

# Exactly Energy Conserving Semi-Implicit Particle in Cell Formulation

Giovanni Lapenta<sup>a,\*</sup>

*<sup>a</sup>Department of Mathematics, KU Leuven, University of Leuven, Belgium*

---

## Abstract

We report a new particle in cell (PIC) method based on the semi-implicit approach. The novelty of the new method is that unlike any of its semi-implicit predecessors at the same time retains the explicit computational cycle and conserves energy exactly. Recent research has presented fully implicit methods where energy conservation is obtained as part of a non linear iteration procedure. The new method (referred to as Energy Conserving Semi-Implicit Method, ECSIM), instead, does not require any non-linear iteration and its computational cycle is similar to that of explicit PIC.

The properties of the new method are: i) it conserves energy exactly to round-off for any time step or grid spacing; ii) it is unconditionally stable in time, freeing the user from the need to resolve the electron plasma frequency and allowing the user to select any desired time step; iii) it eliminates the constraint of the finite grid instability, allowing the user to select any desired resolution without being forced to resolve the Debye length; iv) the particle mover has a computational complexity identical to that of the explicit PIC, only the field solver has an increased computational cost.

---

\*Corresponding author.  
*E-mail address:* giovanni.lapenta@kuleuven.be

The new ECSIM is tested in a number of benchmarks where accuracy and computational performance are tested.

*Keywords:* particle in cell (PIC), electromagnetic, electrostatic, semi implicit particle in cell, exactly energy conserving

---

## 1. Introduction

The particle in cell (PIC) method has affirmed itself as one of the most widely used numerical methods to model plasmas at the kinetic level [1, 2, 3]. The method is often selected for its simplicity that makes any new user or developer readily understand the spirit and implementation. The simplicity is also a great advantage when considering computer science issues relative to the implementation of PIC. Widely different approaches can be quickly implemented to test new software practices and new computer architectures [4, 5, 6, 7].

While the PIC method has proven over the last several decades its ability of giving the right answers in many scientific and engineering problems, a number of further improvements are being studied. Prominent among them is the conservation of energy.

The most common implementation of the PIC method is the classic explicit PIC presented in every textbook. The idea is to break the link between particles and fields for the duration of one time step. In each time step, the particles are advanced in the old fields and afterwards the fields are advanced with the particle information just updated. This approach is tremendously simple and successful but at one cost: it does not conserve energy. In fact energy tends to grow in time. This limitation is overcome in practice by

judiciously choosing the time step. With sufficiently small, sometimes really very small [8], time steps, energy can be conserved to an acceptable degree.

Recent research has focused on developing a new approach to PIC where energy is exactly conserved[9, 10, 11] (we refer to this approach ECPIC in the following). This important result is obtained at the cost of making the PIC scheme fully implicit, meaning that the particle equations and the field equations have to be solved together, coupled via a non-linear Newton or Picard iteration. Energy conservation is in principle exact and its accuracy in practice is controlled by the tolerance of the iterative Newton-Krylov solver used.

The question we ask here is: can energy conservation be achieved in a scheme that retains the feel of explicit PIC avoiding the non-linear iteration? We ask then if one can retain the marching order of explicit PIC: advancing for one time step the particles first and then the fields, then the particles again, without any iteration.

The answer we find is yes. We can, to our knowledge for the first time, design a method that retains that very desirable simplicity of explicit schemes and yet conserves energy exactly. A way to reach this goal is to use the new implementation of the semi-implicit PIC method presented here, called energy conserving semi-implicit method (ECSIM).

Semi-implicit methods retain some of the particle coupling in the field equations through extra terms in the Maxwell's equations that provide a linear approximation to the particle response. The method still advances in its computational cycle as an explicit method but each step in the cycle is somewhat more complex than in the explicit PIC method. Two types of

semi-implicit methods have been reported in the literature: the direct implicit method (DIM), developed at Lawrence Livermore National Laboratory and implemented in codes such as AVANTI [12, 13] and LSP [14] and the implicit moment method (IMM) developed at Los Alamos National Laboratory [15, 16] and used in a family of codes originating with Venus [15], followed by Celeste3D [17, 18], Parsek2D [19] and now used in Parsek2D-MLMD [20, 21] and iPic3D [22].

Compared with these previous semi-implicit methods, we modify significantly the particle mover and the mathematical procedure used to derive the field equations. The mover is replaced with a new mover that does not require any inner iteration and is in its practical implementation of the same level of complexity as the mover of explicit PIC. The field solver on the other hand is significantly more advanced than in the standard explicit methods requiring a significant amount of extra CPU effort. But it remains formally simple, similar to previous semi-implicit methods. Recalling that the cost of the PIC method is primarily coming from the particle mover, the new method has much potential in terms of computing performance.

The previous semi-implicit methods had many good properties of stability and ability to handle multiple time scale problems [16, 15, 13]. But they did not conserve energy [23, 24]. The new method reported here, instead conserves energy exactly.

The practical implication of full energy conservation will be fully assessed in real world applications that we hope future research using the new method will attempt. We report here results based on a simple but complete MATLAB implementation. Using it we show the following points:

1. The method indeed conserves energy exactly as the mathematical proof we provide promises. We use a direct solver so the conservation is exact to round-off. When iterative linear solvers are used, the accuracy of energy conservation would be controlled by the tolerance of the linear solver.
2. The so-called finite grid instability that in explicit PIC forces users to resolve the Debye length can be completely overcome. We show examples where the grid spacing is 16 orders of magnitude larger than the Debye length and the new method continues to remain stable and well behaved. 16 orders is not a limit, we just stopped out of fatigue. There is no reason to believe the system size could not be increased further while still retaining stability.
3. The anisotropic heating and cooling of standard PIC or even of semi-implicit PIC is completely avoided, just as in fully implicit PIC
4. The complexity (in terms of number of cpu operations) of the new method is of the same order as in previous semi-implicit PIC

Below, Section 2 is the central section of the present manuscript and derives the new method, the ECSIM. Section 3 reports the formal proof of energy conservation, exact and to round off. Section 4 provides several benchmarks that confirm the ability of the new method to conserve energy exactly. Section 5 focuses on some additional properties relative to the ability of ECSIM to eradicate completely the grid spacing limitation of the finite grid instability and the unphysical and anisotropic heating/cooling of species. Section 6 draws the conclusions and suggest further needed research.

The appendices report the details of the stability analysis of the new

method (Appendix A), the details of the spatial discretization used in the tests reported in Section 4 and 5 (Appendix B) and the order of accuracy in space and time of the specific implementation used here (Appendix C). Appendix D summarizes all the equations comprising the computational cycle of the new method, providing a concise summary to the interested code developer.

## 2. Energy Conserving Moment Implicit Method

Our goal is to derive a variant of the semi-implicit method that conserves energy exactly, borrowing this important property from fully implicit ECPIC, but avoiding the need for a non-linear iteration between particles and fields. Like previous semi-implicit methods, we still rely on linear field-particle coupling to retain the explicit nature of the PIC computational cycle.

Our starting point is a mover that combines the DIM  $D_1$  scheme [12] with the IMM and ECPIC  $\theta$ -scheme [15]. We advance the particle position as in the  $D_1$  scheme [13], but advance the velocity according to the  $\theta$  scheme [17].

This is the crucial innovation that allows energy conservation and complete stability of the method. We evaluate the fields at the time staggered particle position  $\mathbf{x}_p^{n+1/2}$  known explicitly from the time staggering of the leap-frog but we use the electric field at the  $\theta$  time level:  $\mathbf{E}^{n+\theta}$  exactly as in the standard IMM. The mover has the same stability properties of the IMM (Appendix A) but requires no predictor-corrector iteration between the particle position and particle velocity. This feature is what allows us to write the scheme to be exactly energy conserving. The order of accuracy of the mover is second order in the coupling between the particle position and velocity and

in the coupling with the electric field (if  $\theta = 1/2$ ) but it is first order in the coupling with the magnetic field (Appendix C reports more details).

The proposed new mover is:

$$\begin{aligned} \mathbf{x}_p^{n+1/2} &= \mathbf{x}_p^{n-1/2} + \Delta t \mathbf{v}_p^n \\ \mathbf{v}_p^{n+1} &= \mathbf{v}_p^n + \frac{q_p \Delta t}{m_p} (\mathbf{E}^{n+\theta}(\mathbf{x}_p^{n+1/2}) + \bar{\mathbf{v}}_p \times \mathbf{B}^n(\mathbf{x}_p^{n+1/2})) \end{aligned} \tag{1}$$

where  $\bar{\mathbf{v}}_p = (\mathbf{v}_p^{n+1} + \mathbf{v}_p^n)/2$ . As can be seen the first equation is the same as in the  $D_1$  scheme (but in this regard, identical also to the simple leap-frog scheme), it can be solved using the known velocity from the previous time step. The second equation is identical to that of the  $\theta$ -scheme but the electric and magnetic fields are computed at the known position  $\mathbf{x}_p^{n+1/2}$  rather than at the unknown position  $\bar{\mathbf{x}}_p$  that in the standard IMM requires the predictor-corrector iteration. These two positions are conceptually the same, they express the particle position at the mid-time between the old and new evaluations of the velocity. But one is computed explicitly while the other is computed as part of a predictor-correctors iteration. Both are second order accurate but the scheme in eq. (1) is simpler to compute.

For stability what counts is how the acceleration is computed. The magnetic field term is unconditionally stable even in the standard Boris mover [2]. The term possibly giving a stability constraint is the electric field term. Similarly, the magnetic field does not do any work on the particles and causes no worry in energy conservation, only the electric field term does. For these reasons the force term is written using the magnetic field at the initial time level  $\mathbf{B}^n(\mathbf{x}_p^{n+1/2})$  but the electric field is written at the advanced intermediate level  $\mathbf{E}^{n+\theta}(\mathbf{x}_p^{n+1/2})$ . This same approach is used in the standard  $\theta$  scheme so

the stability analysis remains the same as in the  $\theta$  scheme and the method is unconditionally stable, as we show in Appendix A.

The fields at the particle positions are computed by interpolation:

$$\mathbf{E}_p^{n+\theta} = \mathbf{E}^{n+\theta}(\mathbf{x}_p^{n+1/2}) = \sum_g \mathbf{E}_g^{n+\theta} W(\mathbf{x}_p^{n+1/2} - \mathbf{x}_g) \quad (2)$$

where we have assumed that the field equations are discretized on a grid with a generic index  $g$  (that can be for example the ensemble of the indices  $i, j, k$  in 3D Cartesian grids). Appendix B reports the actual discretization used in the examples below. A similar expression holds for the magnetic field:

$$\mathbf{B}_p^{n+\theta} = \mathbf{B}^{n+\theta}(\mathbf{x}_p^{n+1/2}) = \sum_g \mathbf{B}_g^{n+\theta} W(\mathbf{x}_p^{n+1/2} - \mathbf{x}_g) \quad (3)$$

A short hand notation has been introduced:  $\mathbf{B}_p^{n+\theta} = \mathbf{B}^{n+\theta}(\mathbf{x}_p^{n+1/2})$  and  $\mathbf{E}_p^{n+\theta} = \mathbf{E}^{n+\theta}(\mathbf{x}_p^{n+1/2})$ .

The interpolation function  $W$  can be chosen in many different ways but we consider here the case of b-splines of order  $\ell$  [25] and we use the same interpolation functions for both electric and magnetic field (some methods use different interpolations for the two). For example in 3D Cartesian grids this becomes:

$$W(\mathbf{x}_p - \mathbf{x}_g) = b_\ell(x_p - x_g) b_\ell(y_p - y_g) b_\ell(z_p - z_g) \quad (4)$$

We will use b-spline of order one. This expression reduces trivially in 1D for the examples reported below.

For the Maxwell's equation we use the same discretization as in the standard IMM [22]. The two curl Maxwell equations are discretized in time with



another  $\theta$ -scheme:

$$\begin{aligned}\nabla_g \times \mathbf{E}^{n+\theta} + \frac{1}{c} \frac{\mathbf{B}_g^{n+1} - \mathbf{B}_g^n}{\Delta t} &= 0 \\ \nabla_g \times \mathbf{B}^{n+\theta} - \frac{1}{c} \frac{\mathbf{E}_g^{n+1} - \mathbf{E}_g^n}{\Delta t} &= \frac{4\pi}{c} \bar{\mathbf{J}}_g\end{aligned}\tag{5}$$

The spatial operators in eq. (5) are discretized on a grid. Consistent with the notation introduced earlier, we index the grid elements generically with the index  $g$ .  $\nabla_g$  is a shorthand for the spatial discretization used. For example in the IMM family of codes the discretization used is different from the Yee scheme [26] and locates on the cell centers the magnetic field and on the cell vertices for the electric field [27, 22, 28]. Appendix B reports the actual discretization used in the examples below. But all the derivations below are not critically dependent on which spatial discretization is used.

