EUROfusion

EUROFUSION WP14ER–PR(15)66

R. Kleiber et al.

An Explicit Large Time Step Particle–In–Cell Scheme for Nonlinear Gyrokinetic Simulations in the Electromagnetic Regime

Preprint of Paper to be submitted for publication in
Physics of Plasmas

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

An explicit large time step particle-in-cell scheme for nonlinear gyrokinetic simulations in the electromagnetic regime

R. Kleiber,[1,*] R. Hatzky,[2] A. Könies,[1] A. Mishchenko,[1] and E. Sonnendrücker[2]

[1]*Max-Planck-Institut für Plasmaphysik, EURATOM Association, 17491 Greifswald, Germany*

[2]*Max-Planck-Institut für Plasmaphysik, EURATOM Association, 85748 Garching, Germany*

(Dated: October 19, 2015)

Abstract

A new algorithm for electromagnetic gyrokinetic simulations, the so called "pullback transformation scheme" proposed by Mishchenko et al. [Physics of Plasmas, **21**, 092110 (2014)] is motivated as an elaborated explicit time integrator and investigated in detail. Using a numerical dispersion relation valid in slab geometry, it is shown that the linear properties of the scheme are comparable to those of an implicit $v_\parallel$-scheme. A nonlinear extension of the mixed variable formulation, derived consistently from a field Lagrangian, is proposed. The scheme shows excellent numerical properties with a low statistical noise level and a large time step especially for MHD modes. The example of a nonlinear slab tearing mode simulation is used to illustrate the properties of different formulations of the physical model equations.

*Electronic address: ralf.kleiber@ipp.mpg.de

I. INTRODUCTION

Simulating the electromagnetic gyrokinetic equation using particle-in-cell methods is an ongoing challenge. While significant noise reduction was achieved with the introduction of the δf-splitting, which is equivalent to a control variate method [1], electromagnetic simulations including the perturbed parallel vector potential $A_\parallel$ still proved difficult. Early attempts to use the particle velocity $v_\parallel$ as variable ($v_\parallel$-scheme) had severe problems [2] because the partial time derivative of $A_\parallel$ in the equations of motion made their numerical treatment difficult. After Hahm [3] proposed the canonical momentum ($p_\parallel = m_s v_\parallel + q_s A_\parallel$, with m_s, q_s the species mass and charge) as variable, simulations were still possible only for very small values of the plasma-β or if the grid size was smaller than the collisionless skin depth [4]. In this $p_\parallel$-formulation the term $\partial A_\parallel / \partial t$ is not present in the equations of motion but Ampère's law gained an additional, so called "skin", term: a consequence of the first moment of the distribution function with respect to $p_\parallel$ resulting in a non-physical current. Instead, the current obtained using $p_\parallel$ contains an adiabatic part that exactly cancels the skin term. In many applications this "artificial" skin term exceeds the physical term; especially for large scale electromagnetic modes (as e.g. toroidal Alfvén eigenmodes). Moreover, in particle simulations the current is represented by particles while the skin term (being part of the differential operator acting on $A_\parallel$) is represented by quantities on a grid. This makes the cancellation incomplete thus leaving a statistical residual contaminating the physical solution; the origin of the notorious cancellation problem (see e.g. [5]).

It was shown in [6] exemplarily for a slab geometry that the cancellation problem could be mitigated by a proper re-normalisation of the coefficients in Ampère's law in order to match the numerical coverage of velocity space by particles. Later, significant improvement was achieved by the introduction of an adjustable control variate (ACV) scheme [7], i.e. an iterative procedure to obtain a control variate for the adiabatic current. This method made electromagnetic simulations in tokamaks and stellarators practical. Nonetheless, for simulations in certain parameter regimes the necessary time step could become very restrictive. Also the scheme is very sensitive to small inconsistencies in a simulation code; these need to be eliminated.

An alternative to the ACV was developed in [8] where a new set of variables, so called "mixed variables", was proposed and successfully applied to gyrokinetic simulations of MHD

modes. This approach was further extended in the "pullback transformation scheme" (PT-scheme) to also apply to non-MHD modes [9] and showed good performance for different configurations and parameter regimes [10]. Nevertheless, some issues regarding this scheme remained and will be investigated here. Also the original scheme was linearised; so we will derive a nonlinear version of the scheme based on a field theoretical Lagrangian formulation.

The paper is organised as follows: First we give a pedagogical derivation as well as an interpretation of the pullback transformation scheme as a resetting algorithm. In Section III we derive its numerical dispersion relation for slab geometry and compare it with the one resulting from an implicit $v_\parallel$-scheme. In Section IV, we show that the mixed variable formulation can also be derived using a field Lagrangian. Finally, we show simulation results for three simple cases to illustrate the behaviour of the scheme.

II. THE PULLBACK TRANSFORMATION SCHEME

We derive the mixed variable formulation starting from a phase-space particle Lagrangian formulated in $v_\parallel$ and assuming the field equations as given (a derivation using a field Lagrangian ensuring the consistency of the equations and the existence of conservation laws will be presented in Section IV). The result will then be used to construct the PT-scheme.

The equation to be solved for each species s (ions and electrons) is the gyrokinetic equation for the distribution function f_s in $(\vec{R}, v_\parallel, \mu, \alpha)$-space (note that μ is the specific magnetic moment, $\mu = v_\perp^2/(2B)$)

$$\frac{\mathrm{d}f_s}{\mathrm{d}t} = \frac{\partial f_s}{\partial t} + \dot{\vec{R}}\cdot\nabla f_s + \dot{v}_\parallel \frac{\partial f_s}{\partial v_\parallel} + \dot{\mu}\frac{\partial f_s}{\partial \mu} = 0. \tag{1}$$

The equations of motion are defined via the phase-space particle Lagrangian (see e.g. Ref. [11])

$$L_{v_\parallel} = q_s\left[\vec{A}_0 + \left(\frac{m_s}{q_s}v_\parallel + \langle A_\parallel\rangle\right)\vec{b}\right]\cdot\dot{\vec{R}} + \frac{m_s^2}{q_s}\mu\dot{\alpha} - H_{v_\parallel} \tag{2}$$

$$H_{v_\parallel} = \frac{m_s}{2}v_\parallel^2 + m_s\mu B + q_s\langle\Phi\rangle. \tag{3}$$

Here $\vec{A}_0, \vec{b}$ and B are the fixed background's vector potential, unit vector and strength of magnetic field, while $A_\parallel$ describes the perturbation of the parallel vector potential. The gyro-average is indicated by angular brackets.

The field equation for the electrostatic potential perturbation Φ follows, as usual, from quasineutrality. Neglecting the polarisation term for the electrons, linearising the remaining

polarisation term of the ions and using a long wavelength approximation gives

$$-\nabla \cdot \left(\frac{m_{\mathrm{i}} n_0}{B^2} \nabla_\perp \Phi \right) = \sum_s q_s \langle n_s \rangle, \tag{4}$$

where n_0 is the background density, n_s the species density and the summation index s runs over all species. Ampère's law is given by

$$-\nabla \cdot \nabla_\perp (A_{0,\parallel} + A_\parallel) = \mu_0 \sum_s \langle j_{\parallel,s} \rangle \tag{5}$$

with the total current $j_{\parallel,s}$ and $A_{0,\parallel} = \vec{b} \cdot \vec{A}_0$.

We now proceed by decomposing $A_\parallel$ arbitrarily into a symplectic and Hamiltonian part $A_\parallel = A_\parallel^{\mathrm{s}} + A_\parallel^{\mathrm{h}}$ thereby earning an additional degree of freedom, which will be fixed later. After transforming to the new variable $u_\parallel = v_\parallel + \frac{q_s}{m_s} \langle A_\parallel^{\mathrm{h}} \rangle$ one arrives at

$$L_{\mathrm{mixed}} = q_s \vec{A}^* \cdot \dot{\vec{R}} + \frac{m_s^2}{q_s} \mu \dot{\alpha} - H_{\mathrm{mixed}} \tag{6}$$

$$H_{\mathrm{mixed}} = \frac{m_s}{2} u_\parallel^2 + m_s \mu B + q_s (\langle \Phi \rangle - u_\parallel \langle A_\parallel^{\mathrm{h}} \rangle) + \frac{q_s^2}{2 m_s} \langle A_\parallel^{\mathrm{h}} \rangle^2 \tag{7}$$

with

$$\vec{A}^* = \vec{A}_0 + \left(\frac{m_s}{q_s} u_\parallel + \langle A_\parallel^{\mathrm{s}} \rangle \right) \vec{b}. \tag{8}$$

Apart from the quadratic $A_\parallel^{\mathrm{h}}$ term this Lagrangian is identical to the one in Ref. [9]. Contrary to the $p_\parallel$-formulation, where $p_\parallel$ contains $A_\parallel$, in the mixed formulation only the Hamiltonian part $A_\parallel^{\mathrm{h}}$ is absorbed into the new variable $u_\parallel$. The equations of motion (and $\dot{\mu} = 0$) then directly follow from L_{mixed} as

$$\dot{\vec{R}} = \frac{1}{m_s} \frac{\vec{B}^*}{B_\parallel^*} \frac{\partial H}{\partial u_\parallel} + \frac{1}{q_s B_\parallel^*} \vec{b} \times \nabla H \tag{9}$$

$$\dot{u}_\parallel = -\frac{1}{m_s} \frac{\vec{B}^*}{B_\parallel^*} \cdot \nabla H - \frac{q_s}{m_s} \frac{\partial \langle A_\parallel^{\mathrm{s}} \rangle}{\partial t} \tag{10}$$

with $\vec{B}^* = \nabla \times \vec{A}^*$ and $B_\parallel^* = \vec{b} \cdot \vec{B}^*$. From these equations phase-space conservation

$$\frac{\partial B_\parallel^*}{\partial t} + \nabla \cdot (B_\parallel^* \dot{\vec{R}}) + \frac{\partial (B_\parallel^* \dot{u}_\parallel)}{\partial u_\parallel} = 0 \tag{11}$$

follows directly.

