The relativistic implicit Particle-in-Cell method

Fabio Bacchini, Jorge Amaya, and Giovanni Lapenta
Centre for mathematical Plasma Astrophysics, Department of Mathematics, KU Leuven, Celestijnenlaan 200B, B-3001 Leuven, Belgium
E-mail: fabio.bacchini@kuleuven.be

Abstract. Relativistic Particle-in-Cell (PiC) methods are among the most reliable methods for the investigation of plasma phenomena at the particle scale. Standard explicit and semi-implicit PiC methods are affected by numerical instabilities that restrain the range of admissible simulation parameters, and prevent their application to large domains over long time scales. Here, we present a three-dimensional, fully-implicit algorithm for relativistic PiC simulations that conserves energy exactly (to machine precision) and eliminates numerical instabilities, allowing long-term calculations. We describe the full numerical solution procedure and test the algorithm with several one- and two-dimensional case-studies of kinetic instabilities, comparing the results with theoretical predictions.

1. Introduction
High-energy astrophysical environments where plasmas reach relativistic speeds are a natural laboratory where the interplay of strong electromagnetic fields and the dynamics of particles can be studied. Such scenarios are considered realistic for plasma flows in the magnetosphere of compact objects (black holes and neutron stars, see e.g. [1]), pulsar wind nebulae (PWNe, see [2]), shock fronts of relativistic jets [3], and many others. The modeling of relativistic plasmas can be efficiently carried out with numerical codes, simulating the relevant dynamics on a global scale with magnetohydrodynamic (MHD) approaches, or at the particle scale with a fully-kinetic description.

Particle-in-Cell (PiC) methods are the most successful tool for the simulation of the kinetic dynamics of plasmas. PiC algorithms in a broad sense represent a class of numerical methods for the solution of partial differential equations in contexts where advection of physical quantities plays an important role in the relevant phenomena. As such, PiC methods can be applied to the modeling of fluids [4], to solid mechanics problems [5] and, most importantly, to the kinetic description of plasma processes. The latter context is by far the most widely renowned field of application of PiC codes, hence the common association of the term “PiC” with kinetic physics. PiC codes exist in a broad variety of formulations both for simulations in the nonrelativistic limit [6][7][8], and for relativistic applications [9][10][11].

Despite their success, commonly employed relativistic PiC codes are based on the explicit solution of particle and field equations, which is well-known for exhibiting spurious heating of the lighter particles due to numerical instabilities [12]. In the most severe cases, such instabilities can completely invalidate the numerical results, which become unphysical and often cause the failure of the simulation run. Early semi-implicit schemes, based on the implicit solution of the field equations (with the particles evolved explicitly), relax the constraints imposed by numerical
instabilities, but are still prone to destabilization when specific criteria for the chosen spatial and temporal resolutions are not met. In practice, one is often bound to choose a grid spacing that resolves the electron Debye length \( \lambda_{De} = \sqrt{\varepsilon_0 kT_e/e^2n_e} \), calculated from the electron temperature \( kT_e \), electron charge \( e \), and number density \( n_e \). For electron-positron or electron-ion plasmas with low electron temperatures, however, the dynamics of specific phenomena such as magnetic reconnection can take place over the much larger length and time scales of the (electron or ion) skin depth \[8\]. The requirement of resolving the unimportant Debye length scales imposes a detrimental increase in the computational effort required by explicit codes.

Energy-conserving PiC algorithms are not subjected to this issue, due to the elimination of instabilities related to the spurious non-conservation of energy. The exact conservation of global energy is attained via specific choices of discretization of the grid and particle equations. Energy-conserving semi- and fully-implicit methods have been proposed and successfully applied for plasma simulations in the nonrelativistic limit \[13\] \[6\] \[8\]. In the relativistic framework, only fully-implicit algorithms allow for attaining exact energy conservation, due to the implicit nature of the momentum equation for the particles \[14\].

In this work, we present a multi-dimensional, fully implicit algorithm for relativistic plasma simulations that conserves energy exactly (to machine accuracy), exhibiting superior stability properties. The new scheme is based on that presented by \[14\], but differs in the solution procedure that is adapted to handle coarse temporal and spatial resolutions, as well as realistic ion-to-electron mass ratios, and achieves affordable computational costs in multi-dimensional runs. In the next Section, we present in full the numerical scheme and two alternative solution strategies for the discretized Vlasov-Maxwell system of equations. Then, we test our algorithm against analytic calculations by simulating several types of kinetic instabilities and comparing to the predictions from linear theory. Finally, we provide an example application to a magnetic reconnection case-study, involving large separation of scales between the ion and electron dynamics.

2. Relativistic implicit Particle-in-Cell

2.1. The Vlasov-Maxwell system and the finite particle approach

The physics of plasmas at the microscopic scale is described by the system composed of Maxwell’s equations for the electromagnetic fields \( E, B \),

\[
\frac{\partial E}{\partial t} = \nabla \times B - \mu_0 J, \tag{1}
\]

\[
\frac{\partial B}{\partial t} = -\nabla \times E, \tag{2}
\]

\[
\nabla \cdot E = \frac{\rho}{\varepsilon_0}, \tag{3}
\]

\[
\nabla \cdot B = 0, \tag{4}
\]

and the Vlasov equation for the distribution function \( f_s(x, u, t) \) of each particle species \( s \),

\[
\frac{\partial f_s}{\partial t} + \frac{dx}{dt} \cdot \frac{\partial f_s}{\partial x} + \frac{du}{dt} \cdot \frac{\partial f_s}{\partial u} = 0. \tag{5}
\]

Here, all quantities are expressed in SI units, with the speed of light as a normalization constant.