The set of Maxwell's and Newton's equations are coupled. Eq. (5) requires the new particle velocities to compute the current and eq.(1) requires the new advanced electric field to move the particles. In the spirit of the semi-implicit method, we do not want to solve two coupled sets with a single non-linear iteration and find instead a way to extract analytically from the equations of motion the information needed for computing the current without first moving the particles. In previous semi-implicit methods this is done via a linearization procedure. The new mover used here allows us, instead, to derive the current rigorously without any approximation.

Form its definition, the current for each species in each grid location is:

$$\bar{\mathbf{J}}_{sg} = \frac{1}{V_g} \sum_{p \in s} q_p \bar{\mathbf{v}}_p W(\mathbf{x}_p^{n+1/2} - \mathbf{x}_g)\tag{6}$$

where the summation is over the particles of the same species, labeled by  $s$ .

With easy vector manipulations [17], the velocity equation can be rewritten in the equivalent form:

$$\bar{\mathbf{v}}_p = \hat{\mathbf{v}}_p + \beta_s \hat{\mathbf{E}}_p \quad (7)$$

where hatted quantities have been rotated by the magnetic field:

$$\hat{\mathbf{v}}_p = \alpha_p^n \mathbf{v}_p^n \quad (8)$$

$$\hat{\mathbf{E}}_p = \alpha_p^n \mathbf{E}_p^{n+\theta}$$

via a rotation matrix  $\alpha_p^n$  defined as:

$$\alpha_p^n = \frac{1}{1 + (\beta_s B_p^n)^2} (\mathbb{I} - \beta_s \mathbb{I} \times \mathbf{B}_p^n + \beta_s^2 \mathbf{B}_p^n \mathbf{B}_p^n) \quad (9)$$

where  $\mathbb{I}$  is the dyadic tensor (matrix with diagonal of 1) and  $\beta_s = q_p \Delta t / 2m_p$  (independent of the particle weight and unique to a given species). The elements of the rotation matrix are indicated as  $\alpha_p^{ij,n}$  with label  $i$  and  $j$  referring to the 3 components of the vector space ( $x, y, z$ ).

Substituting then eq. (7) into eq. (6), we obtain without any approximation or linearization:

$$\bar{\mathbf{J}}_{sg} = \frac{1}{V_g} \sum_p q_p \hat{\mathbf{v}}_p W_{pg} + \frac{\beta_s}{V_g} \sum_p q_p \hat{\mathbf{E}}_p^{n+\theta} W_{pg} \quad (10)$$

where we shortened the notation  $W_{pg} = W(\mathbf{x}_p^{n+1/2} - \mathbf{x}_g)$  and the summation is intended over all particles of species  $s$ .

We introduce the following hatted current:

$$\hat{\mathbf{J}}_{sg} = \sum_p q_p \hat{\mathbf{v}}_p W_{pg} \quad (11)$$

with the obvious meaning of current based on the hatted velocities.

Using eq. (8) and recalling the definition of eq. (11), the expression for the current becomes:

$$\bar{\mathbf{J}}_{sg} = \hat{\mathbf{J}}_{sg} + \frac{\beta_s}{V_g} \sum_p q_p \alpha_p^n \mathbf{E}_p^{n+\theta} W_{pg} \quad (12)$$

Computing then the electric field on the particles by interpolation from the grid as in eq. (2), it follows that:

$$\bar{\mathbf{J}}_{sg} = \hat{\mathbf{J}}_{sg} + \frac{\beta_s}{V_g} \sum_p \sum_{g'} q_p \alpha_p^n \mathbf{E}_{g'}^{n+\theta} W_{pg'} W_{pg} \quad (13)$$

Exchanging the order of summation and introducing the elements of the mass matrices [29] as:

$$M_{s,gg'}^{ij} = \sum_p q_p \alpha_p^{ij,n} W_{pg'} W_{pg} \quad (14)$$

we obtain, in matrix form:

$$\bar{\mathbf{J}}_{sg} = \hat{\mathbf{J}}_{sg} + \frac{\beta_s}{V_g} \sum_{g'} M_{s,gg'} \mathbf{E}_{g'}^{n+\theta} \quad (15)$$

In matrix notation the  $3v$  mass matrices are written as  $M_{gg'}$ , that is without the indices  $i, j$  for the vector directions.

Eq. (14) defines the elements of the mass matrices  $M_{s,gg'}$  that are the most peculiar characteristic of the method proposed here. There are  $3v$  such matrices, where  $v$  is the dimensionality of the magnetic field and velocity vector, not to be confused with the dimensionality of the geometry used for space  $d$ . The indices  $i$  and  $j$  in eq. (14) vary in the  $3v$ -space. For example for full 3-components vectors,  $i, j = 1, 2, 3$  and there are 9 mass matrices. Each matrix is symmetric and very sparse with just  $2d$  diagonals.

Eq. (15) allows us to retain symmetry between field interpolation to the particles and current interpolation to the grid: with it the computation of the energy exchange between the plasma species and the fields is identical when computed on the particles and when computed on the grid. With this formula we have an explicit linear link between the advanced current at the mid-point of the time step and the electric field at the advanced time. This linear relationship can be substituted into the discretized Maxwell's equations (5) to form a linear set of equations to be solved on the grid:

$$\begin{cases} \nabla_g \times \mathbf{E}^{n+\theta} + \frac{1}{c} \frac{\mathbf{B}^{n+1} - \mathbf{B}^n}{\Delta t} = 0 \\ \nabla_g \times \mathbf{B}^{n+\theta} - \frac{1}{c} \frac{\mathbf{E}^{n+1} - \mathbf{E}^n}{\Delta t} = \frac{4\pi}{c} \left( \hat{\mathbf{J}}_g + \sum_{g'} M_{gg'} \mathbf{E}_{g'}^{n+\theta} \right) \end{cases} \quad (16)$$

where we have introduced the total current  $\hat{\mathbf{J}}_g = \sum_s \hat{\mathbf{J}}_{sg}$  and the species summed mass matrices, that written by elements are:

$$M_{gg'}^{ij} = \sum_s \frac{\beta_s}{V_g} M_{s,gg'}^{ij} \quad (17)$$

that can be more usefully held in memory, reducing the memory consumption by a factor equal to the number of species.

The linear relationship in eq. (15) represents the response of the particles to changes in the fields. Using it in the Maxwell's equations allows one to step over the time scales of wave characteristics if needed, retaining the implicit nature of the scheme. This feature is in stark variance with the fully implicit ECPIC where the full non-linear coupling is kept. The linear response represented by eq. (15) is sufficient to eliminate the stability requirements. In

ECPIC is also possible to use similar linear relationships as preconditions [? ].

This same advantage is also present in all previously known semi-implicit methods. But formula (15) has one critical new advantage over all other previous semi-implicit methods: it leads to exact energy conservation, just as in the fully implicit ECPIC but without its non-linear complexity.

Before moving on to the formal proof of the energy conservation, it is important to realize fully the relative complexity of the new ECSIM with respect to the other schemes available.

The particle mover is identical in complexity to the standard Boris scheme of the explicit PIC. In particular it does not need the fixed number of PC iterations used in the IMM as implemented for example in iPic3D [22]. In the typical iPic3D run, 3 steps of PC are used, although other iteration schemes are also possible [30, 31]. This new method is then exactly 1/3 in cost in terms of the particle mover step, and by this we include both the motion itself and the field interpolation to the particles. The particle mover is by far the largest user of CPU time in the typical cycle. This reduction in cost is thus very significant.

The field solver is identical to that of other semi-implicit methods. iPic3D for example uses GMRES as solver for the Maxwell's equations. The choice of GMRES is suggested by the non-symmetric nature of the Maxwell matrix and by the diagonal dominance of the matrix which allows a rapid convergence of GMRES. This is in sharp contrast to the solution of elliptic problems, such as the Poisson's equation in electrostatic PIC. In that case, Krylov schemes converge very slowly, requiring the deployment of suitable preconditioner

[32, 33]. The new ECSIM has one extra cost not present in the standard IMM. The mass matrices need to be evaluated but the extra CPU time and memory is manageable.

These matrices are computed each time step and used in the solution of the Maxwell equations. The mass matrices are themselves diagonally dominant a feature that retains the diagonal dominance of the overall Maxwell solver leading to a rapid convergence of iterative solvers.

### 3. Formal Proof of Energy Conservation

Energy conservation means that the amount of energy transferred between the particles and the fields is exactly the same when measured from the particle equations of motions or when measured from the field equations. The result applies only when  $\theta = 1/2$ . Only in that case energy is in fact conserved. A fact we now proceed to prove.

Starting from the particle equations of motions, eq.(1), dotting the velocity equation by the average speed,  $\bar{\mathbf{v}}_p$ , and summing over all particles we obtain:

$$\frac{1}{2} \sum_p (m_p (\mathbf{v}_p^{n+1})^2 - (\mathbf{v}_p^n)^2) = \Delta t \sum_p \left( q_p \sum_g \bar{\mathbf{v}}_p \cdot \bar{\mathbf{E}}_g W_{pg} \right) \quad (18)$$

where the electric field is computed as average consistent with the choice  $\theta = 1/2$  and the magnetic field drops out as obvious from the properties of the cross product. Exchanging the summation over particles and cells, which is perfectly legitimate for finite sums without rising any convergence issue, we obtain:

$$\frac{1}{2} \sum_p (m_p (\mathbf{v}_p^{n+1})^2 - (\mathbf{v}_p^n)^2) = \Delta t \sum_g \bar{\mathbf{J}}_g \cdot \bar{\mathbf{E}}_g \quad (19)$$

where it is recognized that  $\bar{\mathbf{J}}_g = \sum_p q_p \bar{\mathbf{v}}_p W_{pg}$ .

Starting now from the discretized Maxwell equations, eq. (5), and assuming again  $\theta = 1/2$ , we multiply the first equation by  $\bar{\mathbf{B}}_g$  and the second by  $\bar{\mathbf{E}}_g$ . Summing the two equations it follows that:

$$\frac{(\mathbf{B}_g^{n+1})^2 - (\mathbf{B}_g^n)^2}{2c} + \frac{(\mathbf{E}_g^{n+1})^2 - (\mathbf{E}_g^n)^2}{2c} = \Delta t \left( \frac{4\pi}{c} \bar{\mathbf{J}}_g \cdot \bar{\mathbf{E}}_g + \mathbf{E}_g \cdot \nabla_g \times \mathbf{B} - \mathbf{B}_g \cdot \nabla_g \times \mathbf{E} \right) \quad (20)$$

Summing over all grid points and using the properties of the curl operator that are assumed to be replicated by its grid discretization, it follows that:

$$\sum_g \frac{(\mathbf{B}_g^{n+1})^2 - (\mathbf{B}_g^n)^2}{4\pi} + \sum_g \frac{(\mathbf{E}_g^{n+1})^2 - (\mathbf{E}_g^n)^2}{4\pi} = \Delta t \sum_g \bar{\mathbf{J}}_g \cdot \bar{\mathbf{E}}_g + \frac{c\Delta t}{4\pi} \sum_g \nabla_g \cdot (\mathbf{E}_g \times \mathbf{B}_g) \quad (21)$$

The two terms on the left hand side are recognized as the variation of the magnetic and electric energy as measured on the grid. The first term on the right hand side is the energy exchange term with the particles and the last term expresses the divergence of the Poynting flux. This is the usual energy equation for electromagnetic energy, discretized on the grid.

For energy to be conserved in the system, the energy exchange term on the particle equations (eq. (19)) needs to be identical to that on the field equations, (eq. (21)). This term is indeed identically equal to  $\sum_g \bar{\mathbf{J}}_g \cdot \bar{\mathbf{E}}_g$  in both equations. Energy conservation is enforced exactly, to round off.

The proof above is based on two key elements. First, the current in the field equations, eq. (19), must be computed based on the mean speeds

$\bar{\nabla}_p$ . This is the guiding principle followed by the fully implicit ECPIC and by the new method presented here. Second, the curl operator on the grid must retain the property of the continuum curl operator that  $\nabla \cdot (\mathbf{A} \times \mathbf{B}) = \mathbf{B} \cdot \nabla \times \mathbf{A} - \mathbf{A} \cdot \nabla \times \mathbf{B}$ . Not all discretizations would have this property but most discretization used in practice have this mimetic property. The discretization used in the example section below does for example.