While the potential equation (4) remains unchanged, Ampère's equation needs to be transformed into the mixed formulation giving (with the usual simplification of neglecting

the double gyro-average of $A_\parallel^{\rm h}$)

$$-\nabla\cdot\nabla_\perp A_\parallel^{\rm h} + A_\parallel^{\rm h}\sum_s \frac{\mu_0 n_0 q_s^2}{m_s} = \mu_0\sum_s \langle \delta j_{\parallel,s}\rangle + \nabla\cdot\nabla_\perp A_\parallel^{\rm s}. \tag{12}$$

Here f_s has been split into an equilibrium part plus a perturbation, the equilibrium contributions have been cancelled and the nonlinear skin term has been linearised (i.e. n has been replaced by n_0). Since the perturbed current $\delta j_{\parallel,s}$ is now defined as a $u_\parallel$ moment, the skin term only involves $A_\parallel^{\rm h}$. Here, Eq. (12) is regarded as a field equation for $A_\parallel^{\rm h}$ and the $A_\parallel^{\rm s}$ term has been moved to the right-hand side becoming a source term.

It is useful to write out Eqs. (9, 10) for a simple slab geometry, i.e. neglecting the variation of the background magnetic field:

$$\begin{aligned}
\dot{\vec{R}} \;=\; & \left(u_\parallel - \frac{q_s}{m_s}\langle A_\parallel^{\rm h}\rangle\right)\vec{b} + \frac{1}{B_\parallel^*}\vec{b}\times\nabla\left[\langle\Phi\rangle - u_\parallel\left(\langle A_\parallel^{\rm s}\rangle + \langle A_\parallel^{\rm h}\rangle\right) + \frac{q_s}{2m_s}\langle A_\parallel^{\rm h}\rangle^2\right] \\
& + \frac{q_s}{m_s B_\parallel^*}\langle A_\parallel^{\rm h}\rangle\vec{b}\times\nabla\langle A_\parallel^{\rm s}\rangle
\end{aligned} \tag{13}$$

$$\begin{aligned}
\dot{u}_\parallel \;=\; & -\frac{q_s}{m_s}\left\{\vec{b}\cdot\nabla\left[\langle\Phi\rangle - u_\parallel\langle A_\parallel^{\rm h}\rangle + \frac{q_s}{2m_s}\langle A_\parallel^{\rm h}\rangle^2\right] + \frac{\partial\langle A_\parallel^{\rm s}\rangle}{\partial t} + \right. \\
& \left. + \frac{1}{B_\parallel^*}\nabla\langle A_\parallel^{\rm s}\rangle\cdot\vec{b}\times\nabla\left[\langle\Phi\rangle - u_\parallel\langle A_\parallel^{\rm h}\rangle + \frac{q_s}{2m_s}\langle A_\parallel^{\rm h}\rangle^2\right]\right\}.
\end{aligned} \tag{14}$$

Note that these equations reduce to the $v_\parallel$ or $p_\parallel$ equations if $A_\parallel^{\rm h}$, respectively $A_\parallel^{\rm s}$, is set to zero and the remaining field is renamed to $A_\parallel$.

To arrive at an equation more suitable for a numerical treatment it would be advantageous to, firstly eliminate the partial time derivative from Eq. (14) in order to be able to use an explicit time discretisation scheme and, secondly, simplify the remaining equation by removing the dominant term $\vec{b}\cdot\nabla\langle\Phi\rangle$ (as we will see later, with the pullback algorithm $A_\parallel^{\rm h}$ is generally small). One expects the latter to allow for larger time steps in the numerical integration.

The arbitrariness introduced by the splitting of $A_\parallel$ now allows to remove the partial time derivative by postulating a time evolution equation for $A_\parallel^{\rm s}$ which is then used to simplify Eq. (14). One is thus led to postulate

$$\frac{\partial A_\parallel^{\rm s}}{\partial t} + \vec{b}\cdot\nabla\Phi = 0 \tag{15}$$

in order to close the system. Superficially this equation looks like the magneto-hydrodynamic condition $E_\parallel = 0$ but, unlike the MHD case, here it does not tie $A_\parallel$ and Φ together since

$A_\parallel^{\mathrm{h}}$ is still free to evolve. Thus, this postulated relation poses no restriction on the class of possible modes described by the theory. Equation (15) can rather be written as

$$E_\parallel = -\frac{\partial A_\parallel^{\mathrm{h}}}{\partial t}, \tag{16}$$

showing that the parallel electric field is determined by $A_\parallel^{\mathrm{h}}$. The simplified Eq. (14) then becomes

$$\begin{aligned} \dot{u}_\parallel = &- \frac{q_s}{m_s}\Bigg\{\vec{b}\cdot\nabla\left[-u_\parallel\langle A_\parallel^{\mathrm{h}}\rangle + \frac{q_s}{2m_s}\langle A_\parallel^{\mathrm{h}}\rangle^2\right] + \\ &+ \frac{1}{B_\parallel^*}\nabla\langle A_\parallel^{\mathrm{s}}\rangle\cdot\vec{b}\times\nabla\left[\langle\Phi\rangle - u_\parallel\langle A_\parallel^{\mathrm{h}}\rangle + \frac{q_s}{2m_s}\langle A_\parallel^{\mathrm{h}}\rangle^2\right] + \vec{b}\cdot\nabla\langle\Phi\rangle - \langle\vec{b}\cdot\nabla\Phi\rangle\Bigg\}. \end{aligned} \tag{17}$$

When the gyro-average can be neglected (as e.g. for electrons), the last two terms cancel. For arriving at the final equations to be solved, i.e. Eqs. (13, 17) together with Eqs. (4, 12, 15), no numerical argument was involved; the whole procedure served only to remove the time derivative of the parallel vector potential from the equations of motion.

A. Numerical properties of the pullback transformation scheme

We now turn to the question of how the above equations influence a potential numerical algorithm. This is strongly connected to the fact that $A_\parallel^{\mathrm{h}}$ should be kept as small as possible which will be achieved by the so-called resetting procedure.

One immediately observes that, if one neglects the nonlinear terms in Eq. (17), for $A_\parallel^{\mathrm{h}} = 0$ this equation becomes trivial what should allow to use a larger time step in a numerical integrator.

Another advantage of keeping $A_\parallel^{\mathrm{h}}$ small can be seen by splitting f_s into a time independent part $f_{s,0}$ and a time dependent part δf_s according to $f_s = f_{s,0} + \delta f_s$ (the usual δf-splitting used in many PIC codes). The kinetic equation (1) then reads

$$\frac{\mathrm{d}\delta f_s}{\mathrm{d}t} = -\frac{\mathrm{d}f_{s,0}}{\mathrm{d}t}. \tag{18}$$

After linearising this equation, keeping only the term proportional to $\vec{b}\cdot\nabla\langle A_\parallel^{\mathrm{h}}\rangle$ in Eq. (17) and neglecting space gradients in f_0, it reads

$$\frac{\mathrm{d}\delta f_s}{\mathrm{d}t} = \frac{\partial\delta f_s}{\partial t} + u_\parallel\vec{b}\cdot\nabla\delta f_s = -\frac{q_s}{m_s}u_\parallel\vec{b}\cdot\nabla\langle A_\parallel^{\mathrm{h}}\rangle\frac{\partial f_{0,s}}{\partial u_\parallel} \tag{19}$$

what can be written as

$$\frac{\mathrm{d}\delta f_s}{\mathrm{d}t} = -\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{q_s}{m_s}\langle A_\parallel^\mathrm{h}\rangle\frac{\partial f_{0,s}}{\partial u_\parallel}\right) + \frac{q_s}{m_s}E_\parallel\frac{\partial f_{0,s}}{\partial u_\parallel}. \tag{20}$$

Neglecting the $E_\parallel$ term for the moment and integrating along a trajectory gives a change in δf proportional to $(q_s/T_s)u_\parallel\langle A_\parallel^\mathrm{h}\rangle f_{0,s}$ (having assumed $f_{0,s}$ to be a Maxwellian with temperature T_s) which is exactly the adiabatic part of δf. Consequently, by minimising $A_\parallel^\mathrm{h}$ the development of an adiabatic current, one ingredient of the cancellation problem, is mitigated, while the evolving parallel electric field provides the non-adiabatic part. Also when the full equations of motion are taken into account, this mechanism is still at work, leading to a benign numerical behaviour.

Finally, the beneficial influence of having a small $A_\parallel^\mathrm{h}$ also follows from Ampère's law: Neglecting the ion contribution and working in Fourier space it reads in the $p_\parallel$-formulation (defining the electron gyro-radius and the plasma-β as $\rho_\mathrm{e} = \sqrt{2k_\mathrm{B}T_\mathrm{e}m_\mathrm{e}}/(|q_\mathrm{e}|B)$ and $\beta = 2\mu_0 n_0 k_\mathrm{B}T_\mathrm{e}/B^2$)

$$k_\perp^2 A_\parallel + \frac{\beta}{\rho_\mathrm{e}^2}A_\parallel = \mu_0(\delta j_{\parallel,\mathrm{nonad}} + \delta j_{\parallel,\mathrm{ad}}). \tag{21}$$

Here the current was separated into a non-adiabatic part $\delta j_{\parallel,\mathrm{nonad}} = \frac{1}{\mu_0}k_\perp^2 A_\parallel$ and an adiabatic part $\delta j_{\parallel,\mathrm{ad}} = \frac{\beta}{\mu_0\rho_\mathrm{e}^2}A_\parallel$. If the latter is discretised using N_p particles, it can be written as $\delta j_{\parallel,\mathrm{ad}} = \overline{\delta j}_\mathrm{ad} + \epsilon_\mathrm{ad}$, were the expected value $\overline{\delta j}_\mathrm{ad} = \frac{\beta}{\mu_0\rho_\mathrm{e}^2}A_\parallel$ exactly cancels the skin term on the left-hand side of Ampère's law while the statistical error is left over. This is proportional to the variance, which can be computed explicitly here using the background distribution and thus is of the form $\epsilon_\mathrm{ad} = c\frac{\beta}{\mu_0\rho_\mathrm{e}^2}A_\parallel$, with $c \sim \mathcal{O}(1/\sqrt{N_\mathrm{p}})$. Neglecting the statistical error of the non-adiabatic part,

$$\frac{\epsilon_\mathrm{ad}}{\delta j_{\parallel,\mathrm{nonad}}} \sim c\frac{\beta}{k_\perp^2\rho_\mathrm{e}^2} \tag{22}$$

follows. This shows that, especially for small $k_\perp$ and large β, the statistical error of the adiabatic current can become large and swamp the physically more important non-adiabatic part.