A most successful numerical approach for solving the Vlasov-Maxwell system of equations above is the so-called kinetic Particle-in-Cell (PiC) method \[12\]. This class of numerical
methods employs a finite number $N_s$ of “superparticles” for each species $s$, with the fundamental assumption that the function $f_s$ can be sampled by the superposition

$$f_s(x, u, t) = \sum_{p}^{N_s} S(x - x_p) \delta(u - u_p),$$

where $S$ is a shape function characterizing each superparticle $p$. Because of the linearity of the superposition, the evolution of $f_s$ from equation (6) is still governed by the Vlasov equation (5). PiC methods evolve the position and momentum, $x_p$ and $u_p$ of each superparticle in time. The particle position and momentum obey the equations of motion,

$$\frac{dx_p}{dt} = \frac{u_p}{\gamma_p},$$

$$\frac{du_p}{dt} = \frac{q_p}{m_p} \left( E(x_p) + \frac{u_p}{\gamma_p} \times B(x_p) \right),$$

where $\gamma_p = \sqrt{1 + u^2_p}$ is the Lorentz factor, and the right-hand side of equation (8) is the Lorentz force for a particle of mass $m_p$ and charge $q_p$. All the computational (super)particles are evolved by solving equations (7) and (8).

The evolution of the electromagnetic fields is calculated by solving finite-difference approximations of Maxwell’s equations (1)-(4) on a computational grid of $N_g$ elements. Each computational cell has volume $V_g$. Because the shape function $S$ defines the finite size of each particle, the average electric and magnetic field acting on a single particle are defined as

$$E(x_p) = \int_{V_g} E(x) S(x - x_p) dx = \sum_g E_g W(x_g - x_p),$$

$$B(x_p) = \int_{V_g} B(x) S(x - x_p) dx = \sum_g B_g W(x_g - x_p),$$

where the electric and magnetic fields on each grid element $E_g$ and $B_g$ contribute to the particle according to the value assumed by the interpolation function $W(x_g - x_p)$. The interpolation function is related to the shape function by

$$W(x_g - x_p) = \int_{V_g} S(x - x_p) dx.$$

In one dimension, a typical choice for the shape function $S$ is the zeroth-order b-spline, and as a consequence the interpolation function $W$ simply results in the first-order b-spline. In multi-dimensional simulations, the shape function and interpolation function are simply taken as the convolution of b-splines along each spatial dimension. Finally, the interpolation function defines the contribution of the particles to the source terms of Maxwell’s equations at each grid point,

$$\rho_g = \frac{1}{V_g} \sum_s \sum_p^{N_s} q_p W(x_g - x_p)$$

$$J_g = \frac{1}{V_g} \sum_s \sum_p^{N_s} \frac{u_p}{\gamma_p} W(x_g - x_p).$$
2.2. Implicit time discretization

Implicit PiC schemes employ an implicit time discretization of Maxwell’s equations and of the particle equations of motion. Our choice of discretization, in the most general case, consists of a $\theta$-scheme for the fields,

$$\frac{E_{g}^{n+1} - E_{g}^{n}}{\Delta t} = \nabla \times B_{g}^{n+\theta} - \mu_{0}\tilde{J}_{g}, \quad (14)$$

$$\frac{B_{g}^{n+1} - B_{g}^{n}}{\Delta t} = -\nabla \times E_{g}^{n+\theta}, \quad (15)$$

which eliminates the restrictions on the time step $\Delta t$ imposed by the CFL condition when $\theta > 1/2$ [7]. The particle equations of motion are discretized with a midpoint scheme [14],

$$\frac{x_{p}^{n+1} - x_{p}^{n}}{\Delta t} = \frac{\bar{u}_{p}}{\gamma_{p}}, \quad (16)$$

$$\frac{u_{p}^{n+1} - u_{p}^{n}}{\Delta t} = \frac{q_{p}}{m_{p}} \left( E_{g}^{n+\theta}(\bar{x}_{p}) + \frac{\bar{u}}{\gamma} \times B_{g}^{n+\theta}(\bar{x}_{p}) \right), \quad (17)$$

Here, quantities at time level $n + \theta$ are taken as a linear combination of values between time levels $n$ and $n + 1$, e.g. $E^{n+\theta} = \theta E^{n+1} + (1 - \theta) E^{n}$, with the decentering parameter $\theta = [1/2, 1]$. For the case $\theta = 1/2$, the scheme above is of second-order accurate in time. Barred quantities thus represent values at $n + 1/2$, e.g. $\bar{u}_{p} = (u_{p}^{n+1} + u_{p}^{n})/2$. A special case is represented by the average current $\tilde{J}$, which we define as

$$\tilde{J}_{g} = \sum_{s} \sum_{p} q_{p} \bar{u}_{p} W(\bar{x}_{p} - x_{g}). \quad (18)$$

Note that in the system above, we do not consider the two divergence equations [3] and [4] for simplicity. It can be shown that the divergence-free condition for the magnetic field is automatically satisfied by our choice of formulation. Gauss’s law for the electric field, instead, can be enforced by either applying a divergence cleaning step [7] or by subcycling during the current deposition step (as demonstrated in [6] in the Newtonian limit). Note that the former technique, however standard, violates the energy conservation properties of the scheme [13]; the latter, instead, ensures energy conservation at the cost of an increased computational cost. Alternatively, one can rely on charge-conserving deposition schemes [15], which however mutually exclude the energy-conserving formulation of the current (18). Very recently a second alternative approach was proposed for nonrelativistic, semi-implicit energy-conserving PiC methods that enforces Gauss’s law by modifying the particle position at each time step [16]. Whether this is applicable in our algorithm is an interesting possibility that will be considered in the future. Despite the absence of a method to enforce charge conservation, in the results of the tests presented here we did not detect severe unphysical features or spurious oscillations in the field quantities.