This proof just reported is confirmed by the results below that confirm energy is conserved to round off.

#### 4. Benchmarks

We have implemented the method described above in two different programs, written in MATLAB. Obviously, production codes based on the new ECSIM formulation will eventually be implemented in more efficient and more easily parallelizable languages.

Below we report results of test cases from two codes. The first is electrostatic and uses 1 dimension in space,  $x$  as well as one dimension in velocity,  $v_x$ . The electric field and the current also has one component. In this case, in absence of a magnetic field the mass matrices are all the same and only one mass matrix needs to be computed, instead of 9. This condition is referred to as 1D1V.

The second program is still one dimensional in space but it has full 3 components of the particle velocity, of the current and of the fields. Obviously, in 1D with the only non-ignorable coordinate being  $x$ , the  $B_x$  component is constant in time (because the curl operator has no  $x$  component) and uniform in space (to satisfy the divergence condition) but not necessarily 0. In this



case all 9 mass matrices need to be computed. This second case is referred to as 1D3V.

In all cases, the ions are a fixed uniform background and only the electrons are followed. All tests are performed on a Apple MacBook Pro with processor Intel Core i7 running at 2.6 GHz and with 16GB of memory.

#### *4.1. Electrostatic Example: Two-Stream Instability*

A classic benchmark for electrostatic kinetic problems is the two stream instability. The initial condition is formed by a uniform plasma where the particles are distributed in two Maxwellian beams with thermal speed  $v_{th}/c = 0.01$  and speed  $v_0/c = \pm 0.2$  and with equal density  $n_0/2$ . As mentioned above the system evolution is followed in 1D1V with a system size of  $L = 2\pi$  and periodic boundary conditions. The analytical solution for the dispersion relation is well known [34]: longitudinal modes with  $kv_0/\omega_{pe0} < 1$  are unstable and the two beams become destabilized and mix. In the present case, the fastest growing mode has  $m = 3$  (corresponding to  $k = 2\pi m/L$ ) and its growth rate is  $\gamma/\omega_{pe} = 0.35$ . In the initialization, the particle velocity is displaced by a small ( $\Xi = 0.1$ ) amount to favor the onset of the instability:

$$\delta v_p = \Xi v_{th} \sin(2\pi m x_p / L)$$

This problem is extremely repeatable and the evolution of the electrostatic energy has been often used in the literature to assess the correctness of a PIC implementation. We compare four runs described below:

1. The present ECSIM implemented in a MATLAB code.
2. The fully implicit ECPIC described in [9], implemented also in MATLAB and thus providing also a direct comparison in terms of language

of implementation. The tolerance for the Newton-Krylov iteration is set to  $10^{-7}$ .

3. The standard explicit PIC based on the classic textbook implementation, also implemented in MATLAB for direct performance comparison.
4. The standard electrostatic implementation of the IMM similar to that used in the implementations of Venus, Celeste, Parsek2D and iPic3D. Here we use an electrostatic implementation in MATLAB for consistency with the other runs.

The grid discretization scheme used in all codes above is summarized in Appendix B. All codes use the same time step  $\omega_{pe}\Delta t = 0.125$  for 1000 cycles and the same number of cells  $N = 64$  and particles per cell  $N_{ppc} = 154$ . Such low resolution might give the reader a pause. This is a benchmark and this number of cells has only the purpose of testing the method at low resolution. We want to assess the ability of energy conserving schemes to work even at low resolution where non-energy conserving codes start to see degraded results.

Figure 1 summarizes all the results from the four different simulations. The results are all similar and all are in good agreement with linear theory.

Figure 2 reports the error in the total energy of the same runs. More in detail, Table 1 reports quantitatively the execution times and energy conservation statistics.

Energy conservation of the ECPIC and ECSIM is perfect. The fully implicit ECPIC conserves energy to the level of accuracy in the Newton iteration but the ECSIM conserves energy to machine precision. This difference is evident at times  $\omega_{pe}t > 10$  when the error in the fully energy conserving

risks to the level of  $10^{-9}$  which corresponds to the allowed tolerance in the Newton iteration. The difference between the accuracy of the ECPIC and ECSIM is not of any great practical consequence and it is due to the fact that the ECSIM solves the matrix for the Maxwell equation with direct methods that reduce the residual to machine precision.

The explicit method by comparison has an acceptable energy error of less than 1%. But 1% of energy error is still an energy error whose consequences need to be verified in real world simulations to avoid it leading to possible unphysical effects [8].

In the present case, the standard implicit moment method fares worse in terms of energy conservation. Using the new ECSIM is especially important in electrostatic cases where the standard IMM is notoriously challenged.

Table 1: Table of cost effectiveness of Energy conservation in different PIC implementations. Four runs are compared for the same problem of the two-stream instability: the ECSIM presented here, the standard explicit PIC, the fully implicit ECPIC and the standard IMM.

Method	CPU time (s)	$\max  \Delta E /E(0)$
ECSIM	3.47	8.8057e-15
Explicit PIC	2.14	0.0099
EC-PIC	68.78	1.5283e-09
Standard IM	2.43	0.0281

The computation time of the explicit method is smallest. The actual timing observed are just indicative because we made no specific attempt to optimize any of the methods. All methods are implemented in MATLAB

using common programming practices appropriate for that language. Each method can be further optimized and future research can focus on the best optimization method for the new approach to be compared with the others. Especially important is the point that ECPIC can be made much more efficient with preconditioning [? ]: the timing comparison provided below is not indicative then of the expected performance of a preconditioned ECPIC.

As noted above the ECSIM is still an explicit method, the additional cost compared with explicit PIC is the computation of two diagonals for the mass matrix. The matrix is symmetric and 3-diagonal so only 2 diagonals need to be computed. As noted above in absence of magnetic fields there is only one mass matrix rather than 9. The cost of computing the mass matrix is proportional to the number of particles and is comparable to that of an interpolation step in the explicit PIC. This cost is not negligible but the tests above show that even with it the computational time does not increase disproportionately. Future research will need to study the most effective way of coding the mass matrix implementation in modern architectures, using for example vectorization.

Figure 3 reports the error in the momentum conservation as a fraction of the initial momentum of each beam. In the ECSIM momentum is not conserved exactly, just as in the full ECPIC and in the IMM. Unfortunately conservation of energy and momentum appear to be mutually exclusive [23]. In the results reported, the non conservation of momentum is a tiny fraction of the initial momentum of each beam and it does not accumulate in time. The error is of the order of one part per a thousand for each particle or for the total momentum, since there is no tendency for the error to build up. For

the typical applications to space and astrophysical plasmas an error of 1 part per thousand on the momentum is completely negligible. There might be applications where momentum conservation is more important than energy conservation and in that case the explicit momentum conserving method has an advantage.

Another perspective on this is to evaluate the rate of change in the momentum and assess the numerical collisionality that is implicitly leading to that. We can define a numerical collision frequency  $\nu$  from:

$$\frac{dP}{dt} = -\nu P \quad (22)$$

where  $P$  is of the momentum (1D in the present example). This equation needs to be considered with care because the value of  $\nu$  is actually oscillating in time and leads to a sort of conservation when averaged in time, without a diffusive trend leading to an overall decay of momentum. Also in terms of numerical collisionality the effect is of the order of less than one part per thousand without leading to a secular trend.

For this reason, Figure 3 reports the standard deviation of  $\nu$  as it oscillates around 0. This quantifies a random walk with no average trend but a standard deviation.

#### *4.2. Electromagnetic Example: Transverse Beam (Weibel) Instability*

Two counter streaming beams excite also a transverse instability, called Weibel or filamentation, present also in the case of temperature anisotropy [35, 36]. The instability leads to the excitement of a magnetic field and it is considered one of the fundamental process in Nature (alongside the Bierman battery process) that have allowed the primordial formation of a magnetic

field in the Universe. It is a form of dynamo driven by anisotropic velocity distributions due to temperature anisotropy or beams.

We consider here an initially uniform and unmagnetized plasma in a 1D domain of size  $L/d_e = 2\pi$  discretized in 64 cells with 154 particles per cell and a time step size of  $\omega_{pe}\Delta t = .125$ . The beams are counterstreaming with speed  $v_0/c = 0.8$  with a thermal spread of each beam in each direction corresponding to  $v_{th0}/c = 0.01$ . Relativistic effects are moderate at these speeds and we can use the classical methods described above.

We focus here on a choice of grid and time step where  $c\Delta t > \Delta x$  and exclude explicit PIC, focusing only on implicit and semi-implicit codes. We use:

1. The present ECSIM method implemented in a MATLAB code for the fully electromagnetic case.
2. The electromagnetic fully implicit EC-PIC described in [9], implemented also in MATLAB. The tolerance for the Newton-Krylov iteration is set to  $10^{-8}$  in the present case.
3. The standard electromagnetic implementation of the IMM similar to that used in the implementations of Venus, Celeste, Parsek2D and iPic3D. Here we use a fully electromagnetic implementation in MATLAB for consistency with the other runs.

The codes used here are different from those of the previous electrostatic example, and are fully electromagnetic and 1D3V.

All modes are unstable and we focus especially on the growth of mode  $m = 3$ . Figure 4 reports the evolution in time of the mode for different runs. Obviously all runs considered agree with each other and correctly reproduce

the growth rate predicted by linear theory.

Again energy conservation is perfect in the ECSIM and is determined by the accuracy of the Newton iteration tolerance in the ECPIC. The standard IMM has a good conservation but not precise. Table 2 reports the quantitative information on the energy conservation of the different runs, confirming the qualitative impression of Fig. 4.

Table 2 reports also the computational cost of the different runs. Two aspects might appear surprising. First, the fully implicit ECPIC is not much more expensive than the standard IMM. This effect is due to this specific case that does not tax the Newton solver. The standard IMM uses a fixed 3 iteration PC scheme for the particle equations of motion so its cost is not much smaller than that of the Newton iteration in this case.

The second surprising aspect is that the ECSIM is much less expensive than the standard IMM. The surprise in this is that the ECSIM has the extra cost of computing the 9 mass matrices needed in the scheme. This part costs about 30s of CPU time. But this extra cost not present in the standard IMM is more than compensated for by the fact that the new scheme does not require any PC mover and retains the same numerical cost of a straight through explicit leap-frog scheme. Matrix operations are grid operations and usually in PIC tend to be a minority of the cost, while the lion share is done by the particles.

## 5. Elimination of the Finite Grid Instability

To consider the ability of the ECSIM to deal with large system sizes, under-resolving the small scales is tested considering a box of size  $L = 2\pi\Xi$

Table 2: Table of cost effectiveness of energy conservation in different PIC implementations. Three runs are compared for the same problem of the transverse instability: the ECSIM presented here, the fully implicit ECPIC [9] and the standard IMM [22].

Method	CPU time (s)	$\max  \Delta E /E(0)$
ECSIM	31.87	8.6119e-15
EC-PIC	137.59	2.8467e-08
Standard IM	107.28	0.0189

divided into 64 cells and using 10,000 particles distributed initially uniformly. For small  $\Xi$ , the grid resolves well the Debye length,  $\Delta x/\lambda_{De} = \omega_{pe} L 2\pi \Xi / V_{th} 64$ . But as  $\Xi$  is increased the grid spacing progressively exceed more and more the stability constraint of explicit PIC:  $\Delta x/\lambda_{De} < \xi\pi$ .

We consider a uniform plasma with thermal speed  $V_{th} = 0.01c$  assuming immobile ions. The plasma is in thermal equilibrium and it is stable to all Vlasov-Poisson modes.  $\Xi$  is varied by 15 orders of magnitude, from 1 to  $10^{15}$ , resulting in  $\Delta x/\lambda_{De}$  ranging from 10 to  $10^{16}$ . The time step is selected to  $\omega_{pe}\Delta t = .125\Xi$  so that the CFL number is constant in all runs and equal to 0.0127. The results are presented in Fig. 6.

A second case is selected with  $\omega_{pe}\Delta t = 1.25\Xi$  with CFL= 0.127. The results of the second ensemble of runs are presented in Fig. 7.

In all cases the simulation remains completely stable. The total energy is exactly conserved and the temperature of the electrons remains exactly constant. This result is remarkable for the wide range of scales covered: the method is capable of describing a plasma over 15 orders of magnitude,



exceeding the explicit stability limit by 16 orders of magnitude. We stopped at 16, but there is no reason to believe this stable behavior would not continue indefinitely, allowing in principle to model any system size.