On the other hand, in the mixed formulation Ampère's law

$$k_\perp^2(A_\parallel^\mathrm{h} + A_\parallel^\mathrm{s}) + \frac{\beta}{\rho_\mathrm{e}^2}A_\parallel^\mathrm{h} = \mu_0(\delta j_{\parallel,\mathrm{nonad}} + \delta j_{\parallel,\mathrm{ad}}) \tag{23}$$

gives (assuming $A_\parallel^\mathrm{h} \ll A_\parallel^\mathrm{s}$)

$$\frac{\epsilon_\mathrm{ad}}{\delta j_{\parallel,\mathrm{nonad}}} = c\frac{\beta}{k_\perp^2\rho_\mathrm{e}^2}\frac{A_\parallel^\mathrm{h}}{(A_\parallel^\mathrm{h} + A_\parallel^\mathrm{s})} \sim c\frac{\beta}{k_\perp^2\rho_\mathrm{e}^2}\frac{A_\parallel^\mathrm{h}}{A_\parallel^\mathrm{s}} \tag{24}$$

because in this case $\delta j_{\parallel,\mathrm{ad}}$ is proportional to $A_\parallel^\mathrm{h}$ while still $\delta j_{\parallel,\mathrm{nonad}} \sim A_\parallel \approx A_\parallel^\mathrm{s}$ (the statistical error of $A_\parallel^\mathrm{s}$ has been assumed to be negligible since it is mainly determined by the quasi-neutrality equation, which has no cancellation problem). Consequently, the ratio is reduced by a factor of $A_\parallel^\mathrm{h}/A_\parallel^\mathrm{s}$ compared with Eq. (22). Note that the ratio of the two terms involving $A_\parallel^\mathrm{h}$ on the left-hand side of Eq. (23) still scales unfavourably with $\beta/(k_\perp^2\rho_\mathrm{e}^2)$.

B. Resetting procedure

While integrating the system of equations in time, $A_\parallel^\mathrm{h}$ evolves and becomes larger eventually, as we have seen in the last paragraphs, leading to a numerically unfavourable behaviour. This growth of $A_\parallel^\mathrm{h}$ can be prevented by using a resetting algorithm: After a certain number of time steps (with a minimum of one) $A_\parallel^\mathrm{h}$ is reset to zero. Thus, only between the resets can $A_\parallel^\mathrm{h}$ evolve and acquire a finite value. This value is determined by the size of the time step: making the time step smaller can push $A_\parallel^\mathrm{h}$ below every limit.

In order to see what this implies, it is useful to notice that $v_\parallel$ and $u_\parallel$ define two different coordinate systems of the same configuration space, where the $u_\parallel$ coordinate system continuously depends on $A_\parallel^\mathrm{h}$. At the initial time $A_\parallel^\mathrm{h}$ being zero, the two coordinate systems are identical, but then evolve away from each other as the equations above are integrated. Setting $A_\parallel^\mathrm{h}$ to zero has two consequences: $u_\parallel$ must be transformed back to $v_\parallel$ using the transform

$$v_\parallel = u_\parallel - \frac{q_s}{m_s}\langle A_\parallel^\mathrm{h}\rangle \tag{25}$$

and f_s must be pulled back with this coordinate transform. Since f_s is a scalar quantity this pullback transformation is simply given by

$$f_s(v_\parallel) = f_s(u_\parallel). \tag{26}$$

When $A_\parallel^\mathrm{h}$ is reset to zero it is necessary not to lose its contribution to $A_\parallel$. Thus, before the reset is done, the value of $A_\parallel^\mathrm{h}$ must be added to $A_\parallel^\mathrm{s}$, which then becomes equivalent to $A_\parallel$.

A simple graphical representation of the scheme just described is shown in Fig. 1, where the x-axis represents the $v_\parallel$ and $u_\parallel$ coordinate systems with their origins identified. Here it is assumed that the reset procedure is applied after each Δt but the integration of the equations of motion in the $u_\parallel$-system was done with an arbitrarily small time step in order

to be able to draw a continuous trajectory (solid line). A particle follows its trajectory in the $u_\parallel$-system for a time Δt; then it is transformed back into the $v_\parallel$-system (open circles) and $A_\parallel^{\mathrm{h}}$ is reset to zero. Then $u_\parallel = v_\parallel$ and the trajectory in the $u_\parallel$-system starts at this last point for the next Δt. So, in the $v_\parallel$-system the PT-scheme only gives values at discrete time points sampling the continuous trajectory (dashed line), which itself cannot be obtained. In the $u_\parallel$-system trajectories are discontinuous only because this system is shifting with respect to the $v_\parallel$-system and reset each Δt.

In simulations a numerical integration scheme with time step Δt must be used to solve the coupled set of equations consisting of the equations of motion and the potential equations. The above resetting procedure can then be applied after each Δt so that all quantities are given in $v_\parallel$-space with the according magnetic potential $A_\parallel^{\mathrm{s}} = A_\parallel$. Therefore, after finishing the time step by performing the resetting procedure, we end up again in the $v_\parallel$-system. This means that the $u_\parallel$-system is only used to efficiently integrate our set of equations. For the physical interpretation of the results it has no relevance as the diagnostics of the simulation only act on the distribution function $f_s(v_\parallel)$ in the $v_\parallel$-system. Thus, the scheme presented here can be interpreted as a sophisticated time integrator of the gyrokinetic equation in the $v_\parallel$-formulation.

Being back to the $v_\parallel$-system at the end of each time step may also simplify the implementation of collisions since the collision operator can keep its usual form as differential operator in $v_\parallel$. As another beneficial consequence of resetting $A_\parallel^{\mathrm{h}}$ to zero, all terms involving $A_\parallel^{\mathrm{h}}$ vanish when continuing the integration of the equations of motion at the beginning of the new time step.

C. Numerical implementation

In PIC algorithms the numerical particles, so-called markers, are distributed in phase-space by the marker distribution function g_s. In case of an importance sampling method the g_s can differ from the phase-space distribution function f_s of the physical markers. Nevertheless, the markers have to follow the trajectories of the physical particles and hence they obey the same evolution equation, Eq. (1), as f_s: $\mathrm{d}g_s/\mathrm{d}t = 0$. Each marker with index p carries a full-f weight given by the ratio $c_p = f_{s,p}/g_{s,p}$ along its trajectory, which can be used to calculate e.g. the moments of the distribution function. As both, $f_{s,p}$ and

$g_{s,p}$, are conserved along the trajectory of the marker, also c_p is conserved (full-f method). Consequently, the c_p do not change once they have been initialized at the beginning of the simulation. This is true along all trajectories in the $u_\parallel$-system. What is even more important, c_p stays constant when one transforms the velocity from $u_\parallel$ to $v_\parallel$ (Eq. (25)) as part of the pullback transformation Eq. (26). Thus, the markers follow the trajectory sketched in Fig. 1 without changing their full-f weight which can be easily exploited in the the so-called "direct δf-method" (see Ref. [12]) to obtain the noise reduced δf-weight w_p:

$$w_p = \frac{\delta f_{s,p}}{g_{s,p}} = \frac{f_{s,p}(t_0) - f_{s,0}}{g_{s,p}} = c_p - \frac{f_{s,0}}{g_{s,p}}. \tag{27}$$

The pullback transformation can also be formulated in a pure δf-setting: Using the δf-splitting one obtains the pullback transformation, without any approximation, as

$$\delta f_s(v_\parallel) = \delta f_s(u_\parallel) + f_{s,0}\left(v_\parallel + \frac{q_s}{m_s}\langle A_\parallel^{\mathrm{h}}\rangle\right) - f_{s,0}(v_\parallel). \tag{28}$$

Accordingly, the noise reduced weights in the $v_\parallel$-system can be calculated by $w_p = \delta f_{s,p}(v_\parallel)/g_{s,p}$. For a linear simulation Eq. (28) needs to be linearised, resulting in

$$\delta f_s(v_\parallel) = \delta f_s(u_\parallel) + \frac{q_s}{m_s}\frac{\partial f_{s,0}}{\partial u_\parallel}\langle A_\parallel^{\mathrm{h}}\rangle, \tag{29}$$

which is identical to the transformation used in Ref. [9].

For completeness we give here the $u_\parallel$-equations of motion to second order in general geometry:

$$\begin{aligned}
\dot{\vec{R}} \;=\; & \left[u_\parallel - \frac{q_s}{m_s}\langle A_\parallel^{\mathrm{h}}\rangle\right]\vec{b} + \frac{1}{B_\parallel^*}\vec{b}\times\nabla\left[\phi - u_\parallel\left(\langle A_\parallel^{\mathrm{h}}\rangle + \langle A_\parallel^{\mathrm{s}}\rangle\right) + \frac{q_s}{2m_s}\langle A_\parallel^{\mathrm{h}}\rangle^2\right] + \\
& + \frac{q_s}{m_s}\frac{1}{B_\parallel^*}\langle A_\parallel^{\mathrm{h}}\rangle\vec{b}\times\nabla\langle A_\parallel^{\mathrm{s}}\rangle + \frac{m_s}{q_s}\left[\frac{\mu B + u_\parallel^2}{BB_\parallel^*}\,\vec{b}\times\nabla B + \frac{u_\parallel^2}{BB_\parallel^*}(\nabla\times\vec{B})_\perp\right] + \\
& + \vec{b}\times\kappa\left[\frac{u_\parallel}{B_\parallel^*}\left(\langle A_\parallel^{\mathrm{s}}\rangle - \langle A_\parallel^{\mathrm{h}}\rangle\right) - \frac{q_s}{m_s}\frac{1}{B_\parallel^*}\langle A_\parallel^{\mathrm{h}}\rangle\langle A_\parallel^{\mathrm{s}}\rangle\right]
\end{aligned} \tag{30}$$