For the remainder of this work, we focus on the case $\theta = 1/2$. With our choice, the system of equations to be solved is

$$\frac{E_{g}^{n+1} - E_{g}^{n}}{\Delta t} = \nabla \times B_{g}^{n+1/2} - \mu_{0}\tilde{J}_{g}, \quad (19)$$

$$\frac{B_{g}^{n+1} - B_{g}^{n}}{\Delta t} = -\nabla \times E_{g}^{n+1/2}, \quad (20)$$

$$\frac{x_{p}^{n+1} - x_{p}^{n}}{\Delta t} = \frac{\bar{u}_{p}}{\gamma_{p}}, \quad (21)$$
\[
\frac{\mathbf{u}_p^{n+1} - \mathbf{u}_p^n}{\Delta t} = q_p \frac{m_p}{\mathbf{u}_p^\gamma} \left( \mathbf{E}(\mathbf{x}_p) + \mathbf{u} \times \mathbf{B}(\mathbf{x}_p) \right).
\]  

This choice makes the overall scheme globally energy-conserving, as shown in the original work by Lapenta & Markidis [14].

2.3. Solution procedure

The numerical scheme proposed above is inherently implicit in both the field and particle variables. Compared to the relatively simple time-marching algorithm of explicit PiC schemes, implicit methods involve a higher degree of difficulty in the solution procedure. A comparison between the two solution algorithms is sketched in Figure 1. Explicit PiC codes (left) typically employ a four-step sequential scheme to interpolate the fields at the particle position, advance the particle quantities, gather the source term for the fields from the particles, and advance the field quantities on the grid. Implicit PiC schemes (right), instead, must solve particle and field equations simultaneously, clearly involving a higher computational cost. Additionally, for the relativistic case considered here, the particle equations are also nonlinear, due to the presence of the Lorentz factor in equation (22). Because the field equations are coupled implicitly to the particle equations, the whole system of particle and field equations inherits the nonlinear character, and thus is typically solved iteratively.

Figure 1: Schematic comparison between the typical computational cycle of an explicit PiC method (left) and an implicit method (right). Black boxes indicate explicit (non-iterative) operations, while red boxes indicate the use of an iterative solution process.

The most direct method for solving the implicit equations consists of handling the full system (19)–(22) with any preferred nonlinear solver. This solution was shown to be applicable for one-dimensional systems in the original work [14]. This strategy, however, is not suitable for realistic three-dimensional simulations. In this case, the full system consists of \(6N_g + 6N_p\) coupled nonlinear equations. For typical PiC runs where the number of particles \(N_p\) is of order millions or billions, the computational cost of such a direct approach is simply prohibitive.

Here, we consider instead two alternative strategies that reduce the complexity of the system to \(3N_g\) coupled linear or nonlinear equations for the fields, and \(3N_p\) independent nonlinear equations for the particles. The implicit solution of such a system involves a computational cost
that surely exceeds that of an explicit scheme, but that can still be handled by state-of-the-art linear/nonlinear solvers, contrary to the original $6N_g + 6N_p$ nonlinear system. The workflow of the two algorithms is compared in Figure 2. In the next Sections, we describe the two strategies in detail and discuss their features.

2.3.1. System reduction

As a first step, we can reduce the total size of the system of equations from $6N_g + 6N_p$ by a factor 2. The field equations can be reduced with a standard procedure employed in implicit PiC codes [13] as follows. First, substitute the electric and magnetic fields at time level $n + 1$ with the definitions in terms of the corresponding barred quantities into equations (19) and (20),

\[
\frac{\mathbf{E}_g^n - \mathbf{E}_g}{\Delta t/2} = \nabla \times \mathbf{B}_g - \mu_0 \tilde{\mathbf{J}}_g,
\]

(23)

\[
\frac{\mathbf{B}_g^n - \mathbf{B}_g}{\Delta t/2} = -\nabla \times \mathbf{E}_g.
\]

(24)

Then, take the curl of equation (24) and substitute it in equation (23) to obtain a second-order equation in $\mathbf{E}$ only,

\[
\frac{\mathbf{E}_g^n - \mathbf{E}_g}{\Delta t/2} = \nabla \times \mathbf{B}_g^n - \frac{\Delta t}{2} \nabla \times \nabla \times \mathbf{E}_g - \mu_0 \tilde{\mathbf{J}}_g.
\]

(25)

With such a manipulation, the field part of the system is reduced to $3N_g$ equations. The magnetic field can then be computed with equation (24).