The behavior of the fluctuations is significant. As the system size is increased, the thermal fluctuations and the electric field fluctuations are nearly exactly overlapping, as required by the total energy conservation. But at large scales, both fluctuations become progressively smaller. This is the desirable behavior: as the system size becomes larger, the model should transition to a fluid macroscopic behavior where the electric field and thermal energy fluctuations are zero. It happens exactly as one should expect.

The thermal energy has a lower minimum near the machine precision for the linear direct solver,  $10^{-15}$  used to solve Maxwell's equations. Beyond this level, the thermal energy becomes limited by the total energy error while the electric field fluctuation continues to decrease. The momentum conservation is equally well behaved, with the momentum conservation improving as the system size is increased, another indication of convergence towards a fluid, noiseless, behavior.

Comparing the two different ensembles of runs with the two different CFL values, the results differ only in detail, presenting the same qualitative behavior. With larger CFL, meaning a larger time step for the same grid spacing, the transition towards reduced fluctuations and convergence towards a macroscopic fluid behavior happens at smaller spatial scales.

### *5.1. Artificial Cooling and Artificial Anisotropy*

A peculiar drawback of the lack of energy conservation is the tendency for the particles to heat or cool, anisotropically.

Explicit codes tend to heat, a tendency that can be especially deleterious providing extra free energy to drive numerical instabilities and to accelerate particles via numerical effects. The terminal stage of this illness being the finite grid instability.

Semi-Implicit codes tend to cool particles, even though in certain circumstances for large CFL numbers in non-equilibrium situations, numerical heating can be observed as well [24].

In both cases, the numerical energy error tends to be anisotropic, even though it tends to be gyrotropic. The temperature parallel and perpendicular can change due to numerical effects differently (anisotropy) while the distribution function in the direction perpendicular to the magnetic field remains symmetric (gyrotropic).

The method presented here, as well as the fully implicit methods [9, 10] completely eliminate this artificial source of anisotropy. To test this point we consider a condition very typical in the plasma of the Earth space environment (magnetosphere) [28, 37]. The initial electron population is assumed to be uniform and magnetized with  $\omega_{ce} = 0.005\omega_{pe}$  and  $v_{th0}/c = 0.01$ . The ions are assumed to be immobile. The magnetic field is directed along  $y$  and uniform.

We consider the typical resolution requested in a implicit simulation of the Earth environment:  $\Delta x/d_e = 0.5$  and  $\omega_{ce}\Delta t = 0.05$ . As a simple test we use a small 1D box with  $L/d_e = 32$  and use the 1D3V MATLAB codes described above.

Explicit codes need to resolve the Debye length, but as shown in the previous section, this is not the case of implicit and semi-implicit methods.

We assume the electron skin depth needs to be marginally resolved in the case of large magnetospheric simulations where reconnection takes place. The typical scale of the so-called electron diffusion region is of order  $d_e$  and needs to be resolved [38]. An accurate simulation would need to have a smaller grid spacing (see e.g. [39]) than the one chosen above, but when one pushes the limit for large scale simulations, the resolution set above would be reasonable [40, 41, 42].

Figure 8 reports the comparison of two simulations with the MATLAB implementations of the standard IMM and of the new ECSIM method. The total energy in the standard IMM method is poorly conserved. In the present case  $v_{th0}\Delta t/\Delta x = 0.2$  and the IMM is well within its stable range. The explicit method would not work because  $\omega_{pe}\Delta t = 10$  and  $\Delta x/\lambda_{De} = 50$ , the stability limits of explicit methods widely exceeded.

During the course of the run, the standard IMM loses over 10% of the energy and it does so all in the perpendicular plane. Eventually the numerical effect would bring the plasma in a regime where instabilities can be driven by the numerical artificial anisotropy. If this were a production run with iPic3D, we would simply have to cut the time step, improving energy conservation. The case selected above is especially pathological and an extreme case of lack of energy conservation not representative of any simulation in the previous literature. We selected it only to make the point clear. The typical workflow of a iPic3D simulation requires a preliminary study of the optimal  $\Delta t$  that allows the longest possible simulation with the largest possible  $\Delta t$ , while retaining sufficient energy conservation and avoiding artificial anisotropy of the type in Fig. 8. This concern is completely eliminated by the new ECSIM

making the use of the new ECSIM simpler. Of course, we have only done a few examples here and only experience will confirm if this initial assessment is correct and it will continue to hold in complex production runs.

## 6. Conclusions

We have introduced a new type of semi-implicit PIC method. The method is based on a new mover and a new approach to compute the plasma response using the mass matrix instead of the plasma susceptibility used in previous methods. The complete algorithm derived above is summarized in Appendix D.

Let us review what are the novelties and new properties of the new method.

- I We use a new mover. The new mover is a hybridization of the  $D_1$  mover of DIM and the  $\theta$  mover of IMM. It advances the particle positions just as in the leap-frog mover, stepping the positions over the velocity in a time centered, second order accurate, way. But it advances the velocity using an implicit expression for the electric field, computed as average between start and end of the time step. This feature makes the scheme unconditionally stable and implicit.
- II We handle the implicitness in the spirit of previous semi-implicit methods but we modify the procedure to linearize the coupling of particles and fields. We do so by introducing the mass matrices that express the particle response to the electric field. The presence of a magnetic field requires to used  $3v$  mass matrices, where  $v$  is the dimensionality of the fields and velocity vectors. The mass matrices express the plasma re-

sponse obtained from the particle equations of motion. Once substituted into the Maxwell equations, a complete self consistent linear problem is obtained for the fields.

- III The new method conserves energy exactly. We provide a direct mathematical proof and we test that in the examples presented energy is indeed conserved to round off.
- IV The new system is unconditionally stable. The only constraint is that the particles should not move too far in one time step so that the mover remains accurate. The same constraint as in the standard IMM remains valid,  $v_{the}\Delta t < \Delta x$ . The thermal speed should be changed with the average velocity for a fast moving plasma. This is an accuracy constraint not a stability constraint. With the new approach the current is computed exactly from the mover without the Taylor series expansion typical of the standard IMM. We do not need to worry about the convergence of the truncated series. The condition is purely one of accuracy. In the IMM, instead, violating this condition leads to severe energy error. In the ECSIM, the solution might not be accurate if  $v_{the}\Delta t > \Delta x$  but energy would still be exactly conserved.
- V A consequence of the finding above is that the need to resolve the Debye length to avoid the onset of the finite grid instability is completely eliminated. We have shown this point to an exaggerated degree proving stability and accuracy of the solution with cells up to 16 orders of magnitude larger than the Debye length. The ECSIM could model macroscopically large systems with coarse resolution, conserving energy to round off. This results is obtained for a simple maxwellian plasma

without drifts: other problems need to be tested to investigate how this property holds up in more complex setups that include drifts.

VI We have tested the method for accuracy and energy conservation in a few standard benchmark problems, electrostatic and electromagnetic. We reported two here, more have been done but are not reported.

VII The method presented is exactly energy conserving but it lacks two other conservation properties: momentum and charge. The future users of the method are recommended to monitor if those quantities are well conserved in realistic applications.

VIII The numerical complexity of the ECSIM has been compared with that of the explicit PIC, semi-implicit standard IMM, fully implicit ECPIC. Comparing with explicit PIC, the method has the same complexity in term of particle motion, the most expensive operation for a PIC. But it has the extra cost of the linear solution of the Maxwell equations and the computation of the mass matrices. When compared with standard IMM, the complexity is reduced by a factor of 3 for the particle mover because the ECSIM mover has the same complexity of the leap-frog and does not require the PC iterations of the standard IMM. The Maxwell solver has the same complexity but the ECSIM requires to compute  $3v$  extra matrices. Overall we find that the reduced cost of the mover outweighs the cost of the matrix computation. To make a full performance comparison with ECPIC, one needs to include the effect of preconditioners [? ]. In the present work we have used an unpreconditioned ECPIC but future work will investigate the issue more in depth.

Further research is ongoing to build on the results presented above. First,

we are looking at the conservation of other quantities, such as the local divergence of the electric field (Gauss theorem) and that of momentum. Both can be improved if the velocity update is done over substeps [43] rather than in a single jump as in the mover used now (1).

However, the most important future development is implementing the method in a 3D production code. The new method is sufficiently simple that can be readily implemented in new codes or adapted to existing codes. Appendix D summarizes all the steps needed for readers interested in coding the new method. The particle mover is very similar to that of explicit PIC and simpler than in the previous IMM. The field solver can be easily implemented with the now standard GMRES matrix solver.

## Acknowledgments

The present work is supported by the European Commission via the Deep-ER project (<http://www.deep-er.eu>), by the Onderzoekfonds KU Leuven (Research Fund KU Leuven, GOA scheme and Space Weaves RUN project), by the US Air Force EOARD Project FA2386-14-1-0052 and by the Interuniversity Attraction Poles Programme of the Belgian Science Policy Office (IAP P7/08 CHARM). The author is affiliated with Gruppo Nazionale per la Fisica Matematica of the Italian INdAM.

## Appendix A. Von Neumann Stability Analysis of the New Mover

The new mover of the ECSIM is a modification of the standard  $\theta$ -mover and it has similar stability properties. We proceed here to show the stability

for the paradigmatic case of Langmuir waves used in every textbook analysis of PIC stability [2, 1].

We proceed using the Von Neumann analysis [46], meaning we assume the temporal evolution to be harmonic so that all quantities depend on time as  $e^{i\omega t}$ :

$$\begin{aligned}\mathbf{x}_p^{n+1/2} &= \tilde{\mathbf{x}}_p e^{i\omega(n+1/2)\Delta t} \\ \mathbf{v}_p^{n+1} &= \tilde{\mathbf{v}}_p e^{i\omega(n+1)\Delta t}\end{aligned}\tag{A.1}$$

$$\mathbf{E}_p^{n+\theta} = \tilde{\mathbf{E}}_p \left( (1-\theta)e^{i\omega n\Delta t} + \theta e^{i\omega(n+1)\Delta t} \right)$$

We consider the mover introduced in eq. (1) repeated here for the case of a Langmuir wave (i.e. without the magnetic field):

$$\begin{aligned}\mathbf{x}_p^{n+1/2} &= \mathbf{x}_p^{n-1/2} + \Delta t \mathbf{v}_p^n \\ \mathbf{v}_p^{n+1} &= \mathbf{v}_p^n + \frac{q_p \Delta t}{m_p} \mathbf{E}_p^{n+\theta}\end{aligned}\tag{A.2}$$

Setting  $\theta = 1/2$  and substituting in each term the harmonic dependence, we obtain in the Fourier transformed space:

$$\begin{aligned}\tilde{\mathbf{x}}_p \left( e^{i\omega\Delta t/2} - e^{-i\omega\Delta t/2} \right) &= \tilde{\mathbf{v}}_p \Delta t \\ \tilde{\mathbf{v}}_p \left( e^{i\omega\Delta t} - 1 \right) &= \frac{q_s \Delta t}{m_s} \frac{e^{i\omega\Delta t} + 1}{2} \tilde{\mathbf{E}}_p\end{aligned}\tag{A.3}$$

Recalling now the basic property of cold plasma Langmuir waves that relate the variation of the particle displacement with the electric field in Fourier space as [47]:

$$\frac{q_s \tilde{\mathbf{E}}}{m_s} = -\omega_{pe}^2 \tilde{\mathbf{x}}\tag{A.4}$$



the final dispersion relation can be computed as the determinant of the system matrix:

$$\tilde{\mathbf{x}}_p (e^{i\omega\Delta t/2} - e^{-i\omega\Delta t/2}) = \tilde{\mathbf{v}}_p \Delta t \quad (\text{A.5})$$

$$\tilde{\mathbf{v}}_p (e^{i\omega\Delta t} - 1) = -\omega_{pe}^2 \Delta t \frac{e^{i\omega\Delta t} + 1}{2} \tilde{\mathbf{x}}_p$$

Using the Euler definition of the trigonometric functions, the determinant is:

$$\begin{vmatrix} 2i \sin \frac{\omega\Delta t}{2} & -\Delta t \\ \omega_{pe}^2 \cos \frac{\omega\Delta t}{2} & 2i \sin \frac{\omega\Delta t}{2} \end{vmatrix} = 0 \quad (\text{A.6})$$

Expanding the determinant we obtain the final answer:

$$4 \sin \frac{\omega\Delta t}{2} \tan \frac{\omega\Delta t}{2} = \Delta t^2 \omega_{pe}^2 \quad (\text{A.7})$$

This result differs slightly from that of the  $\theta$ -scheme that is (also using  $\theta = 1/2$ ) [15]:

$$4 \tan^2 \frac{\omega\Delta t}{2} = \Delta t^2 \omega_{pe}^2 \quad (\text{A.8})$$

but remains fundamentally different from the result corresponding to the explicit leapfrog algorithm [2]:

$$4 \sin^2 \frac{\omega\Delta t}{2} = \Delta t^2 \omega_{pe}^2 \quad (\text{A.9})$$

As can be seen, for  $\omega_{pe}\Delta t > 2$  the leap frog algorithm leads to an impossibility in the real axis, as the sin function cannot exceed 1 in amplitude. The resulting solutions for  $\omega$  are complex conjugate, one resulting in unphysical growth and signaling numerical instability. Both the new scheme and the  $\theta$ -scheme, instead, have real solutions for  $\omega$  for all values of  $\Delta t$  and are unconditionally stable.