$$
\begin{aligned}
\dot{u}_\parallel \;=\; & -\frac{q_s}{m_s}\Bigg\{\nabla_\parallel\left[\phi - u_\parallel\langle A_\parallel^{\mathrm{h}}\rangle + \frac{q_s}{2m_s}\langle A_\parallel^{\mathrm{h}}\rangle^2\right] + \partial_t\langle A_\parallel^{\mathrm{s}}\rangle + \\
& + \frac{1}{B_\parallel^*}\nabla\langle A_\parallel^{\mathrm{s}}\rangle\cdot b\times\nabla\left[\phi - u_\parallel\langle A_\parallel^{\mathrm{h}}\rangle + \frac{q_s}{2m_s}\langle A_\parallel^{\mathrm{h}}\rangle^2\right]\Bigg\} + \\
& - \mu\,\nabla B\cdot\left[\vec{b} + \frac{m_s}{q_s}\frac{u_\parallel}{BB_\parallel^*}(\nabla\times\vec{B})_\perp\right] + \\
& - \frac{u_\parallel}{B_\parallel^*}\vec{b}\times\kappa\cdot\nabla\left(\phi - u_\parallel\langle A_\parallel^{\mathrm{h}}\rangle\right) - \frac{\mu}{B_\parallel^*}\left[\vec{b}\times\nabla B\cdot\nabla\langle A_\parallel^{\mathrm{s}}\rangle + \frac{1}{B}\nabla B\cdot(\nabla\times B)_\perp\langle A_\parallel^{\mathrm{s}}\rangle\right] + \\
& - \frac{q_s}{m_s}\frac{1}{B_\parallel^*}\vec{b}\times\kappa\cdot\left[\langle A_\parallel^{\mathrm{s}}\rangle\nabla\left(\phi - u_\parallel\langle A_\parallel^{\mathrm{h}}\rangle\right) + \langle A_\parallel^{\mathrm{h}}\rangle\nabla\langle A_\parallel^{\mathrm{h}}\rangle\left(u_\parallel + \frac{q_s}{m_s}\langle A_\parallel^{\mathrm{s}}\rangle\right)\right]
\end{aligned}
\tag{31}
$$

with

$$
\vec{b}\times\kappa = \frac{1}{B}\left(\vec{b}\times\nabla B + (\nabla\times\vec{B})_\perp\right) \tag{32}
$$

$$
B_\parallel^* = B + \left[\frac{m_s}{q_s}u_\parallel + \langle A_\parallel^{\mathrm{s}}\rangle\right]\vec{b}\cdot(\nabla\times\vec{b}). \tag{33}
$$

Note that all high order terms have to be kept in a nonlinear simulation although they might be small. If not, the proposed scheme may lose its conservation properties, like conservation of phase space volume, energy or total canonical angular momentum.

In summary the pullback transformation scheme consists of three steps: integrating the equations of motion (30), (31) together with the kinetic equation (1); solving the field equations Eqs. (4), (12), (15); resetting, i.e. applying the pullback transformation (25), (26) and, finally, performing the assignment $A_\parallel^{\mathrm{s}} + A_\parallel^{\mathrm{h}} \to A_\parallel^{\mathrm{s}}$ followed by $A_\parallel^{\mathrm{h}} = 0$.

III. NUMERICAL DISPERSION RELATION

In this section, a numerical dispersion relation (see e.g. [13]) taking into account the discrete time integration by an Euler scheme for the $v_\parallel$- and PT-scheme is derived (it would also be possible to include the spatial discretisation, but this is not done here). In order not to make this analysis too cumbersome a kinetic shear Alfvén wave in simple slab geometry is taken as an example. So, we assume all equations to be linearised, written for a simple one-dimensional geometry and all background quantities taken as constant. Furthermore, f_0 is a centred Maxwellian with temperature T and density n; the background magnetic field points into the z-direction; $\Phi, A_\parallel, \delta f$ vary as $\mathrm{e}^{\mathrm{i}(k_\perp x + k_\parallel z)}$; electrons are the only dynamical species (the ions provide the neutralising background) and the gyro-average is neglected.

In the $v_\parallel$-formulation one thus has (a superscript "(1)" indicates the perturbed part of the trajectory)

$$\dot{z} = v_\parallel \tag{34}$$

$$\dot{v}_\parallel = 0 \tag{35}$$

$$\dot{v}_\parallel^{(1)} = -\frac{q_\mathrm{e}}{m_\mathrm{e}}\left(\frac{\partial \Phi}{\partial z} + \frac{\partial A_\parallel}{\partial t}\right) \tag{36}$$

$$\frac{\mathrm{d}\delta f_\mathrm{e}}{\mathrm{d}t} = \frac{\partial \delta f_\mathrm{e}}{\partial t} + \dot{z}\frac{\partial \delta f_\mathrm{e}}{\partial z} = -\dot{v}_\parallel^{(1)}\frac{\partial f_{\mathrm{e},0}}{\partial v_\parallel} \tag{37}$$

$$\frac{m_\mathrm{i} n_0}{B^2} k_\perp^2 \Phi = q_\mathrm{e} \int \delta f_\mathrm{e}\, \mathrm{d}v_\parallel \tag{38}$$

$$k_\perp^2 A_\parallel = \mu_0 q_\mathrm{e} \int v_\parallel \delta f_\mathrm{e}\, \mathrm{d}v_\parallel. \tag{39}$$

Assuming a time dependency $\mathrm{e}^{-\mathrm{i}\omega t}$ and normalising (denoted by a bar) $k_\perp$ and ω to ρ_e and $k_\parallel v_{\mathrm{th,e}}$, the well known dispersion relation

$$D_\mathrm{exact} = 1 - \frac{2\beta}{\bar{k}_\perp^2}\left(\bar{\omega}^2 - \frac{\tilde{\mu}}{\beta}\right)(1 + \bar{\omega} Z(\bar{\omega})) = 0 \tag{40}$$

follows ($v_{\mathrm{th,e}} = \sqrt{2T_\mathrm{e}/m_\mathrm{e}}$ being the electron thermal velocity and $\tilde{\mu}$ the electron to ion mass ratio).

A. Implicit $v_\parallel$-scheme

Using an implicit Euler scheme with step-size Δt, the time derivatives of e.g. $A_\parallel$ (with the above dependency on space and time) can be written as (a subscript n indicates the value at time $t = t_n = n\Delta t$)

$$\left.\frac{\partial A_\parallel}{\partial t}\right|_n \approx \frac{A_{\parallel,n} - A_{\parallel,n-1}}{\Delta t} = -\mathrm{i}\Omega_+ A_{\parallel,n} \tag{41}$$

with

$$\Omega_\pm = \pm\mathrm{i}\frac{1 - \mathrm{e}^{\pm\mathrm{i}\omega\Delta t}}{\Delta t}. \tag{42}$$

The equations of motions become simply $z_n = z_{n-1} + v_{\parallel,n-1}\Delta t$ and $v_{\parallel,n} = \text{const.} = v_\parallel$, what is used to calculate the total derivative of δf_e along the trajectories

$$\left.\frac{\mathrm{d}\delta f_\mathrm{e}}{\mathrm{d}t}\right|_n \approx \frac{\delta f_{\mathrm{e},n} - \delta f_{\mathrm{e},n-1}}{\Delta t} = -\mathrm{i}K_+ \delta f_{\mathrm{e},n} \tag{43}$$

with

$$K_\pm = \pm \mathrm{i} \frac{1 - \mathrm{e}^{\pm \mathrm{i}(\omega - k_\parallel v_\parallel)\Delta t}}{\Delta t}. \tag{44}$$

Computing the Fourier series of the 2π-periodic function whose value on $[-\pi, \pi[$ is $\cos(ax)/\sin(a\pi)$ for any $a \in \mathbb{R}\setminus\mathbb{Z}$, and equating the Fourier series to the function at the point $x = 0$ yields the following identity

$$\frac{\pi}{\sin(a\pi)} = \sum_{n=-\infty}^{+\infty} \frac{(-1)^n}{n+a}. \tag{45}$$

Using this, a straightforward calculation yields the following expression for $K_\pm^{-1}$:

$$K_\pm^{-1} = -\sum_{l=-\infty}^{\infty} \frac{\mathrm{e}^{\pm \frac{\mathrm{i}}{2}(k_\parallel v_\parallel - \omega + \frac{2\pi l}{\Delta t})\Delta t}}{k_\parallel v_\parallel - \omega + \frac{2\pi l}{\Delta t}}. \tag{46}$$

Evaluating the right-hand side of Eq. (37) and combining with Eq. (43) gives

$$\delta f_{\mathrm{e},n} = \frac{q_\mathrm{e}}{m_\mathrm{e}} K_+^{-1} (k_\parallel \Phi_n - \Omega_+ A_{\parallel,n}) \frac{\partial f_{\mathrm{e},0}}{\partial v_\parallel}. \tag{47}$$

Putting Eqs. (46, 47) into the field equations leads to the numerical dispersion relation for the $v_\parallel$-scheme (with $\Omega_\pm$ and $1/\Delta t$ normalised to $k_\parallel v_{\mathrm{th,e}}$)

$$D_{v_\parallel} = 1 - \frac{2\beta}{\bar{k}_\perp^2} \sum_{l=-\infty}^{\infty} \left(\bar{\Omega}_+ F_{1,l} - \frac{\tilde{\mu}}{\beta} F_{0,l} \right) = 0. \tag{48}$$

The integrals $F_{0,l}$ and $F_{1,l}$ defined by

$$F_{m,l} = -\frac{k_\parallel v_{\mathrm{th,e}}^{2-m}}{2n} \int_{-\infty}^{\infty} \frac{\mathrm{e}^{\frac{\mathrm{i}}{2}(k_\parallel v_\parallel - \omega + \frac{2\pi l}{\Delta t})\Delta t}}{k_\parallel v_\parallel - \omega + \frac{2\pi l}{\Delta t}} \frac{\partial f_0}{\partial v_\parallel} v_\parallel^m \, \mathrm{d}v_\parallel \tag{49}$$

can be evaluated to

$$F_{0,l} = \mathrm{e}^{-y^2 + 2\mathrm{i}xy} \left[1 + xZ(x + \mathrm{i}y) \right] \tag{50}$$

$$F_{1,l} = \mathrm{e}^{-y^2 + 2\mathrm{i}xy} \left[x^2 Z(x + \mathrm{i}y) + x - \mathrm{i}y \right] \tag{51}$$

with $x = \bar{\omega} - 2\pi l/\Delta\bar{t}$ and $y = -\Delta\bar{t}/4$ (with $\bar{t} = t k_\parallel v_{\mathrm{th},e}$). For $\Delta\bar{t} \to 0$, which entails $l = 0$, Eq. (48) converges to Eq. (40).