The particle equations can be reduced in two ways. A first, straightforward approach is to employ a leap-frog scheme for the position and momentum, such that the position equation (21) is written

\[
\frac{x_p^{n+1/2} - x_p^{n-1/2}}{\Delta t} = \frac{u_p^n}{\gamma_p^n}.
\]

(26)

In this way, the average position $\bar{x}_p \equiv x_p^{n+1/2}$ does not depend on the unknown $\bar{u}_p$. A second, more consistent approach is to substitute for $\bar{x}_p$ into equation (21), such that

\[
\frac{\bar{x}_p - x_p^n}{\Delta t/2} = \bar{u}_p.
\]

(27)

Then, by substituting this expression into equation (22), we are left with a system of equations in the unknown $\bar{u}_p$ only ($u_p^{n+1}$ can also be substituted by inverting $\bar{u}_p = (u_p^{n+1} + u_p^n)/2$). Either approach reduces the system size to $3N_p$ nonlinear equations.

2.3.2. Field hiding

The field hiding (FH) strategy (Figure 2, left) proposed recently for nonrelativistic implicit PiC [17] decouples the field and particle parts of the initial set of equations (19)-(22). The particle equations of motion (21)-(22) are solved independently for each particle with an iterative process, until convergence below a prescribed tolerance. With an initial guess for $\bar{u}$, the current, $\tilde{\mathbf{J}}$, is computed on the grid. Then, equations (25) and (24) are solved for the electromagnetic fields, $\mathbf{E}$ and $\mathbf{B}$. With the new values of the fields, the momentum $\bar{u}$ is computed from equation (22), and the iteration is restarted.

The fundamental assumption of the FH strategy is that the motion of each particle can be resolved while neglecting the motion of all other particles. This dependence is in fact “hidden” in the field terms. In practice, when solving equation (22) for $\bar{u}$, one neglects the dependence $\mathbf{E}_g(\bar{u})$ and $\mathbf{B}_g(\bar{u})$ (arising from the expression of $\tilde{\mathbf{J}}_g$), and the field terms are essentially considered fixed for the current iteration. This effectively reduces the global problem to the solution of $6N_p$ ($3N_p$,

after the reduction outlined above) *independent* equations. Should this dependence be taken into account, the system of equations to be solved would consist instead of $6N_p (3N_p)$ *coupled* equations, which is clearly prohibitive, given the large number of particles employed in typical three-dimensional PiC simulations.

The strategy above implies, as a secondary assumption, that when solving equation (25) for $\mathbf{E}$ the source term, $\tilde{\mathbf{J}}$, is kept fixed. This simplification makes equation (25) linear, such that it can be solved with any preferred solver for linear systems.

In summary, in the FH approach it is sufficient to solve a system of $3N_g$ coupled linear equations for the fields, and $3N_p$ independent nonlinear equations for the particles, at each iteration step. The reduction in computational cost, compared to the initial system size, is of orders of magnitude. The resulting system of equations can be easily handled by modern linear solvers. However, neglecting the nonlinear dependence that characterizes the global system of equations can have a detrimental effect on the results. The FH algorithm is indeed not suitable for all applications; for specific cases, e.g. electron-proton plasmas with realistic mass ratios, it requires a critical reduction of the time step, thus neglecting the advantage of the implicit solution procedure. On the other hand, if the FH scheme can be applied, it allows for neglecting the nonlinear coupling between all the particles, which introduces a much higher complexity in the overall algorithm (millions or billions of coupled nonlinear equations).

### 2.3.3. Kinetic enslavement

An alternative approach to the FH strategy consists of retaining the nonlinearity of the initial system, but confining it to the field equations instead of the particle equations. With the kinetic enslavement (KE) approach (Figure 2, right), an iterative solution is carried out on the field quantities only. At each time step, an initial guess is given for $\mathbf{E}_g$ and $\mathbf{B}_g$. Then, the particles are evolved independently of each other, by solving equation (22), and the current $\tilde{\mathbf{J}}_g$ is gathered on the grid. With the new values of the current, the fields are recalculated by solving equation (25). The iteration is carried out until convergence of the fields below a prescribed tolerance. Such a technique was first developed for nonrelativistic implicit PiC methods [18], and successfully applied in the Newtonian regime [6][13][19]. To the best of our knowledge, here we show for the first time the application of the KE strategy to relativistic implicit PiC simulations, which also implies a substantially different approach for the solution of the momentum equation (see next Section).

Here, the key concept is that the nonlinear character of equation (25) is retained by considering the nonlinear dependence $\tilde{\mathbf{J}}_g(\mathbf{E}_g)$ (arising via the particle equations). As a consequence, one needs to solve a system of $3N_g$ coupled *nonlinear* equations, contrary to the FH strategy where the system is linear. Therefore, a nonlinear solver must be employed for the field advancing step. This approach shows much higher robustness for typical implicit PiC applications, at a cost that does not outmatch dramatically the FH strategy. In summary, the overall complexity of this strategy involves solving $3N_g$ (after reduction) coupled nonlinear equations for the fields, and $3N_p$ independent nonlinear equations for the particles.