Figure A.9 compares the numerical frequency produced by the three different discretization schemes above. For large  $\omega_{pe}\Delta t$ , the implicit schemes behave similarly, reaching the asymptotic value of  $\omega\Delta t = \pi$ , the Nyquist frequency. The explicit scheme is bound by the constraint  $\omega_{pe}\Delta t < 2$  to prevent the numerical frequency from acquiring an imaginary part.

The two implicit schemes, the new one presented here and the classic  $\theta$  scheme differ slightly but the new scheme is more accurate at intermediate  $\omega_{pe}\Delta t$ , where it remains closer to the physical frequency for the Langmuir waves. The slight improvement is unlikely to have great practical consequences. Nevertheless, it is better to be better, even if slightly.

## Appendix B. Spatial Discretization and its Properties

The main paper uses a shorthand notation for the grid discretization:  $\nabla_g \times \mathbf{E}$  for example. The results reported in the present paper do not depend critically on the spatial discretization, only the temporal discretization counts. Nevertheless, for the final implementation in the result section we need to use a specific discretization. We use the same type of discretization used in the family of IMM codes: Venus, Celeste, Parsek2D and iPic3D. The grid is formed by cells, the centers of the cells host the magnetic field and the vertices of the cells host the electric field.

In the example section we apply this in 1D. The application is straightforward elementary numerical analysis. Here we very briefly give the details. An interval from  $x \in [0, L]$  with periodic conditions is discretized in  $N$  cells. The generic cell  $i$  is an interval  $[x_i, x_{i+1}]$  with center  $x_{i+1/2}$ . We use a uniform grid, so  $x_i = i\Delta x$  with the integer  $i \in [0, N]$ .

The electric field is then located at  $\mathbf{E}(x_i) = \mathbf{E}_i$  and the magnetic field at  $\mathbf{B}(x_{i+1/2}) = \mathbf{B}_{i+1/2}$ . The shorthand notation  $\nabla_g \times \mathbf{E}$  can then be given a concrete meaning:

$$(\nabla \times \mathbf{E})_{i+1/2} = \frac{1}{\Delta x} \begin{pmatrix} 0 \\ -(E_{z,i+1} - E_{z,i}) \\ (E_{y,i+1} - E_{y,i}) \end{pmatrix} \quad (\text{B.1})$$

Analogously, for the curl of the magnetic field,  $\nabla_g \times \mathbf{B}$ , we have:

$$(\nabla \times \mathbf{B})_i = \frac{1}{\Delta x} \begin{pmatrix} 0 \\ -(B_{z,i+1/2} - B_{z,i-1/2}) \\ (B_{y,i+1/2} - B_{y,i-1/2}) \end{pmatrix} \quad (\text{B.2})$$

These expression can be used to explicitly write out the Maxwell's equations. It is interesting to express the Poynting flux term in the energy equation:

$$\begin{aligned} \sum_g V_g (\mathbf{E}_g \cdot \nabla_g \times \mathbf{B} - \mathbf{B}_g \cdot \nabla_g \times \mathbf{E}) \rightarrow \\ \sum_i (B_{y,i+1/2}(E_{z,i+1} - E_{z,i}) - B_{z,i+1/2}(E_{y,i+1} - E_{y,i}) \\ - E_{y,i}(B_{z,i+1/2} - B_{z,i-1/2}) + E_{z,i}(B_{y,i+1/2} - B_{y,i-1/2})) = 0 \end{aligned} \quad (\text{B.3})$$

All terms  $i \in [2, N-1]$  cancel out, leaving only the boundary values corresponding to the end values  $i = 0$  and  $i = N$ . Via the periodic boundary conditions these terms are the same and cancel out as well. In presence of periodic boundary conditions the Poynting flux terms gives an exactly 0 balance as physically meaningful.

## Appendix C. Rate of Convergence in Space and Time of the EC-SIM

A powerful tool to detect coding errors [48] is to test the actual observed rate of convergence and compare it with the order of accuracy theoretically predicted by the modified equation approach [49]. We apply the approach to each module of the code separately: the field equations and the mover. We do not make a full coupling analysis because it depends also on the number of particles and their shape. We analyse the Maxwell solver with given sources and the mover with a given electric fields (that of linear Langmuir waves), in absence of magnetic fields. Note that the coupling with the magnetic field is instead first order in time.

The spatial discretization used for the Maxwell equations is centered and produces a second order accurate convergence. The temporal discretization of the Maxwell equations and of the equations of motion is also centered (in the case of  $\theta = 1/2$ ) and leads also to a second order accurate scheme.

Except for coding error, the rate of convergence obtained solving the method described above with progressively smaller  $\Delta t$  and  $\Delta x$  should therefore also be second order.

Figure C.10 confirms that the discretization error in the mover decreases with the second power of both  $\Delta t$  and  $\Delta x$ . The results are obtained decreasing  $\Delta t$  at fixed  $\Delta x$  (left panel) and viceversa (right panel). The error is measured in  $L_2$  norm with respect to a converged solution.

The combined confirmation of exact energy conservation and second order accuracy in space and time can be used as powerful debugging tool. We have experienced that small errors even just in the boundary conditions break

energy conservation.

## Appendix D. Summary of the ECSIM Algorithm

We summarize below the steps needed in the construction of the ECSIM algorithm

1. Advance the particle position based on the know velocity, as in the leap-frog scheme of explicit PIC:

$$\mathbf{x}_p^{n+1/2} = \mathbf{x}_p^{n-1/2} + \Delta t \mathbf{v}_p^n \quad (\text{D.1})$$

2. Interpolate the known magnetic field to the particles:

$$\mathbf{B}_p^n(\mathbf{x}_p) = \sum_g \mathbf{B}_g^n W_{pg} \quad (\text{D.2})$$

where  $W_{pg} = W(\mathbf{x}_p^{n+1/2} - \mathbf{x}_g)$ .

3. Compute the rotation operators. This can be an explicit matrix in the code or it can be a method (in the C++ sense, a routine in Fortran) for matrix-free formulations:

$$\alpha_p^n = \frac{1}{1 + (\beta_s B_p^n)^2} (\mathbb{I} - \beta_s \mathbb{I} \times \mathbf{B}_p^n + \beta_s^2 \mathbf{B}_p^n \mathbf{B}_p^n) \quad (\text{D.3})$$

where  $\beta_s = q_p \Delta t / 2m_p$ .

4. Compute the hatted current based on the known velocities rotated by the known magnetic field:

$$\hat{\mathbf{J}}_{sg} = \sum_p q_p \alpha_p^n \mathbf{v}_p^n W_{pg} \quad (\text{D.4})$$

with the total current  $\hat{\mathbf{J}}_g = \sum_s \hat{\mathbf{J}}_s$ .

- Construct the mass matrices,  $3v$  in number, where  $v$  is the dimensionality of the vectors (not to be confused with the dimensionality of space  $d$ ). Again this can be a matrix stored in memory or a method for matrix free formulations. The elements of the matrices are:

$$M_{gg'}^{ij} = \sum_s \left( \frac{\beta_s}{V_g} \sum_p q_p \alpha_p^{ij,n} W_{pg'} W_{pg} \right) \quad (\text{D.5})$$

where  $V_g$  is the volume of cell  $g$ . In matrix notation the  $3v$  mass matrices are written as  $M_{gg'}$ , that is without the indices  $i, j$  for the vector directions.

- Solve the linear set of Maxwell equations where the plasma response is provided by the mass matrices:

$$\begin{cases} \nabla_g \times \mathbf{E}^{n+\theta} + \frac{1}{c} \frac{\mathbf{B}^{n+1} - \mathbf{B}^n}{\Delta t} = 0 \\ \nabla_g \times \mathbf{B}^{n+\theta} - \frac{1}{c} \frac{\mathbf{E}^{n+1} - \mathbf{E}^n}{\Delta t} = \frac{4\pi}{c} \left( \hat{\mathbf{J}}_g + \sum_{g'} M_{s,gg'} \mathbf{E}_{g'}^{n+\theta} \right) \end{cases} \quad (\text{D.6})$$

- Interpolate the new field from the grid to the particles:

$$\mathbf{E}_p^{n+\theta}(\mathbf{x}_p) = \sum_g \mathbf{E}_g^{n+\theta} W_{pg} \quad (\text{D.7})$$

- Advance the particle velocity using the new electric field:

$$\bar{\mathbf{v}}_p = \alpha_p^n \left( \mathbf{v}_p^n + \beta_s \mathbf{E}_p^{n+\theta}(\mathbf{x}_p^{n+1/2}) \right) \quad (\text{D.8})$$

The new velocity is then  $\mathbf{v}_p^{n+1} = 2\bar{\mathbf{v}}_p - \mathbf{v}_p^n$ .

As can be observed the computational cycle retains the same nature as that of explicit PIC or of other semi-implicit PIC. There is no iteration between particles and field.

## References

- [1] C. Birdsall, A. Langdon, Plasma Physics Via Computer Simulation, Taylor & Francis, London, 2004.
- [2] R. Hockney, J. Eastwood, Computer simulation using particles, Taylor & Francis, 1988.
- [3] Y. N. Grigoryev, V. A. Vshivkov, M. P. Fedoruk, Numerical Particle-in-Cell Methods: Theory and Applications, de Gruyter, 2002.
- [4] S. Markidis, G. Lapenta, W. VanderHeyden, Z. Budimlić, Implementation and performance of a particle-in-cell code written in java, Concurrency and Computation: Practice and Experience 17 (7-8) (2005) 821–837.
- [5] V. K. Decyk, T. V. Singh, Adaptable particle-in-cell algorithms for graphical processing units, Computer Physics Communications 182 (3) (2011) 641–648.
- [6] B. Verleye, P. Henri, R. Wuyts, G. Lapenta, K. Meerbergen, Implementation of a 2d electrostatic particle-in-cell algorithm in unified parallel c with dynamic load-balancing, Computers & Fluids 80 (2013) 10–16.
- [7] R. Wuyts, T. Haber, G. Lapenta, Helsim: A particle-in-cell simulator for highly imbalanced particle distributions, Procedia Computer Science 51 (2015) 2923–2927.
- [8] G. Lapenta, S. Markidis, Particle acceleration and energy conservation

- in particle in cell simulations, *Physics of Plasmas* (1994-present) 18 (7) (2011) 072101.
- [9] S. Markidis, G. Lapenta, The energy conserving particle-in-cell method, *Journal of Computational Physics* 230 (18) (2011) 7037–7052.
  - [10] G. Chen, L. Chacón, D. C. Barnes, An energy-and charge-conserving, implicit, electrostatic particle-in-cell algorithm, *Journal of Computational Physics* 230 (18) (2011) 7018–7036.
  - [11] W. T. Taitano, D. A. Knoll, L. Chacón, G. Chen, Development of a consistent and stable fully implicit moment method for vlasov–ampère particle in cell (pic) system, *SIAM Journal on Scientific Computing* 35 (5) (2013) S126–S149.
  - [12] D. W. Hewett, A. B. Langdon, Electromagnetic direct implicit plasma simulation, *Journal of Computational Physics* 72 (1987) 121–155. doi: 10.1016/0021-9991(87)90075-1.
  - [13] A. Langdon, B. Cohen, A. Friedman, Direct implicit large time-step particle simulation of plasmas, *J. Computat. Phys.* 51 (1983) 107–138.
  - [14] D. R. Welch, D. V. Rose, R. E. Clark, T. C. Genoni, T. Hughes, Implementation of an non-iterative implicit electromagnetic field solver for dense plasma simulation, *Computer physics communications* 164 (1) (2004) 183–188.
  - [15] J. Brackbill, D. Forslund, Simulation of low frequency, electromagnetic phenomena in plasmas, *J. Computat. Phys.* 46 (1982) 271.