B. Pullback transformation scheme

The equations to be solved in this scheme are

$$\dot{z} = u_\parallel \tag{52}$$

$$\dot{u}_\parallel = 0 \tag{53}$$

$$\dot{u}_\parallel^{(1)} = -\frac{q_\mathrm{e}}{m_\mathrm{e}} u_\parallel \frac{\partial A_\parallel^\mathrm{h}}{\partial z} \tag{54}$$

$$\frac{\mathrm{d}\delta f_\mathrm{e}}{\mathrm{d}t} = -\dot{u}_\parallel^{(1)} \frac{\partial f_{\mathrm{e},0}}{\partial u_\parallel} \tag{55}$$

with the field equations

$$\frac{\partial A_\parallel^\mathrm{s}}{\partial t} = -\frac{\partial \Phi}{\partial z} \tag{56}$$

$$\left(k_\perp^2 + \frac{\mu_0 n_0 q_\mathrm{e}^2}{m_\mathrm{e}}\right) A_\parallel^\mathrm{h} = \mu_0 q_\mathrm{e} \int u_\parallel \delta f_\mathrm{e} \,\mathrm{d}u_\parallel - k_\perp^2 A_\parallel^\mathrm{s} \tag{57}$$

$$\frac{m_\mathrm{i} n_0}{B^2} k_\perp^2 \Phi = q_\mathrm{e} \int \delta f \,\mathrm{d}u_\parallel. \tag{58}$$

The pullback transformation has to be applied after the end of each time step, i.e. quantities at time t_n have to be transformed (denoted by an asterisk) according to

$$\delta f_\mathrm{e}^* = \delta f_\mathrm{e} + \frac{q_\mathrm{e}}{m_\mathrm{e}} \frac{\partial f_{\mathrm{e},0}}{\partial u_\parallel} A_\parallel^\mathrm{h} \tag{59}$$

$$A_\parallel^{\mathrm{s}*} = A_\parallel^\mathrm{s} + A_\parallel^\mathrm{h} \tag{60}$$

$$A_\parallel^{\mathrm{h}*} = 0. \tag{61}$$

As a consequence, when discretising the time derivatives in Eqs. (55, 56) the transformed quantities at time step n have to be used, e.g.

$$\left.\frac{\partial A_\parallel^\mathrm{s}}{\partial t}\right|_n \approx \frac{A_{\parallel,n+1}^\mathrm{s} - A_{\parallel,n}^{\mathrm{s}*}}{\Delta t}. \tag{62}$$

For this scheme all equations involving time derivatives will be discretised explicitly.

Due to Eq. (61) the right-hand side of Eq. (55) becomes zero, leading to

$$\delta f_{\mathrm{e},n} = \frac{\mathrm{i}}{\Delta t K_-} \frac{q_\mathrm{e}}{m_\mathrm{e}} A_{\parallel,n}^\mathrm{h} \frac{\partial f_{\mathrm{e},0}}{\partial u_\parallel}, \tag{63}$$

while from Eq. (56) using Eq. (60)

$$A_{\parallel,n}^\mathrm{s} = \frac{\mathrm{i}}{\Delta t \Omega_-} (A_{\parallel,n}^\mathrm{h} - \mathrm{i} k_\parallel \Phi_n \Delta t) \tag{64}$$

follows. From Eq. (64) together with the fact that $A_\parallel^{\mathrm{h}}$ is reset to zero after each time step follows the important result $A_\parallel^{\mathrm{h}} \sim \Delta t$. Putting everything together leads finally to the numerical dispersion relation for the PT-scheme

$$D_{\mathrm{PT}} = 1 - \mathrm{i}\bar{\Omega}_- \Delta t \left(1 + \frac{\beta}{\bar{k}_\perp}\right) - \frac{2\beta}{\bar{k}_\perp^2} \sum_{l=-\infty}^{\infty} \left(\bar{\Omega}_- F_{1,l} - \frac{\tilde{\mu}}{\beta} F_{0,l}\right) = 0 \tag{65}$$

where $F_{0,l}, F_{1,l}$ now must be evaluated using $y = \Delta\bar{t}/4$.

C. Results of the dispersion relation

Figure 2 shows the real and imaginary part of $\bar{\omega}$ obtained from the three dispersion relations as a function of β.

In the case with small $\bar{k}_\perp$ (left-hand side of Figure 2) the time step of $\Delta\bar{t} = 10^{-6}$ is sufficiently small for the numerical schemes to give the same real part as the exact solution. In fact the frequency is very robust with respect to changes in time step: For the relative frequency error to stay below $\approx 1\%$ (at $\beta = 0.1$) the time step needs only to be smaller than approximately 10^{-2}. However, the (very small) damping rate is not as well approximated as the frequency: the numerical dispersion relations show an artificial damping about 70% higher than the exact one for $\beta = 0.1$. Decreasing $\Delta\bar{t}$ by a factor of ten decreases the artificial damping by roughly the same factor. In total, the numerical schemes consistently give a lower imaginary part than the exact solution, showing that they are always numerically stable.

In the large $\bar{k}_\perp$ case (right-hand side of Figure 2), where the exact solution is strongly damped (note the different y-axis scaling of the lower left and right plot), a relatively large time step of $\Delta\bar{t} = 1$ has been chosen in order to highlight the behaviour of the imaginary part (here, the relative error of the frequency and damping rate is of the same order): The numerical schemes show artificial damping for $\beta > 0.05$ and act somewhat destabilising for lower β but, due to the strong intrinsic physical damping, this does not lead to numerical instability.

It is also possible to use an explicit discretisation for the $v_\parallel$-scheme but we found this leading to strong numerical instabilities when β was not small enough.

Nevertheless, the most important observation is that the $v_\parallel$- and the PT-scheme give nearly the same results for both sets of parameters (with the latter scheme having slightly

less numerical damping). Thus, at least for this simple case, the two schemes are equivalent but the PT-scheme has the advantage of being explicit.

IV. FIELD THEORETICAL FORMULATION

Following e.g. Ref. [14], [15], as modern gyrokinetic theory is derived so as to conserve some structure of the underlying Vlasov-Maxwell equations, it is desirable with respect to consistency and in order to exactly conserve the main physical invariants, to derive the gyrokinetic equation from a variational principle $\delta S = 0$ where $S = \int \mathcal{L}\,\mathrm{d}t$ is the action of a Lagrangian $\mathcal{L}$. Starting from Eqs. (2, 3) and performing the coordinate transformation $p_\parallel = m_s v_\parallel + q_s A_\parallel$ gives the $p_\parallel$ particle Lagrangian

$$L_{p_\parallel} = q_s\left(\vec{A}_0 + \frac{1}{q_s}p_\parallel\vec{b}\right)\cdot\dot{\vec{R}} + \frac{m_s^2}{q_s}\mu\dot{\alpha} - H_{p_\parallel} \tag{66}$$

$$H_{p_\parallel} = \frac{p_\parallel^2}{2m_s} + m_s\mu B + q_s\left(\langle\Phi\rangle - \frac{p_\parallel}{m_s}\langle A_\parallel\rangle\right) + \frac{q_s^2}{2m_s}\langle A_\parallel\rangle^2. \tag{67}$$

From this we motivate the $p_\parallel$ field-Lagrangian (the polarisation term has been linearised using the equilibrium distribution function $f_{0,s}$)

$$\begin{aligned}\mathcal{L}_{p_\parallel} = \sum_s \Bigg\{ \int f_s \Bigg[q_s\left(\vec{A}_0 + \frac{1}{q_s}p_\parallel\vec{b}\right)\cdot\dot{\vec{R}} + \frac{m_s^2}{q_s}\mu\dot{\alpha} - \frac{p_\parallel^2}{2m_s} - m_s\mu B - q_s\left(\langle\Phi\rangle - \frac{p_\parallel}{m_s}\langle A_\parallel\rangle\right) + \\ - \frac{q_s^2}{2m_s}\langle A_\parallel\rangle^2\Bigg]\mathrm{d}W_0\,\mathrm{d}V_0 + \int f_{0,s}\frac{m_s}{2B^2}(\nabla_\perp\Phi)^2\mathrm{d}W\,\mathrm{d}V\Bigg\} - \frac{1}{2\mu_0}\int(\nabla_\perp A_\parallel)^2\,\mathrm{d}V\end{aligned} \tag{68}$$

where $\mathrm{d}W_0$ and $\mathrm{d}V_0$ are the velocity and position space volume element (a subscript "0" indicates that the integration has to be done with respect to the initial phase-space positions of the trajectories [14]).

After introducing into Eq. (68) the splitting $A_\parallel = A_\parallel^{\mathrm{s}} + A_\parallel^{\mathrm{h}}$ and the new variable $u_\parallel = p_\parallel - \frac{q_s}{m_s}\langle A_\parallel^{\mathrm{s}}\rangle$, a Lagrange multiplier λ is finally used to incorporate $\frac{\partial A_\parallel^{\mathrm{s}}}{\partial t} + \vec{b}\cdot\nabla\Phi = 0$ giving

$$\begin{aligned}\mathcal{L}_{\mathrm{mixed}} = \sum_s \Bigg\{ \int f_s \Bigg[q_s\vec{A}^*\cdot\dot{\vec{R}} + \frac{m_s^2}{q_s}\mu\dot{\alpha} - \frac{m_s}{2}u_\parallel^2 - \mu B - q_s\left(\langle\Phi\rangle - u_\parallel\langle A_\parallel^{\mathrm{h}}\rangle\right) + \\ - \frac{q_s^2}{2m_s}\langle A_\parallel^{\mathrm{h}}\rangle^2\Bigg]\mathrm{d}W_0\,\mathrm{d}V_0 + \int f_{0,s}\frac{m_s}{2B^2}(\nabla_\perp\Phi)^2\mathrm{d}W\,\mathrm{d}V\Bigg\} - \frac{1}{2\mu_0}\int(\nabla_\perp A_\parallel)^2\,\mathrm{d}V + \\ + \int\lambda\left(\frac{\partial A_\parallel^{\mathrm{s}}}{\partial t} + \vec{b}\cdot\nabla\Phi\right)\mathrm{d}V\end{aligned} \tag{69}$$

with

$$\vec{A}^* = \vec{A}_0 + \left(\frac{m_s}{q_s} u_\parallel + \langle A^{\mathrm{s}}_\parallel \rangle \right) \vec{b}. \tag{70}$$

Variation with respect to particle positions, by construction, recovers Eqs. (6, 7). Performing the variation with respect to $\Phi, A^{\mathrm{h}}_\parallel, A^{\mathrm{s}}_\parallel$ and λ gives (using the relation $\vec{b}\cdot\dot{\vec{R}} = \frac{1}{m}\frac{\partial H}{\partial u_\parallel}$ following from Eq. (9)) the field equations.