### 2.3.4. Solution of the momentum equation

The discrete equation of motion (22) differs from typical choices of discretization employed in relativistic particle simulations. Typical schemes are formulated such that the discretized momentum equation, though implicit, can be formally inverted in order to carry out the momentum update explicitly (see e.g. [20]). The main difference between these schemes lies in the definition of the “average velocity” used in the cross product term of the Lorentz force. In our formulation of equation (22), the average velocity is

$$\frac{\mathbf{u}_p}{\gamma_p} = \frac{\mathbf{u}_p^{n+1} + \mathbf{u}_p^n}{\gamma_p^{n+1} + \gamma_p^n}$$

(28)
Figure 2: Schematic comparison between the FH and KE strategy for the implicit solution of the nonlinear discretized equations. The FH approach (left) requires a nonlinear iteration (red box) in the particle momentum, while the fields are solved for in a linear system (black box). The KE approach (right) requires instead a nonlinear iteration in the field quantities, while the momentum update is carried out particle by particle outside of the iteration.

where the barred quantities are defined above as the mean value between time steps \( n \) and \( n+1 \). With this choice, the momentum equation becomes more complicated to handle, and a formal inversion is not possible. Here, we consider an iterative solution for the numerical update of \( \mathbf{u}_p \). One can employ a simple root-finding algorithm (e.g. Picard’s or Newton’s method) for the nonlinear solution of equation (22) in the variable \( \mathbf{u}_p^{n+1} \). This approach has the advantage of being applicable also in case the system size is not reduced with the leap-frog approach. In typical simulations, the number of iterations required remains relatively limited (around 4-5 in the examples of the next Sections). The solution of equation (22) is carried out by iteratively searching for the roots of the residual equation

\[
F(\mathbf{u}_p) = \mathbf{u}_p - \mathbf{u}_p^n - \frac{q_p e \Delta t}{m_p} \left( \mathbf{E}(\mathbf{x}_p) + \frac{\mathbf{u}_p + \mathbf{u}_p^n}{\gamma_p + \gamma_p^n} \times \mathbf{B}(\mathbf{x}_p) \right),
\]

(29)

where \( \gamma_p = \sqrt{1 + u_p^2} \) is a function of the iteration variable \( \mathbf{u}_p \). If the leap-frog approach is not used for the position equation, then also \( \mathbf{x}_p = \mathbf{x}_p + (\Delta t/2)(\mathbf{u}_p + \mathbf{u}_p^n)/(\gamma_p + \gamma_p^n) \) is a function of \( \mathbf{u}_p \). This procedure is employed in all numerical experiments in the next Sections.

3. Validation tests

3.1. Counter-streaming instabilities

As is customary for kinetic PIC code, a set of simple one-dimensional tests is run in order to check the accuracy of the new algorithm. Here we choose to quantitatively assess the performance of the relativistic implicit solver both in the electrostatic and in the full electromagnetic limits, by simulating counter-streaming instabilities in pair plasmas.

For the electrostatic limit, we consider the two-stream instability, a classical benchmark for PIC codes in both Newtonian and relativistic regimes. The simplest such a setup consists of two counter-streaming electron-positron beams of opposite initial speed \( \pm v_0 \), flowing parallel to the wave-vector \( k = k \hat{x} \). The system is unstable to specific perturbation modes, and the beams tend
to interact, converting beam bulk energy into thermal energy. The instability is characterized by a maximum growth rate

$$\Gamma_{\text{max}} = \frac{\omega_p^0}{2\gamma_0^{3/2}}, \tag{30}$$

for symmetric counter-streaming beams with initial velocity $v_0$ and Lorentz factor $\gamma_0 = (1 - v_0^2)^{-1/2}$. The associated most unstable wave-vector is $k_{\text{max}} = \sqrt{3} \omega_p^0/(2v_0 \gamma_0^{3/2})$, where $\omega_p^0 = \sqrt{n_0 e^2/(\epsilon_0 m_e)}$.

The two-stream instability is simulated by initializing two counter-streaming electron-positron beams at initial drift speed $v_0$ in a one-dimensional domain with 1000 particles per cell. A small thermal velocity $v_{th}$ is added to excite all the instability modes. In our set of runs, we consider an increasing initial $\gamma_0 = 2, 10, 50$. The domain is enlarged accordingly as $L = 128d_0^e, 1280d_0^e, 12800d_0^e$ (where $d_0^e = c/\omega_p^0$) in order to allow the maximum growth rate to dominate. For each $\gamma_0$, we vary the spatial resolution from a minimum of 32 to a maximum of 256 total cells. In all cases the Debye length is not resolved, exceeding the stability constraint that characterizes explicit schemes. While the highest resolutions suffice, in some cases, to resolve the nonrelativistic skin depth $d_0^e$, the relativistic skin depth $d_e = c/\omega_p = c\sqrt{\gamma_0}/\omega_p^0$ is never resolved. For all values of $\Delta x$ we keep the CFL ratio $\Delta t/\Delta x = 0.25$ fixed and we run until $t = 100(\omega_p^0)^{-1}$. The results for the $\gamma_0 = 10$ run are shown in Figure 3 (left panel), where we compare the growth of the electric energy associated to the most-unstable wave-vector $k_{\text{max}}$ with the analytic growth rate from equation (30). For almost all resolutions and $\gamma_0$, we observe excellent agreement between numerical results and theoretical predictions. For the lowest resolution of 32 grid points, the numerical results show worse agreement with the analytic prediction, which is expected since the spatial scale of $k_{\text{max}}$ is not accurately resolved. We also show the conservation of total energy (right panel), which is kept to machine accuracy at all times, as expected.