- [16] J. U. Brackbill, B. I. Cohen, Multiple time scales, Vol. 3, Academic Press, 2014.
- [17] H. X. Vu, J. U. Brackbill, Celest1d: An implicit, fully-kinetic model for low-frequency, electromagnetic plasma simulation, *Comput. Phys. Comm.* 69 (1992) 253.
- [18] G. Lapenta, J. U. Brackbill, P. Ricci, Kinetic approach to microscopic-macroscopic coupling in space and laboratory plasmas, *Physics of Plasmas* 13 (5) (2006) 055904. doi:10.1063/1.2173623.
- [19] S. Markidis, E. Camporeale, D. Burgess, Rizwan-Uddin, G. Lapenta, Parsek2D: An Implicit Parallel Particle-in-Cell Code, in: N. V. Pogorelov, E. Audit, P. Colella, & G. P. Zank (Ed.), *Astronomical Society of the Pacific Conference Series*, Vol. 406 of *Astronomical Society of the Pacific Conference Series*, 2009, pp. 237–+.
- [20] A. Beck, M. E. Innocenti, G. Lapenta, S. Markidis, Multi-level multi-domain algorithm implementation for two-dimensional multiscale particle in cell simulations, *Journal of Computational Physics* 271 (2014) 430–443.
- [21] M. Innocenti, A. Beck, T. Ponweiser, S. Markidis, G. Lapenta, Introduction of temporal sub-stepping in the multi-level multi-domain semi-implicit particle-in-cell code parsek2d-mlmd, *Computer Physics Communications* 189 (2015) 47–59.
- [22] S. Markidis, G. Lapenta, Rizwan-uddin, Multi-scale simulations of

- plasma with iPIC3D, *Mathematics and Computers and Simulation* 80 (2010) 1509–1519.
- [23] J. Brackbill, On energy and momentum conservation in particle-in-cell simulation, arXiv preprint arXiv:1510.08741.
  - [24] B. I. Cohen, A. B. Langdon, D. W. Hewett, R. J. Procassini, Performance and Optimization of Direct Implicit Particle Simulation, *Journal of Computational Physics* 81 (1989) 151–+. doi:10.1016/0021-9991(89)90068-5.
  - [25] C. D. Boor, *A practical guide to splines*, Springer, 1978.
  - [26] K. S. Yee, J. S. Chen, The finite-difference time-domain (fdtd) and the finite-volume time-domain (fvtd) methods in solving maxwell’s equations, *Antennas and Propagation, IEEE Transactions on* 45 (3) (1997) 354–363.
  - [27] D. Sulsky, J. Brackbill, A numerical method for suspension flow, *Journal of Computational Physics* 96 (2) (1991) 339–368.
  - [28] G. Lapenta, Particle simulations of space weather, *Journal of Computational Physics* 231 (3) (2012) 795–821.
  - [29] D. Burgess, D. Sulsky, J. Brackbill, Mass matrix formulation of the flip particle-in-cell method, *Journal of Computational Physics* 103 (1) (1992) 1–15.
  - [30] K. Noguchi, C. Tronci, G. Zuccaro, G. Lapenta, Formulation of the

- relativistic moment implicit particle-in-cell method, *Physics of Plasmas* 14 (4) (2007) 042308–+. doi:10.1063/1.2721083.
- [31] I. B. Peng, S. Markidis, A. Vaivads, J. Vencels, J. Amaya, A. Divin, E. Laure, G. Lapenta, The formation of a magnetosphere with implicit particle-in-cell simulations, *Procedia Computer Science* 51 (2015) 1178–1187.
- [32] D. Knoll, G. Lapenta, J. Brackbill, A multilevel iterative field solver for implicit, kinetic, plasma simulation, *Journal of computational physics* 149 (2) (1999) 377–388.
- [33] P. Kumar, S. Markidis, G. Lapenta, K. Meerbergen, D. Roose, High performance solvers for implicit particle in cell simulation, *Procedia Computer Science* 18 (2013) 2251–2258.
- [34] R. Goldston, P. Rutherford, *Introduction to plasma physics*, Taylor & Francis, 1995.
- [35] E. S. Weibel, Spontaneously growing transverse waves in a plasma due to an anisotropic velocity distribution, *Physical Review Letters* 2 (3) (1959) 83.
- [36] B. D. Fried, Mechanism for instability of transverse plasma waves, *Physics of Fluids* 2 (3) (1959) 337–337. doi:http://dx.doi.org/10.1063/1.1705933.  
URL <http://scitation.aip.org/content/aip/journal/pof1/2/3/10.1063/1.1705933>

- [37] G. Lapenta, S. Markidis, S. Poedts, D. Vucinic, Space weather prediction and exascale computing, *Computing in Science & Engineering* 15 (5) (2013) 68–76.
- [38] J. Birn, E. R. Priest, Reconnection of magnetic fields: magnetohydrodynamics and collisionless theory and observations, Cambridge University Press, 2007.
- [39] G. Lapenta, S. Markidis, M. V. Goldman, D. L. Newman, Secondary reconnection sites in reconnection-generated flux ropes and reconnection fronts, *Nature Physics* 11 (8) (2015) 690–695.
- [40] L. K. Daldorff, G. Tóth, T. I. Gombosi, G. Lapenta, J. Amaya, S. Markidis, J. U. Brackbill, Two-way coupling of a global hall magnetohydrodynamics model with a local implicit particle-in-cell model, *Journal of Computational Physics* 268 (2014) 236–254.
- [41] M. Ashour-Abdalla, G. Lapenta, R. J. Walker, M. El-Alaoui, H. Liang, Multiscale study of electron energization during unsteady reconnection events, *Journal of Geophysical Research: Space Physics* 120 (6) (2015) 4784–4799.
- [42] G. Lapenta, M. Ashour-Abdalla, R. J. Walker, M. El Alaoui, A multi-scale study of ion heating in earth’s magnetotail, *Geophysical Research Letters*.
- [43] G. Chen, L. Chacón, D. C. Barnes, An efficient mixed-precision, hybrid cpu–gpu implementation of a nonlinearly implicit one-dimensional

- particle-in-cell algorithm, *Journal of Computational Physics* 231 (16) (2012) 5374–5388.
- [44] D. A. Mallon, N. Eicker, M. E. Innocenti, G. Lapenta, T. Lippert, E. Suarez, On the scalability of the clusters-booster concept: a critical assessment of the deep architecture, in: *Proceedings of the Future HPC Systems: the Challenges of Power-Constrained Performance*, ACM, 2012, p. 3.
  - [45] A. Vapirev, J. Deca, G. Lapenta, S. Markidis, I. Hur, J.-L. Cambier, Initial results on computational performance of intel many integrated core, sandy bridge, and graphical processing unit architectures: implementation of a 1d c++/openmp electrostatic particle-in-cell code, *Concurrency and Computation: Practice and Experience* 27 (3) (2015) 581–593.
  - [46] E. Isaacson, H. B. Keller, *Analysis of numerical methods*, Courier Corporation, 1994.
  - [47] R. O. Dendy, *Plasma dynamics*, Oxford University Press, Oxford, UK, 1990.
  - [48] P. J. Roache, Code verification by the method of manufactured solutions, *Journal of Fluids Engineering* 124 (1) (2002) 4–10.
  - [49] R. Warming, B. Hyett, The modified equation approach to the stability and accuracy analysis of finite-difference methods, *Journal of computational physics* 14 (2) (1974) 159–179.

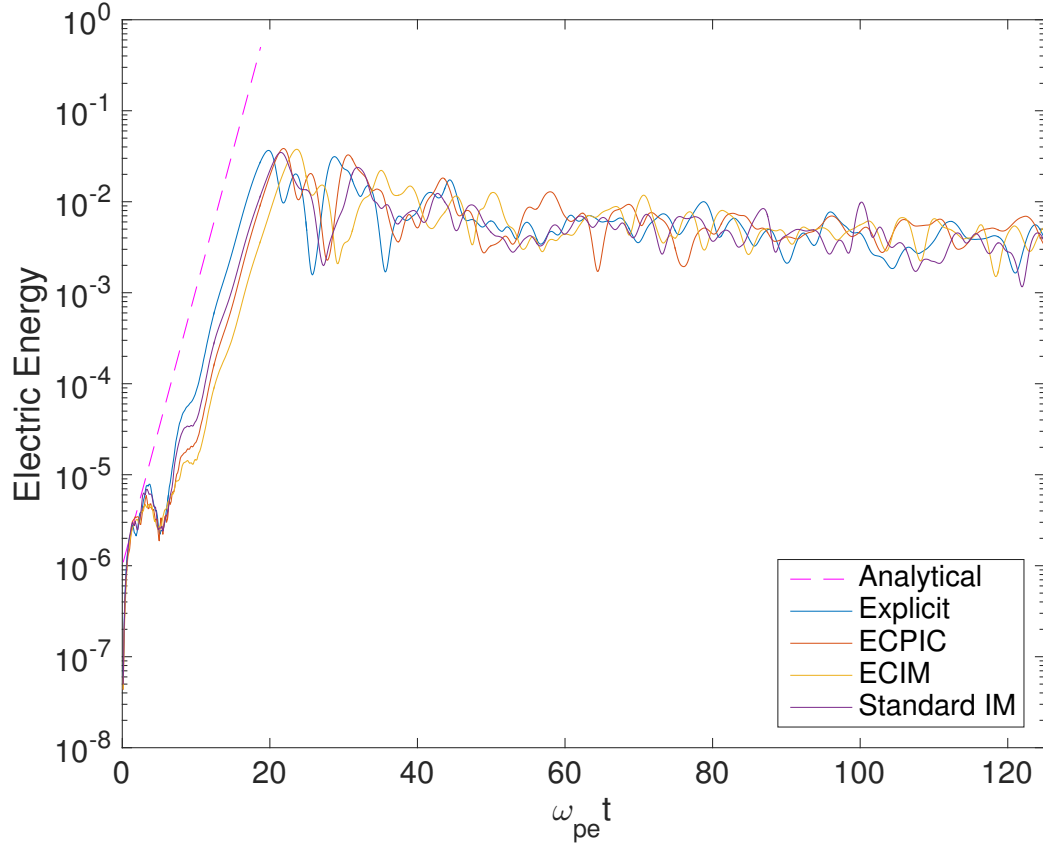


Figure 1: Evolution of the integral of the electric field energy in four runs of the two-stream instability. The linear growth rate predicted by theory is shown by a dashed line. Four runs are compared for the same problem of the two-stream instability: the ECSIM presented here, the standard explicit PIC, the fully implicit ECPIC and the standard IMM.

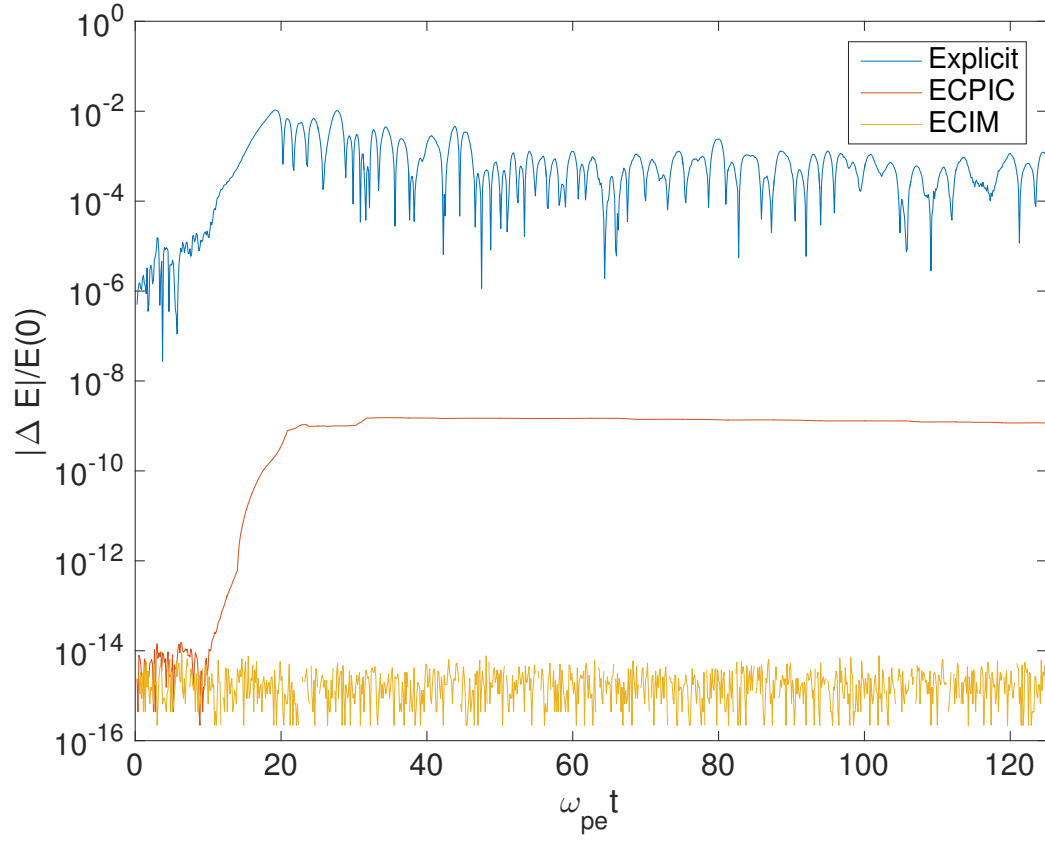


Figure 2: Evolution of the error in the electric field energy in three runs of the two-stream instability: the ECSIM presented here, the standard explicit PIC, the fully implicit ECPIC and the standard IMM.