One needs to be careful here with the gyro-averaging operator: Since it is not self-adjoint (it can be written as an integral operator with non-symmetric kernel), it is necessary to introduce its adjoint denoted by $\langle\cdot\rangle^\dagger$ as

$$\int \langle g \rangle^\dagger h \, \mathrm{d}V = \int g \langle h \rangle \, \mathrm{d}V \tag{71}$$

for arbitrary functions g and h (see also [16]). Also, by a slight abuse of notation, we define $\langle n \rangle^\dagger = \int \langle f \rangle^\dagger \, \mathrm{d}W$ and similarly for the other velocity moments. However, in a finite element PIC code as it is used here, one can easily handle the adjoint operator since the field equations are formulated in a weak form.

Finally, the field equations read as follows

$$-\nabla\cdot\frac{\sum_s n_{0,s} m_s}{B^2}\nabla_\perp \Phi = \sum_s q_s \langle n_s \rangle^\dagger + \vec{b}\cdot\nabla\lambda \tag{72}$$

$$-\frac{1}{\mu_0}\nabla\cdot\nabla_\perp A_\parallel + \sum_s \left(\frac{q_s^2}{m_s} \mathcal{S} A^{\mathrm{h}}_\parallel \right) = \sum_s \langle j_{\parallel,s} \rangle^\dagger \tag{73}$$

$$-\frac{1}{\mu_0}\nabla\cdot\nabla_\perp A_\parallel + \sum_s \left(\frac{q_s^2}{m_s} \mathcal{S} A^{\mathrm{h}}_\parallel \right) = \sum_s \langle j_{\parallel,s} \rangle^\dagger - \frac{\partial\lambda}{\partial t} \tag{74}$$

$$\frac{\partial A^{\mathrm{s}}_\parallel}{\partial t} + \vec{b}\cdot\nabla\Phi = 0 \tag{75}$$

with

$$\mathcal{S} = \int \langle f_s \langle\cdot\rangle \rangle^\dagger \, \mathrm{d}W \tag{76}$$

the skin-depth operator (in order to arrive at the usual field equations one needs to use the approximation $\mathcal{S} \approx n_s$). Combining Eqs. (73, 74) gives the time evolution equation for λ as $\partial\lambda/\partial t = 0$. Utilising the initial conditions for λ we choose $\lambda(t=0) = \lambda_0$ with $\vec{b}\cdot\nabla\lambda_0 = 0$. The λ term in Eq. (72) then vanishes, returning the usual field equation for Φ.

Defining a Lagrangian density $\tilde{L}$ by $\mathcal{L} = \int \tilde{L} \, \mathrm{d}V$ Noether's theorem (see e.g. [14]) gives a conserved quantity C

$$C = \int \left[\left(\sum_j \frac{\partial \tilde{L}}{\partial \eta^j_{,t}} \eta^j_{,t} - \tilde{L} \right) X^t + \sum_{i,j} \frac{\partial \tilde{L}}{\partial \eta^j_{,t}} \frac{\partial \eta^j}{\partial R^i} X^i \right] \mathrm{d}V \tag{77}$$

if the action is invariant with respect to an infinitesimal transformation. Here, $X = (X^i, X^t)$, with $i = 1 \ldots 3$, is the infinitesimal generator of the transformation in four dimensional $\vec{R}$-t-space and η^j denotes the dynamical fields appearing in $\tilde{L}$ (with $\eta^j_{,t}$ their time derivative). From time-translation invariance thus follows the conserved energy, using Eqs. (72, 73, 75), as

$$\begin{aligned} E_{\text{total}} = &\sum_s \int f_s \left[\frac{m_s}{2} u_\parallel^2 + \mu B \right] \mathrm{d}W \, \mathrm{d}V + \\ &+ \frac{1}{2} \sum_s \int \left[q_s \langle n_s \rangle^\dagger \Phi + \langle j_{\parallel,s} \rangle^\dagger (A_\parallel^{\mathrm{s}} - A_\parallel^{\mathrm{h}}) - \frac{q_s^2 n_s}{m_s} A_\parallel^{\mathrm{s}} \mathcal{S} A_\parallel^{\mathrm{h}} \right] \mathrm{d}V. \end{aligned} \tag{78}$$

Due to the term $A_\parallel^{\mathrm{s}} \mathcal{S} A_\parallel^{\mathrm{h}}$, this is a non-trivial extension of the total energies in the $v_\parallel$- or $p_\parallel$-formalism. Similarly, an expression for conservation of angular momentum in axisymmetric systems can be obtained.

Approximate formulations

In this Section, we investigate consequences of approximations regarding higher order terms. These approximations are often done in either the particle Lagrangian or the equations of motion and result in different physical models.

Starting from Eqs. (6, 7) the resulting mixed equations of motion, in slab geometry, are given by Eqs. (13, 14). This, what we call model "M", is the most complete and consistent physical model. In Ref. [9] the equations of motion have been linearised by neglecting quadratic field terms, thereby neglecting especially the "flutter" term $1/B_\parallel^* \nabla \langle A_\parallel^{\mathrm{s}} \rangle \cdot (\vec{b} \times \nabla \langle \Phi \rangle)$ in Eq. (14). We call the resulting equations of motion model "Mlin".

Sometimes the skin term $f_s \frac{q_s^2}{2m_s} \langle A_\parallel \rangle^2$ in Eq. (68) is linearised by replacing f_s by $f_{0,s}$. This approximation has the consequence of removing the nonlinear term from the equations of motion; after transforming to mixed variables this then gives the often used model "PlinM" where Eq. (7) is modified to

$$H_{\text{mixed}} = \frac{m_s}{2} u_\parallel^2 + m_s \mu B + q_s (\langle \Phi \rangle - u_\parallel \langle A_\parallel^{\mathrm{h}} \rangle) - \frac{q_s^2}{m_s} \langle A_\parallel^{\mathrm{s}} \rangle \left(\frac{1}{2} \langle A_\parallel^{\mathrm{s}} \rangle + \langle A_\parallel^{\mathrm{h}} \rangle \right). \tag{79}$$

Consequently, the corresponding equations of motion are Eqs. (13, 14) with $\langle A_\parallel^{\mathrm{h}} \rangle^2/2$ replaced by $-\langle A_\parallel^{\mathrm{s}} \rangle (\langle A_\parallel^{\mathrm{s}} \rangle/2 + \langle A_\parallel^{\mathrm{h}} \rangle)$ which especially introduces the non-negligible term $\langle A_\parallel^{\mathrm{s}} \rangle^2$.

In deriving the Lagrangian density for the mixed formulation this linearisation must not be made, since it breaks some necessary cancellations between quadratic quantities

in the bracket following f_s. In cases where having a linear skin term is desirable, this approximation must be made in the final field equations. Although this might ease numerical implementation, one should be aware of the consequences, i.e. of breaking consistency.

V. NUMERICAL RESULTS

The behaviour of the pullback transformation scheme will now be illustrated for three examples in slab geometry: the linear Alfvén wave as an example for an MHD mode, the linear ion-temperature-gradient-driven mode (ITG) and the (linear and nonlinear) collisionless tearing mode as two examples for non-MHD-like modes. The gyrokinetic code GYGLES [6], [17] was used for these simulations.

A. Convergence with time step and number of markers

The frequency ω (normalised to the analytic frequency ω_A from Eq. (40)) of the Alfvén wave in slab geometry as a function of the time step Δt can be found in Figure 3, left. The same parameters as in Ref. [7] were used: $\beta = 6.08\%$, $k_\perp \rho_\mathrm{e} = 4.54 \cdot 10^{-4}$, $k_\parallel \rho_\mathrm{e} = 1.23 \cdot 10^{-5}$, $B = 2.5\,\mathrm{T}$, $T_\mathrm{e} = 5\,\mathrm{keV}$. Only the electron dynamics was considered, i.e. the ions only provided a neutralising background. In contrast to Ref. [7] the phase factor extraction was not used which required the grid resolution in the z-direction to be increased in order to reach convergence.

Curves are shown for the ACV-scheme (circles) and the PT-scheme (squares) using 10^4 electron markers for both. The ACV-scheme uses one iteration to adjust its control variate. For comparison also the result of the mixed-variable formulation (triangles) without the resetting procedure (corresponding to the scheme proposed in [8]) is shown ($4 \cdot 10^5$ electron markers were used). All schemes converge to the correct result with the same rate (determined by a fourth-order Runge-Kutta time integration) but the PT-scheme allows a time step approximately 50 times larger than the ACV-scheme. Already the mixed-variable formulation gives an improvement: the time step is a factor of 10 larger than that needed by the ACV-scheme. Note that for all curves in Fig. 3 the frequencies are already accurate up to 5%.

Convergence with the number of electron markers $N_\mathrm{p,e}$ is shown in Figure 3, right.

The convergence of the ACV-scheme (circles, $\Delta t = 10^{-9}\,\mathrm{s}$) and the PT-scheme (squares, $\Delta t = 5{\cdot}10^{-8}\,\mathrm{s}$) is very similar. Comparing with the result from the mixed-variable formulation without application of the resetting procedure (triangles, $\Delta t = 10^{-8}\,\mathrm{s}$) shows that the inclusion of the resetting procedure reduces the number of needed markers by a factor of about 50. Thus, the PT-Scheme behaves comparably to the ACV-scheme when it comes to its very low statistical noise property.

For the tearing mode simulations, the basic geometry was the same as in Ref. [18]: A slab with a strong guiding field B_z in the z-direction and a sheared field

$$B_y = B_{y,0}\,\mathrm{erf}\left(\frac{x - \frac{1}{2}L_x}{L_\mathrm{s}}\right) \tag{80}$$

pointing in the y-direction. From B_y follows the electron current used in a shifted Maxwellian. The parameters $B_{y,0}, L_x, L_\mathrm{s}$ determine the strength of the field, the x-width of the slab and the shear length, respectively. The size of the slab is 10×10 in x and y (all lengths normalised to the ion gyro-radius); the other parameters are $L_\mathrm{s} = 0.5, B_{y,0}/B_z = 0.02, \beta = 4\cdot 10^{-3}$. In the simulations 128 grid points in the x-direction and 10^6 electron markers were used (with the ions providing a neutralising background) and only the lowest k_y mode was kept. In the left of Figure 4 the frequency of the linear tearing mode is displayed for the ACV- and the PT-scheme. The latter scheme permits only a modestly larger time step.