![Figure 3: Relativistic two-stream instability for an electron-positron plasma with initial $\gamma_0 = 10$ for several grid resolutions and CFL ratio 0.25. Left: evolution of the electric field energy associated to the most-unstable wave-vector, compared to linear theory. Right: evolution of the relative error in the conservation of the total energy. In all cases the relativistic skin depth $d_e = c/\omega_p$ is underresolved.](image)

In order to test the code in the electromagnetic limit, we consider the filamentation instability, another classical benchmark for PiC codes. The simplest one-dimensional setup consists again
of two counter-streaming electron-positron beams, with initial velocity $v_0 \perp k$, where $k = k\hat{x}$. The filamentation instability generally dominates over the parallel modes, when the two beams have equal density, in the relativistic regime \[21\]. The maximum growth rate saturates at large $k$ to the value
\[
\Gamma_{\text{max}} = \frac{\omega_{\text{p}} v_0}{\sqrt{2 \gamma_0}}.
\] (31)

We test our algorithm by considering variable spatial and temporal resolutions in the simulation of two symmetric, counter-streaming electron-positron beams. The initial drift velocity is set to $v_0$ along the $y$-direction. A small thermal velocity is added to seed the instability. We initialize the system with increasing $\gamma_0 = 2, 10, 50$, enlarging the simulation box accordingly to $L = 12.8d_e^0, 25.6d_e^0, 51.2d_e^0$. We keep a CFL ratio of 0.125 and vary the resolution from 32 to 256 total cells, with 1000 particles per cell and periodic boundary conditions. Since the growth rate is insensitive of $k$, we can measure the total magnetic energy for comparison with linear theory. The results for $\gamma_0 = 10$ are shown in Figure 4 (left panel), where we find excellent agreement between our simulations and the expected growth rate of the instability (left panel). We note that satisfactory agreement with the analytic growth rate from equation (31) is achieved only for well-resolved runs where the grid spacing is below the relativistic skin depth $d_e$. This is not unexpected, since this corresponds to the fastest-growing wavelength for the filamentation instability. However, the nonrelativistic skin depth can be underresolved, without loss of accuracy or stability of the calculation. Figure 4 (right panel) shows the effect of a variation in the CFL ratio for the $\gamma_0 = 10$ case, demonstrating that the algorithm remains stable and accurate even in the extreme situation where $\Delta t \simeq \Delta x$, which exceeds by far the limit of explicit algorithms.

The results show an extreme robustness of the algorithm in both the electrostatic and electromagnetic regimes, with exact energy conservation at all times even for simulations in which both the relativistic and the nonrelativistic skin depth are heavily underresolved. The performance of the implicit solution procedure greatly surpasses that of explicit algorithms, allowing for much larger $\Delta x$ and $\Delta t$ without loss of stability.

For all runs we employ the kinetic enslavement technique explained in Section 2.3.3. The field hiding approach proves unreliable as it produces acceptable results only for highly resolved
runs where $\Delta x$ and $\Delta t$ do not exceed specific limiting values. When this is not verified, the iteration fails to converge and the results become unphysical. The KE technique, instead, allows for less demanding simulation parameters such has large $\Delta t$ and $\Delta x$, while still retaining the physical correctness of the results.

### 3.2. Relativistic tearing instability

In order to test the implementation of the implicit algorithm more extensively, we apply the method to a relativistic reconnection setup. A classic benchmark for this purpose is the onset of tearing modes in a two-dimensional, relativistic Harris equilibrium. Following [11], a double current sheet is initialized in a periodic two-dimensional domain by choosing the magnetic field profile

$$
B(x) = B_0 \left[ -1 - \tanh \left( \frac{x - x_{cs}}{\Delta} \right) + \tanh \left( \frac{x + x_{cs}}{\Delta} \right) \right] \hat{y},
$$

where $B_0$ is the magnetic field strength away from the current sheet and $\pm x_{cs}$ are the locations of the current sheets. The half-width $\Delta$ characterizes the sheet thickness. The magnetic field is sustained by a population of particles flowing at velocity $\pm U_{s,cs}$ (sign depending on the charge of each species $s$). The associated particle density profile in the current sheet reads

$$
n_{s}(x) = n_{s,cs} \left[ \cosh^{-2} \left( \frac{x + x_{cs}}{\Delta} \right) - \cosh^{-2} \left( \frac{x - x_{cs}}{\Delta} \right) \right],
$$

where $n_{s,cs}$ is the density at the center of the current sheet, measured in the simulation frame.

The typical Harris equilibrium is obtained upon ensuring that the magnetic pressure at the center of the current sheet is balanced by thermal pressure. For a pair plasma of electrons and positrons, neutral ($n_{s,cs} = n_{cs}$ for all species), and populated by equally hot particles (with temperature $T_{s,cs}/(m_sc^2) = \Theta_{s,cs} = \Theta_{cs}$) this condition reads [22]

$$
\frac{B_0^2}{2\mu_0} = 2n_{cs}\Theta_{cs}c^2m_e.
$$

A second condition, deriving deriving from Ampère’s law, ensures that the particle current sustains the magnetic field profile at the current sheet, as

$$
\gamma_{s,cs}U_{s,cs} = -2\frac{\Theta_{cs}c^2m_e}{q_sB_0\Delta}.
$$

Upon deriving the temperature and drift speed of the particles in the current sheet, we initialize such particles at the current sheet locations by loading the corresponding Maxwell-Jüttner distribution according to the technique presented by Melzani et al. [11]. An additional, background population of cold particles is initialized with a nonrelativistic Maxwellian distribution.

The electron-positron equilibrium of a Harris current sheet as described above is subject to tearing instability modes. Magnetic field lines in the vicinity of the sheet are pushed close to each other, eventually undergoing reconnection, that converts magnetic energy into particle kinetic energy. The growth rate of the tearing instability, for pair plasmas, is given by the solution of the dispersion relation as given by [22].