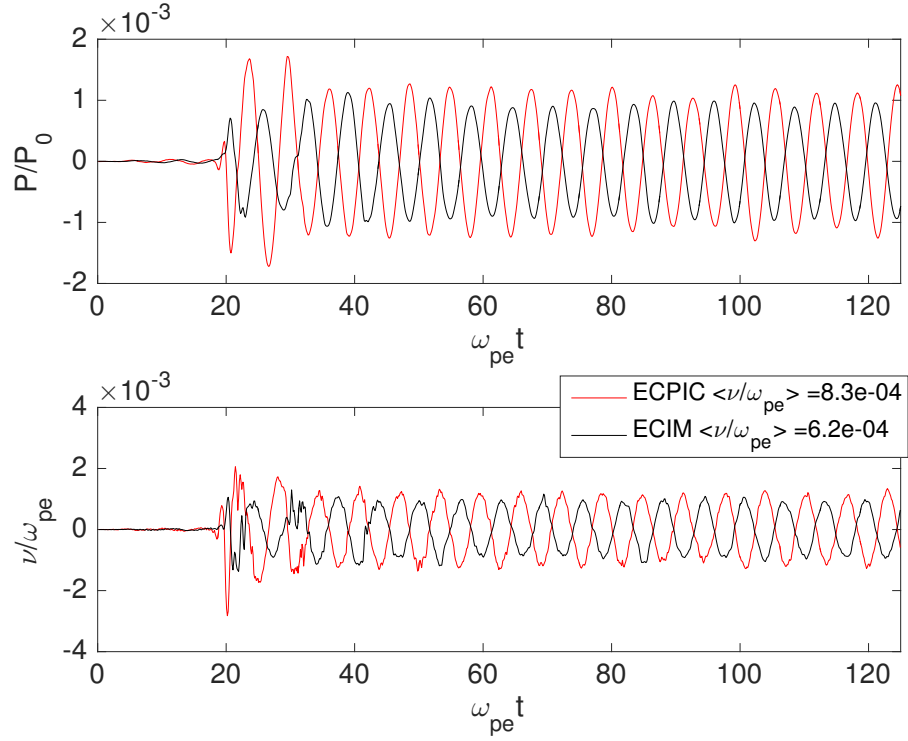


Figure 3: Evolution of the fractional change in momentum. The top panel measures the momentum conservation error as  $P/P_0 = \sum_p m_p u_p / \sum_p m_p V_0$  where  $V_0$  is the initial velocity of the two beams. The bottom panel measures the error in terms of an effective numerical collision probability defined from the empirical formula  $\nu = |dP/dt|/P$ . An average collisionality is derived as the standard deviation of the fluctuating value. The fully implicit ECPIC (in back) and the semi implicit ECSIM (in red) are compared.



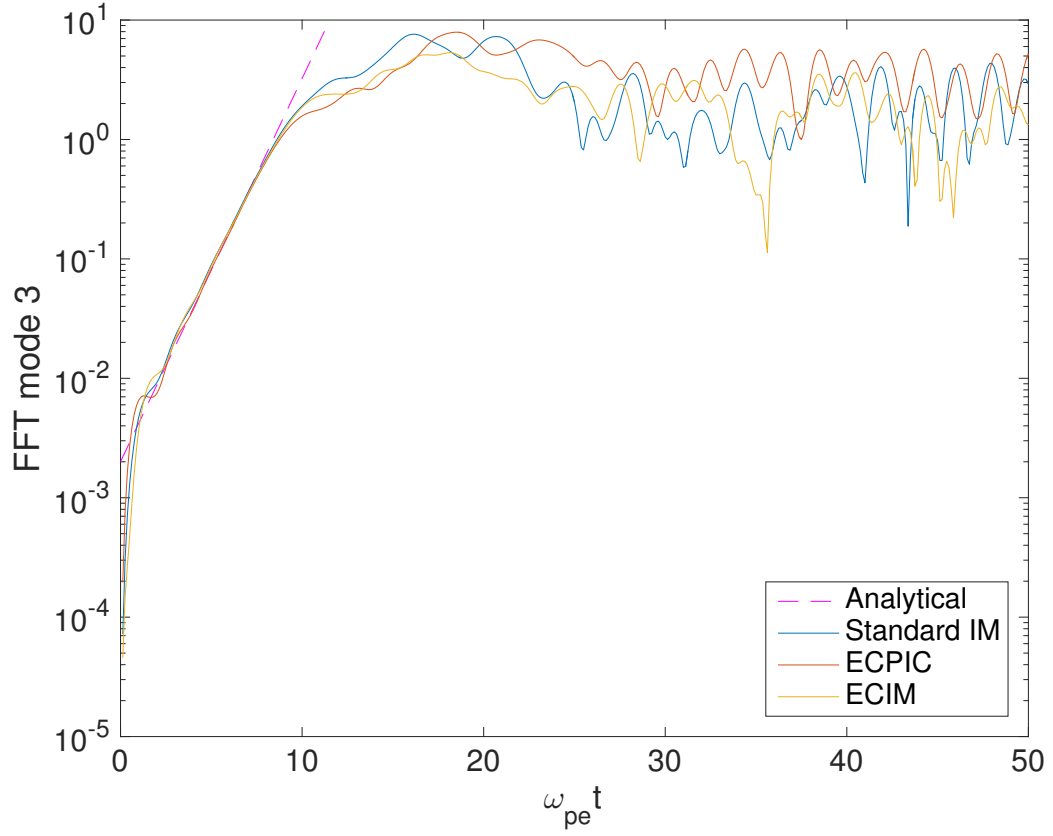


Figure 4: Evolution of the amplitude of Fourier mode  $m = 3$  of the magnetic field component  $B_z$  in different runs of the transverse instability. The linear growth rate predicted by theory is shown by a dashed line. Four runs are compared for the same problem of the two-stream instability: the ECSIM presented here, the fully implicit ECPIC and the standard IMM.

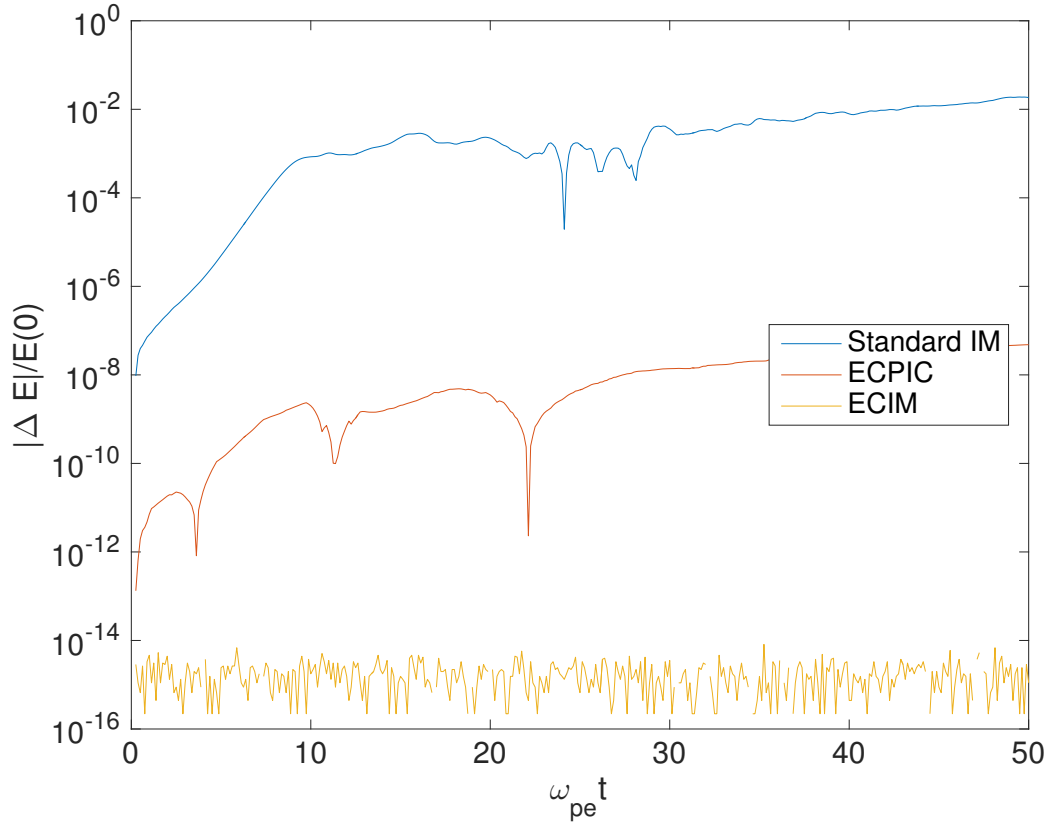


Figure 5: Evolution of the error in the electric field energy in three runs of the transverse instability: the ECSIM presented here, the fully implicit ECPIC and the standard IMM.

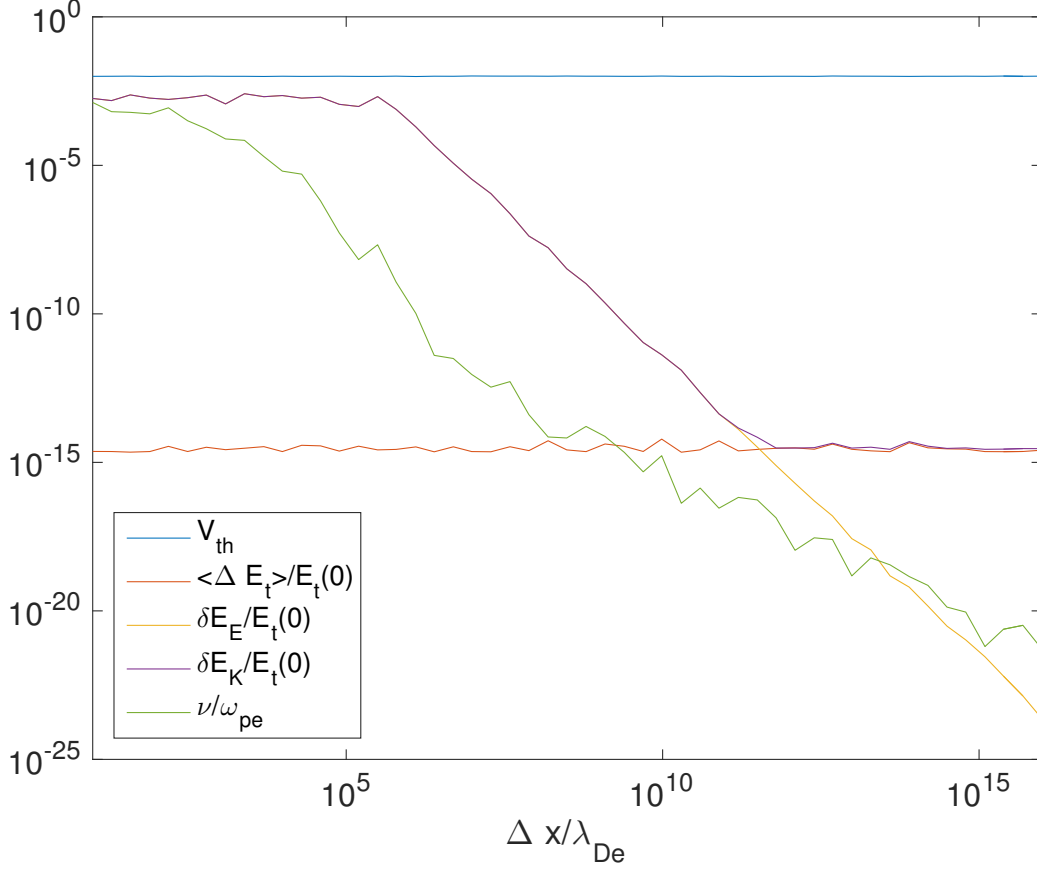


Figure 6: Dependence of several run quality factors on the grid spacing normalized to the Debye length of an initially maxwellian and stable plasma. We report 51 different runs with grid spacing ranging over 16 orders of magnitude from  $\Delta x/\lambda_{De} = 10$  to  $\Delta x/\lambda_{De} = 10^{16}$ . In all runs,  $CFL = 0.0127$ . The parameters reported are: thermal speed at the end of each run, the average total energy error during each run, normalized to the initial total energy, the standard deviation of the fluctuation of the electric field,  $\delta E_E/E_{tot}(0)$  (normalized to the initial total energy), the standard deviation of the fluctuation of the particle kinetic energy,  $\delta E_K/E_{tot}(0)$  (normalized to the initial total energy) and the effective numerical collision frequency  $\nu$  implied by the lack of exact momentum conservation.