Results for the linear electromagnetic ITG mode are shown at the right of Figure 4. The simulation parameters are similar to the ones used in the ITG simulation of a ϑ-pinch in Ref. [19]. However, we take here the slab approximation with a tanh-like profile for the ion and electron temperature, a flat density profile, $\beta = 20\%$ and 10^6 ion (with gyro-averaging neglected) and electron markers. Again, the time step convergence of the PT-scheme is better than that of the ACV-scheme, allowing for approximately a factor of two larger time step. However, the convergence of the frequency of the ITG mode with regard to the marker number shows that both the PT-scheme and the ACV-scheme behave the same.

From these examples one concludes that for MHD modes the PT-scheme allows a much larger time step compared with the ACV-scheme, while for strongly non-MHD-like modes it may give a factor of two or less. In addition, the PT-scheme shares the very low statistical noise level property with the ACV-scheme. This is the result of the resetting procedure which diminishes $A_\parallel^\mathrm{h}$ and therefore the adiabatic part of the distribution function in some

cases dramatically.

B. Nonlinear tearing mode

The collisionless slab tearing mode provides a simple but nontrivial nonlinear test case for comparing different simulation schemes. Except for a smaller time step of $\Delta t = 5{\cdot}10^{-9}\,\mathrm{s}$, which had to be adjusted for the nonlinear simulation, all parameters were the same as used in the last section. However, four times as many, i.e. $4{\cdot}10^6$ electron and ion markers were used in the simulation and gyro-averaging, though merely of small importance for this set of parameters, was taken into account.

The different mixed-variable models described at the end of Section IV are now used within the PT-scheme; i.e. relation Eq. (15) is used to eliminate $\partial A_\parallel^{\mathrm{s}}/\partial t$ and the resetting procedure is applied. Especially note that in the PT-scheme $A_\parallel^{\mathrm{h}}$ is set to zero at the beginning of each time step and will evolve to small values during the sub-time steps of the multi-step integrator. In general, we have $A_\parallel^{\mathrm{h}} \to 0$ for $\Delta t \to 0$. As a consequence, in all the models the $A_\parallel^{\mathrm{h}}$ terms in the equations of motion are typically very small.

The results of the simulations are displayed in Fig. 5, where the total fluctuating field energy for each simulation run is shown. In the linear phase the energies for all models are the same but they differ significantly in the initial nonlinear phase. For later times the energies for models "M" and "Plin" oscillate about a constant value, while the energy for model "Mlin" shows a downward trend and finally reaches the level of model "PlinM" after a long time. At the end of the simulation the energies for models "PlinM" and "Mlin" are about 15% higher than for model "M". In the linear phase the mode structure of $A_\parallel$ for all models is, as expected, identical (Fig. 6, top, left). In the early nonlinear phase (Fig. 6, top, right) differences start to emerge and all models give somewhat different results: The maxima of $A_\parallel$ for models "PlinM" and "Mlin" are about 5%, respectively 15%, higher than for model "M". But more importantly, model "Mlin" gives a significantly flatter shape of $A_\parallel$ near the resonant layer at $x = 0$ compared with the other models. In the late nonlinear phase (Fig. 6, bottom) "Mlin" and "PlinM" give the same amplitude, somewhat in excess of the amplitude from model "M". Nevertheless, "Mlin" results in an unexpected asymmetric shape for $A_\parallel$, while model "PlinM" leads to a somewhat flatter curve than model "M". This behaviour points at an altered dynamics near the resonant layer in model "Mlin". We

conclude that for consistency of the mixed formulation it is necessary to keep all terms up to second order in the equations of motion.

VI. CONCLUSIONS

The pullback transformation scheme proposed in [9] is a new scheme for the solution of the electromagnetic gyrokinetic equation by particle methods. While the scheme could already be used for linear simulations of fusion devices [10], we here aimed at investigating its numerical behaviour more fundamentally.

The scheme combines a splitting of the parallel vector potential $A_\parallel$ into two parts $A_\parallel^{\mathrm{s}}$ and $A_\parallel^{\mathrm{h}}$ with a special coordinate transformation in parallel velocity in order to gain the freedom to postulate a field equation for $A_\parallel^{\mathrm{s}}$. This allows us to shift part of the parallel dynamics from the particles to the field equation, thus reducing the parallel dynamics significantly. The resulting mixed formulation equations are then combined with a resetting algorithm to yield an efficient explicit numerical integration method (the pullback transformation scheme). From a numerical dispersion relation for the model of a slab Alfvén wave we could conclude that the scheme behaves very similarly to an implicit $v_\parallel$-scheme.

The scheme avoids a large adiabatic part in the distribution function which usually triggers an extremely high statistical noise level in conventional PIC methods. Also, it allows the time step to be increased, in some cases dramatically. Especially for MHD modes more than an order of magnitude could be achieved. This makes the scheme superior to other schemes, such as e.g. the adjustable control variate scheme, which only addresses the statistical noise level problem.

We were able to generalize the scheme so that it can be used for nonlinear simulations and to consistently derive the mixed formulation from a field Lagrangian, thus putting it on solid theoretical ground. In addition, an expression for the conserved energy was given. The example of a nonlinear simulation for the slab tearing mode showed that there can be significant differences between simplified physical models, and that linearisation at any stage of the derivation and implementation should be avoided.

Overall, the scheme is now mature and ready to be used for production runs of nonlinear electromagnetic simulations.

Acknowledgements

We would like to thank J. W. Burby, M. Restelli and A. Bottino for helpful discussions. This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014–2018 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

[1] A. Y. Aydemir, A unified Monte Carlo interpretation of particle simulations and applications to non–neutral plasmas, Phys. Plasmas 1 (1994) 822.

[2] J. V. W. Reynders, Gyrokinetic simulations of finite–β plasmas on parallel architectures, PhD Thesis, Princeton University, 1992.

[3] T. S. Hahm, W. W. Lee, A. Brizard, Nonlinear gyrokinetic theory for finite-beta plasma, Phys. Fluids 31 (1988) 1940.

[4] J. C. Cummings, Gyrokinetic simulations of finite–β and self–generated sheared–flow effects on pressure–gradient–driven instabilities, PhD Thesis, Princeton University, 1995.

[5] Y. Chen, S. E. Parker, A δf particle method for gyrokinetic simulations with kinetic electrons and electromagnetic perturbations, J. Comp. Phys. 189 (2003) 463.

[6] A. Mishchenko, R. Hatzky, A. Könies, Conventional δf–particle simulations of electromagnetic perturbations with finite elements, Phys. Plasmas 11 (2004) 5408.

[7] R. Hatzky, A. Könies, A. Mishchenko, Electromagnetic gyrokinetic PIC simulation with an adjustable control variates method, J. Comp. Phys. 225 (2007) 568.

[8] A. Mishchenko, M. Cole, R. Kleiber, A. Könies, New variables for gyrokinetic electromagnetic simulations, Phys. Plasmas 21 (2014) 052113.

[9] A. Mishchenko, A. Könies, R. Kleiber, M. Cole, Pullback transformation in gyrokinetic electromagnetic simulations, Phys. Plasmas 21 (2014) 092110.

[10] A. Mishchenko, M. Borchardt, M. Cole, R. Hatzky, T. Fehér, R. Kleiber, A. Könies, A. Zocco, Global linear gyrokinetic particle–in–cell simulations including electromagnetic effects in shaped plasmas, Nucl. Fusion 55 (2015) 053006.

[11] A. J. Brizard, T. S. Hahm, Foundations of nonlinear gyrokinetic theory, Rev. Mod. Phys. 79 (2007) 421.

[12] S. J. Allfrey, R. Hatzky, A revised δf algorithm for nonlinear PIC simulation, Comp. Phys. Commun. 154 (2003) 98.

[13] A. B. Langdon, Analysis of the time integration in plasma simulation, J. Comp. Phys. 30 (1979) 202.

[14] H. Sugama, Gyrokinetic field theory, Phys. Plasmas 7 (2000) 466.

[15] B. Scott, J. Smirnov, Energetic consistency and momentum conservation in the gyrokinetic

description of tokamak plasmas, Phys. Plasmas 17 (2010) 112302.

[16] A. Bottino, E. Sonnendrücker, Monte Carlo particle–in–cell methods for the simulation of the Vlasov–Maxwell gyrokinetic equations, Journal of Plasma Physics 81 (2015) 435810501.

[17] M. Fivaz, S. Brunner, G. de Ridder, O. Sauter, T. M. Tran, J. Vaclavik, L. Villard, K. Appert, Finite element approach to global gyrokinetic Particle–In–Cell simulations using magnetic coordinates, Comp. Phys. Commun. 111 (1998) 27.

[18] O. Zacharias, R. Kleiber, R. Hatzky, Gyrokinetic simulations of collisionless tearing modes, J. Phys.: Conf. Ser. 401 (2012) 012026.

[19] R. Hatzky, T. M. Tran, A. Könies, R. Kleiber, S. J. Allfrey, Energy conservation in a nonlinear gyrokinetic particle–in–cell code for ion–temperature–gradient–driven modes in θ–pinch geometry, Phys. Plasmas 9 (2002) 898.

Captions

Fig. 1: Graphical representation of the resetting procedure: The solid/dashed line shows the trajectory of a particle in $u_\parallel$, respectively $v_\parallel$-coordinates. After each Δt the resetting is applied and $u_\parallel$ (filled circles) of the particle is transformed to $v_\parallel$ (open circles). The transformation itself is just a shift by $(q_s/m_s)A_\parallel^{\mathrm{h}}$.

Fig. 2: Real (top) and imaginary (bottom) part of the solution $\bar{\omega}$ of the exact (dotted) and of the numerical dispersion relations ($v_\parallel$-scheme: dashed, pullback transformation scheme: solid) as a function of β. Left: $\bar{k}_\perp = 2{\cdot}10^{-4}, \Delta\bar{t} = 10^{-6}$, right: $\bar{k}_\perp = 0.25, \Delta\bar{t} = 1$.