In order to test our algorithm, we run a series of simulations in a two-dimensional periodic domain of size $L_x \times L_y$ discretized with $1024 \times 512$ cells in the $x$ and $y$ directions, respectively. The cell spacing is uniform in both directions and equal to $0.1d_0$. We choose an initial plasma $\beta = 4\mu_0n_0T_0/B_0^2$ of the background equal to 0.01, and we vary the magnetization $\sigma = B_0^2/(2n_0\mu_0c^2) = 5,10,15,20$ while keeping an overdensity ratio $\eta = n_{cs}/n_0 = 3$
constant. In all cases the initial current sheets have a half-width $\Delta = 1d_0^e$ and are located at $x_{cs} = L_x/4, 3L_x/4$. Finally, the time step is taken to be equal to the grid spacing, i.e. $\Delta t/\Delta x = 1$, which is the maximum value allowed by our algorithm. Since these runs only involve pair plasmas with a sufficiently high resolution, we have employed the FH strategy, which is less computationally expensive than the more robust KE scheme.

Figure 6 shows the results of the $\sigma = 20$ run, with the magnitude of the $z$-component of the current $J_z^2$ and the magnetic energy $B^2$ (with selected field lines) shown at several times during and after the linear phase of the instability. After an initial equilibrium phase, thermal noise triggers the tearing instability, which initiates magnetic reconnection. Magnetic islands start growing and merging along the current sheets, in a process that converts magnetic energy into particle kinetic energy.

The conversion of magnetic energy into particle kinetic energy is shown in Figure 5 (left panel) for $\sigma = 20$, where the interplay between particles and fields appears clearly. In order to quantitatively confirm the accuracy of our scheme, we also show the comparison (right panel) between the evolution of the magnetic field energy in the $x$-direction and the theoretical growth rate from [22]. In all cases, we find excellent agreement between the two.

Figure 5: Energy analysis for the $\sigma = 20$ run. Left: time evolution of the magnetic, kinetic, and electric energy, showing the process of conversion of magnetic energy into particle energy. Right: evolution of the magnetic field energy in the $x$-direction, compared to the maximum growth rate from [22]. The magnetic energy in the $y$-direction and the electric energy in the $z$-direction are also shown for comparison.

4. Conclusions
We presented a three-dimensional implementation of the relativistic implicit Particle-in-Cell method for kinetic simulations of plasmas. The algorithm is based on the method presented in [14], but introduces two new strategies for the solution of the implicit equations, that allow for higher-dimensional simulations at accessible computational costs. The algorithm is tested by simulating one- and two-dimensional kinetic instabilities, and shows excellent agreement between analytic predictions for the growth rates and numerical results.

The method conserves energy to machine precision at all times, achieving high physical accuracy of the results. At the same time, this effectively eliminates the constraints imposed on the spatial and temporal resolution by numerical instabilities that affect explicit and semi-implicit PiC methods. The superior stability of the algorithm allows for long-term calculations.
without compromising the accuracy of the results. In all cases where explicit methods are bound to resolve the Debye length, in order to control numerical heating of the particles, our algorithm allows instead for much larger $\Delta x$ and $\Delta t$, which can result in dramatic reductions of the computational cost.

While we refer to a future, more detailed work for a full analysis of computational costs and performances of the various strategies here presented, here we provide a first estimate of the cost of the implicit algorithm, compared to a typical explicit method. In particular, we have considered the two-dimensional reconnection setup from Section 3.2 with the typical production run parameters used for the tearing instability test. First, we have run the setup with the FH strategy (used also for the test) in the implicit implementation; then, we have re-run the same setup with a standard explicit algorithm (namely a Boris push and no iteration on the calculation of the current, which characterises the implicit cycle) implemented in the same code framework. By comparing the runtime of the two algorithms, we have found that a good estimate of the ratio between costs of the implicit and explicit runs is given by the number of nonlinear iterations performed by the latter. This holds essentially because the main cost of one time step is determined by the current gathering step, which is performed only once in the explicit time update. In the implicit cycle, at each time step, a certain number of nonlinear iterations are carried out, each requiring one current gathering. Since the number of nonlinear iterations is typically around 4-5, the resulting computational cost of the implicit run becomes $\sim$4-5 times that of the explicit run. We confirmed this with actual measures of the runtime: the explicit run took $\sim$7900 s to perform 2000 time steps, while the implicit run took $\sim$29200 s, slightly less than 4 times as long.

The high accuracy and reliability of relativistic implicit PiC methods allow for applications in several contexts. Large-scale simulations of cold plasmas, where it is desirable to employ spatial resolutions much coarser than the Debye length (see e.g. [23]), are the ideal field of application of this method. At the same time, long-term calculations where energy errors usually prevent the application of explicit methods, are instead allowed by the energy conservation property of the new scheme (e.g in the case of relativistic turbulence, see e.g. [24]). A final example is provided by coupled algorithms where PiC methods are complemented by magnetohydrodynamic (MHD) calculations (see e.g. [25]). In such cases, the typical MHD length-scales are much larger than the kinetic scales, therefore rendering necessary the employment of PiC methods that can relax the simulation parameters as much as possible, in order to avoid excessive computational costs.

