. Comput., 21 (1999), pp. 441–454.
[27] ———, *Asymptotic preserving (AP) schemes for multiscale kinetic and hyperbolic equations: A review*, Lecture Notes for Summer School on "Methods and Models of Kinetic Theory" (M&MKT), Porto Ercole (Grosseto, Italy), (2010), pp. 177–216.
[28] A. Klar, *A numerical method for nonstationary transport equations in diffusive regimes*, Transport Theor. Stat., 27 (1998), pp. 653–666.
[29] J. A. Krommes, *The gyrokinetic description of microturbulence in magnetized plasmas*, Annu. Rev. Fluid Mech., 44 (2012), pp. 175–201.
[30] W. Lee, *Gyrokinetic approach in particle simulation*, Phys. Fluids, 26 (1983), pp. 556–562.
[31] M. Lutz, *Étude mathématique et numérique d'un modèle gyrocinétique incluant des effets électromagnétiques pour la simulation d'un plasma de Tokamak*, PhD thesis, Université de Strasbourg, 2013.
[32] V. F. Matteo, *Gyrokinetic theory for particle transport in fusion plasmas*, PhD thesis, Università di Roma Tre, 2017.
[33] É. Miot, *On the gyrokinetic limit for the two-dimensional Vlasov–Poisson system*, arXiv preprint, arXiv-1603.04502 (2016).
[34] L. Saint-Raymond, *Control of large velocities in the two-dimensional gyrokinetic approximation*, J. Math. Pure. Appl., 81 (2002), pp. 379–399.
[35] B. D. Scott, *Gyrokinetic field theory as a Gauge transform or: gyrokinetic theory without Lie transforms*, arXiv preprint, arXiv-1708.06265 (2017).
[36] C. Yang and F. Filbet, *Conservative and non-conservative methods based on Hermite weighted essentially non-oscillatory reconstruction for Vlasov equations*, J. Comput. Phys., 279 (2014), pp. 18–36.

Université de Toulouse III & IUF, UMR CNRS 5219, Institut de Mathématiques de Toulouse, 118 route de Narbonne, F-31062 Toulouse cedex, France
Email address: `francis.filbet@math.univ-toulouse.fr`

Université de Rennes 1 & IUF, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France
Email address: `luis-miguel.rodrigues@univ-rennes1.fr`

Université de Toulouse III, UMR CNRS 5219, Institut de Mathématiques de Toulouse, 118 route de Narbonne, F-31062 Toulouse cedex, France
Email address: `seyed_hamed.zakerzadeh@math.univ-toulouse.fr`