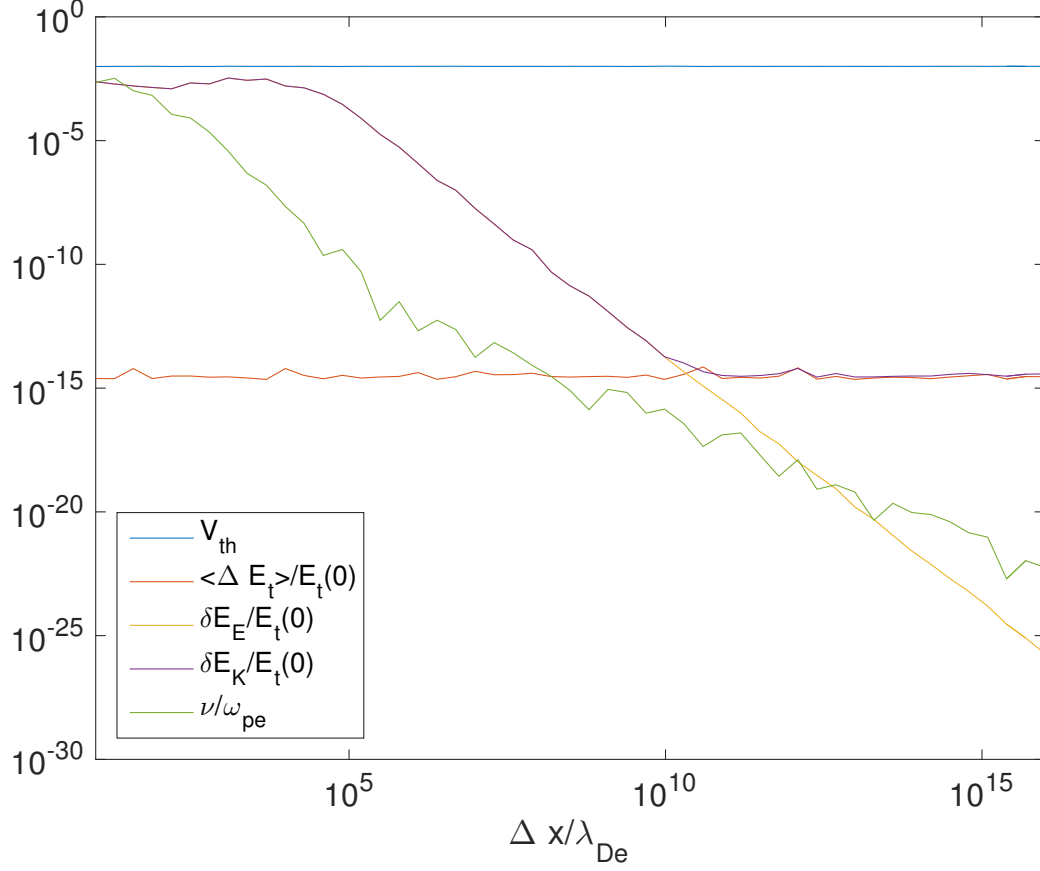


Figure 7: Dependence of several run quality factors on the grid spacing normalized to the Debye length of an initially maxwellian and stable plasma. We report 51 different runs with grid spacing ranging over 16 orders of magnitude from  $\Delta x/\lambda_{De} = 10$  to  $\Delta x/\lambda_{De} = 10^{16}$ . In all runs,  $CFL = 0.127$ . The parameters reported are: thermal speed at the end of each run, the average total energy error during each run, normalized to the initial total energy, the standard deviation of the fluctuation of the electric field,  $\delta E_E/E_{tot}(0)$  (normalized to the initial total energy), the standard deviation of the fluctuation of the particle kinetic energy,  $\delta E_K/E_{tot}(0)$  (normalized to the initial total energy) and the effective numerical collision frequency  $\nu$  implied by the lack of exact momentum conservation.

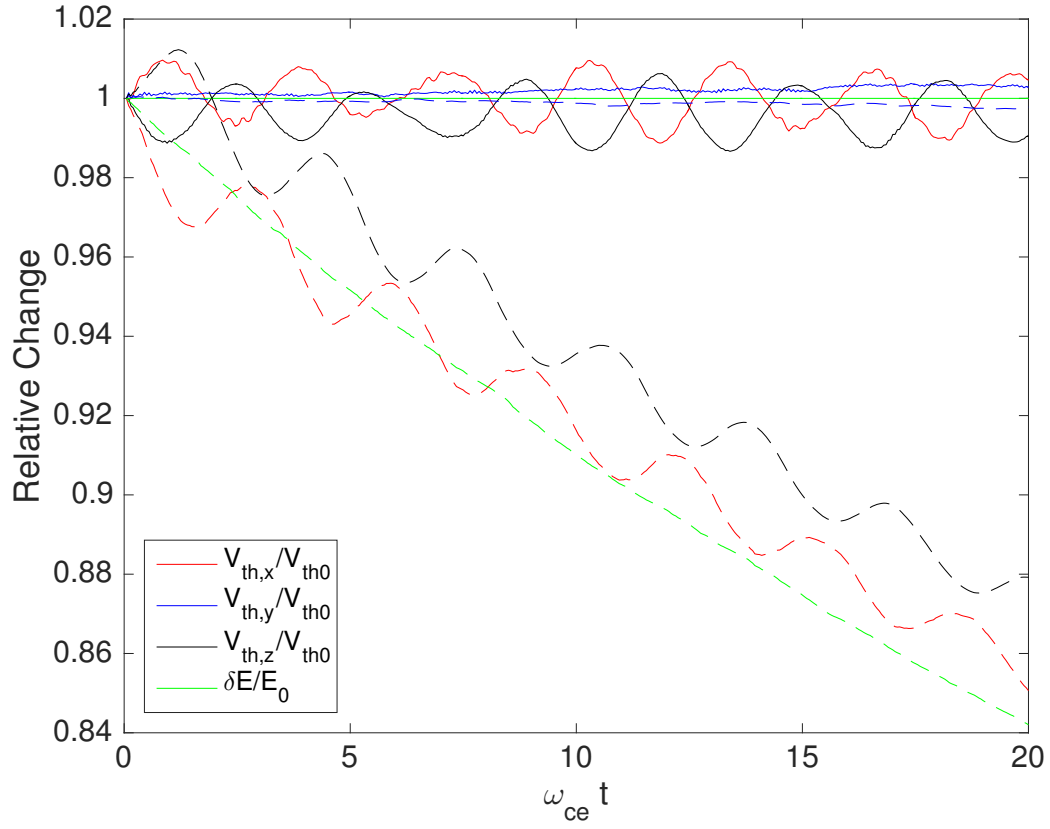


Figure 8: Paradigmatic space plasma simulation. Evolution of the total energy and of the electron temperature in each direction for a standard IMM (dashed lines) and for the new ECSIM method (solid lines).

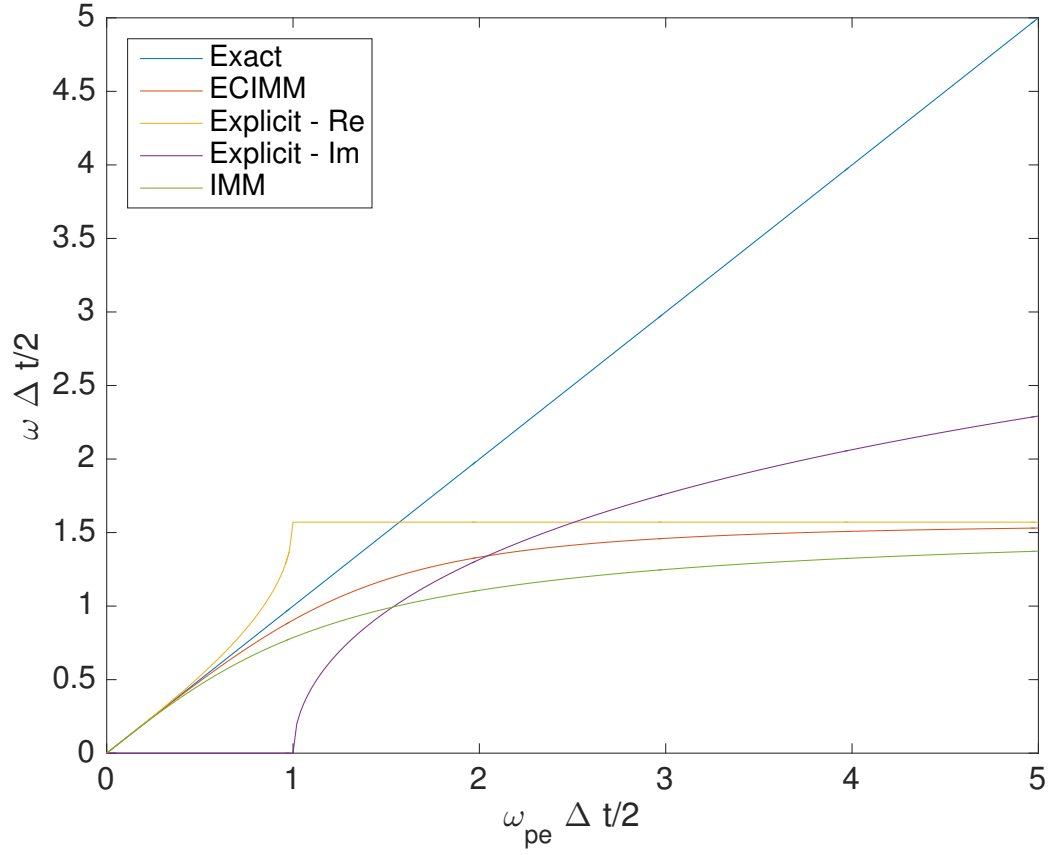


Figure A.9: Stability analysis for 3 movers: explicit, standard IMM and new ECSIM. The numerical frequency  $\omega$  for Langmuir waves is obtained by the Von Neumann analysis as a function of the time step.

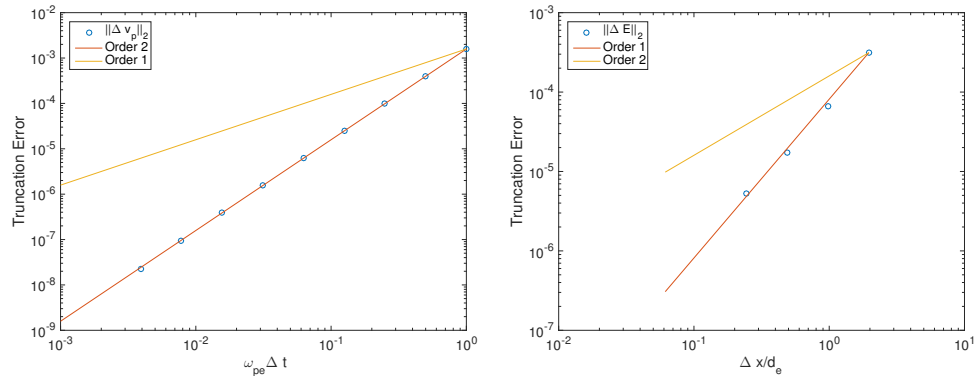


Figure C.10: Stability analysis for 3 movers: explicit, standard IMM and new ECSIM. The numerical frequency  $\omega$  for Langmuir waves is obtained by the Von Neumann analysis as a function of the time step.