Acknowledgements
FB is supported by projects GOA/2015-014 (2014-2018 KU Leuven) and the Interuniversity Attraction Poles Programme by the Belgian Science Policy Office (IAP P7/08 CHARM). FB is also supported by the Research Fund KU Leuven and Space Weaves RUN project. The computational resources and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation Flanders (FWO) and the Flemish Government - department EWI.
Figure 6: Spatial distribution of the current density (left column) and magnetic energy (right column) at several times during and after the linear growth of the tearing instability, for one of the two current sheets in the simulation. The growing and merging of primary and secondary magnetic islands appears clearly, together with the reconnection of magnetic field lines.
References

[1] A. Chael, M. Rowan, R. Narayan, M. Johnson, and L. Sironi. The role of electron heating physics in images and variability of the Galactic Centre black hole Sagittarius A*. *MNRAS*, 478, 4, 2018.

[2] L. Sironi and A. Spitkovsky. Acceleration of particles at the termination shock of a relativistic striped wind. *ApJ*, 741, 39, 2011.

[3] A. Spitkovsky. Particle acceleration in relativistic collisionless shocks: Fermi process at last? *ApJL*, 682, L5, 2008.

[4] F. Bacchini, V. Olshevsky, S. Poetds, and G. Lapenta. A new particle-in-cell method for modeling magnetized fluids. *CPC*, 210, 79, 2017.

[5] S. G. Bardenhagen and E. M. Kober. The generalized interpolation material point method. *CMES*, 5, 6, 2004.

[6] G. Chen, L. Chacón, and D. C. Barnes. An energy- and charge-conserving, implicit, electrostatic particle-in-cell algorithm. *JCP*, 230, 18, 2011.

[7] S. Markidis, G. Lapenta, and Rizwan-uddin. Multi-scale simulations of plasma with iPIC3D. *Mathematics and Computers in Simulation*, 80, 7, 2010.

[8] G. Lapenta, D. Gonzalez-Herrero, and E. Boella. Multiple-scale kinetic simulations with the energy conserving semi-implicit particle in cell method. *J. Plasma Phys.*, 83, 705830205, 2017.

[9] A. Spitkovsky. *AIP Conf. Proc. 801, Astrophysical Sources of High Energy Particles and Radiation*, ed. T. Bulik, B. Rudak, and G. Madejski (Melville, NY: AIP), 345, 2005.

[10] B. Cerutti, G. R. Werner, D. A. Uzdensky, and M. C. Begelman. Simulations of particle acceleration beyond the classical synchrotron burnoff limit in magnetic reconnection: An explanation of the crab flares. *ApJ*, 770, 147, 2013.

[11] M. Melzani, C. Winisdoerffer, R. Walder, D. Folini, J. M Favre, S. Krastanov, and P. Messmer. *Apar-T*: code, validation, and physical interpretation of particle-in-cell results. *A&A*, 558, A133, 2013.

[12] C.K. Birdsall and A.B. Langdon. *Plasma physics via computer simulation*. IoP Publishing, Bristol, 1991.

[13] S. Markidis and G. Lapenta. The energy conserving particle-in-cell method. *J. Comp. Phys.*, 230, 18, 2011.

[14] G. Lapenta and S. Markidis. Particle acceleration and energy conservation in particle in cell simulations. *Phys. Plasmas*, 072101, 18, 2011.

[15] J. Villasenor and O. Buneman. Rigorous charge conservation for local electromagnetic field solvers. *CPC*, 69, 306, 1992.

[16] Y. Chen and G. Tóth. Gauss’s Law Satisfying Energy-Conserving Semi-Implicit Particle-in-Cell Method. *arXiv:1808.05745*, 2018.

[17] L. Siddi, E. Cazzola, and G. Lapenta. Comparison of Preconditioning Strategies in Energy Conserving Implicit Particle in Cell Methods. *arXiv:1806.04454*, 2018.

[18] W. Taitano. Development of a Jacobian-free Newton-Krylov method with kinetic enslavement to implicitly solve Vlasov-Poisson system in plasma physics. *PhD Thesis*, 2010.

[19] G. Chen and L. Chacón. A multi-dimensional, energy- and charge-conserving, nonlinearly implicit, electromagnetic VlasovDarwin particle-in-cell algorithm. *CPC*, 197,73, 2015.

[20] B. Ripperda, F. Bacchini, J. Teunissen, C. Xia, O. Porth, L. Sironi, G. Lapenta, and R. Keppens. A comprehensive comparison of relativistic particle integrators. *ApJS*, 235:21, 2018.
[21] A. Bret, L. Gremillet, and M. E. Dieckmann. Multidimensional electron beam-plasma instabilities in the relativistic regime. *Physics of Plasmas*, 17(12), 2010.

[22] J. Pétri and J. G. Kirk. Growth rates of the weibel and tearing mode instabilities in a relativistic pair plasma. *Plasma Phys. Control. Fusion*, 49, 1885, 2007.

[23] M. André, A. Vaivads, Y. V. Khotyaintsev, T. Laitinen, H. Nilsson, G. Stenberg, A. Fazakerley, and J. G. Trotignon. Magnetic reconnection and cold plasma at the magnetopause. *GRL*, 37, 22, 2010.

[24] V. Zhdankin, D. A. Uzdensky, G. R. Werner, and M. C. Begelman. Numerical investigation of kinetic turbulence in relativistic pair plasmas i. turbulence statistics. *MNRAS*, 474, 2, 2018.

[25] K. Makwana, R. Keppens, and G. Lapenta. Two-way coupled mhd-pic simulations of magnetic reconnection in magnetic island coalescence. *J. Phys.: Conf. Ser.*, 1031, 012019, 2018.