Fig. 3: Convergence with time step Δt (left) and electron marker number $N_{\mathrm{p,e}}$ (right) for the frequency ω of the Alfvén wave using the adjustable control variate scheme (circles), the mixed formulation without pullback transformation (triangles) and the pullback transformation scheme (squares).

Fig. 4: Convergence with time step Δt for the frequency ω of the tearing mode (left) and the ITG mode (right) using the adjustable control variate scheme (circles) and the pullback transformation scheme (squares).

Fig. 5: Total fluctuating field energy as a function of time for the three models "Mlin" (dotted), "PlinM" (dashed) and "M" (solid).

Fig. 6: The magnetic potential $A_\parallel$ at $t = 8{\cdot}10^{-7}\,\mathrm{s}$ (top, left), $t = 2{\cdot}10^{-6}\,\mathrm{s}$ (top, right) and $t = 10^{-5}\,\mathrm{s}$ (bottom) for the three models "Mlin" (dotted), "PlinM" (dashed), "M" (solid).

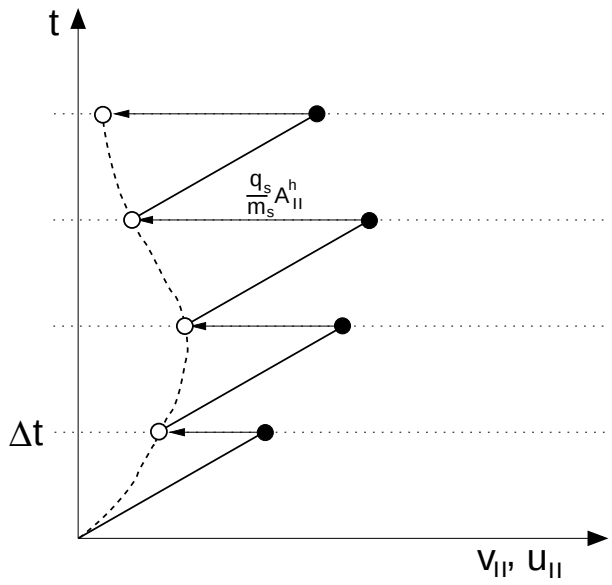

FIG. 1: Kleiber et al.

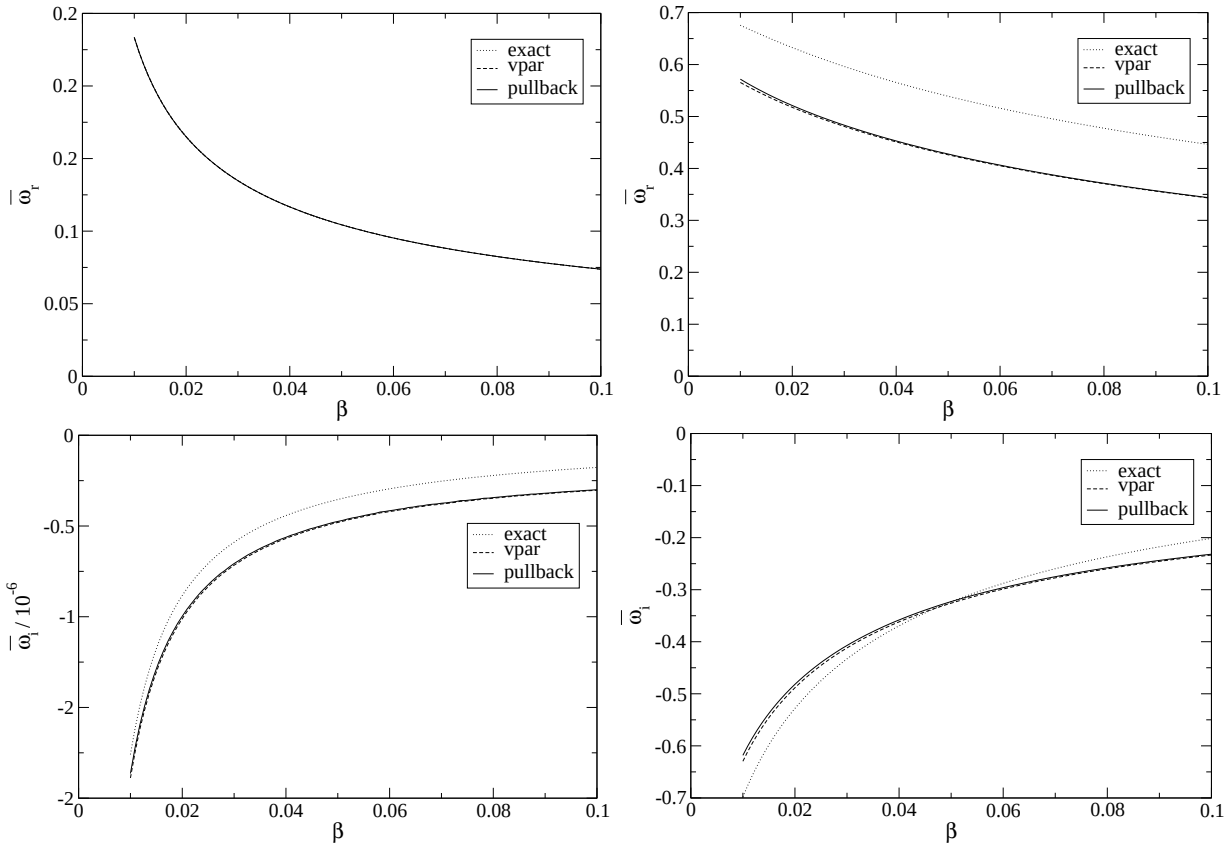

FIG. 2: Kleiber et al.

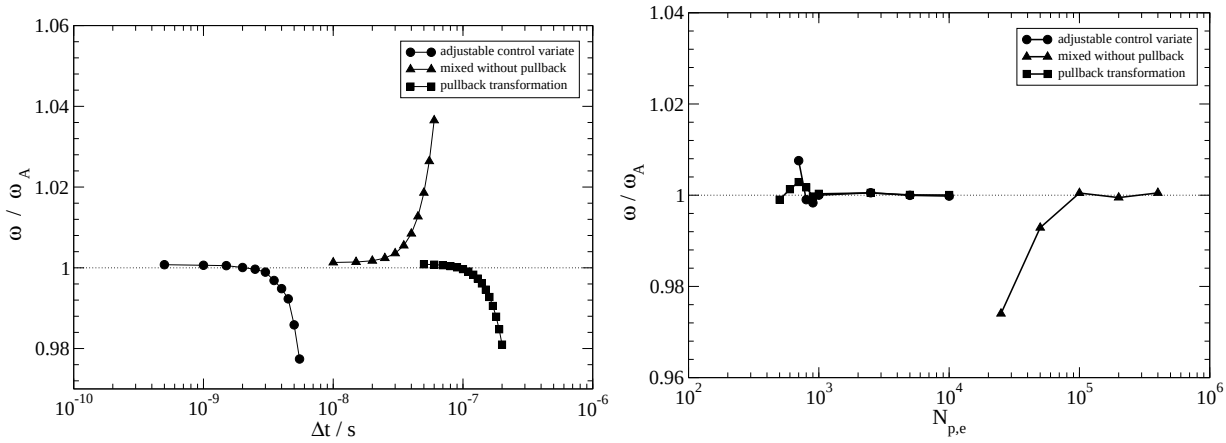

FIG. 3: Kleiber et al.

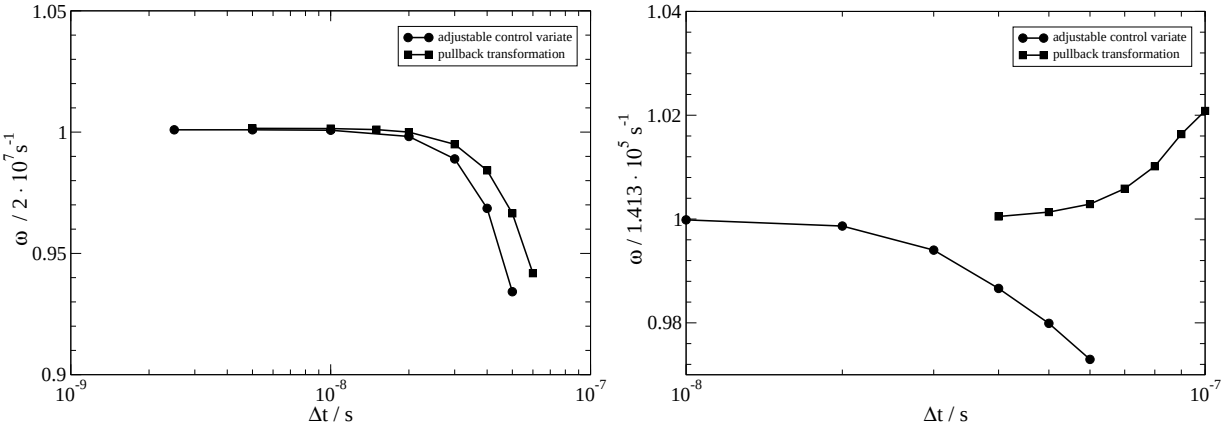

FIG. 4: Kleiber et al.

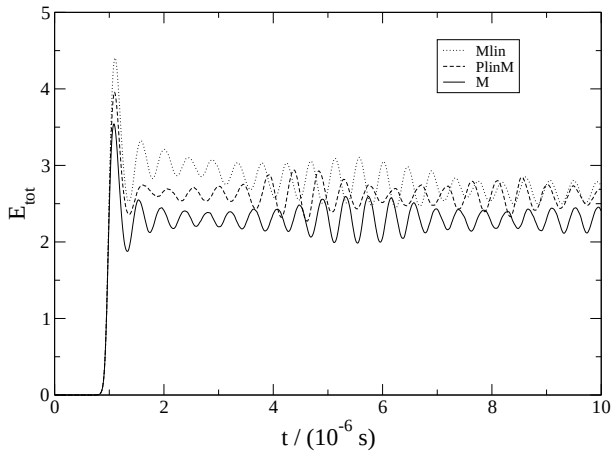

FIG. 5: Kleiber et al.

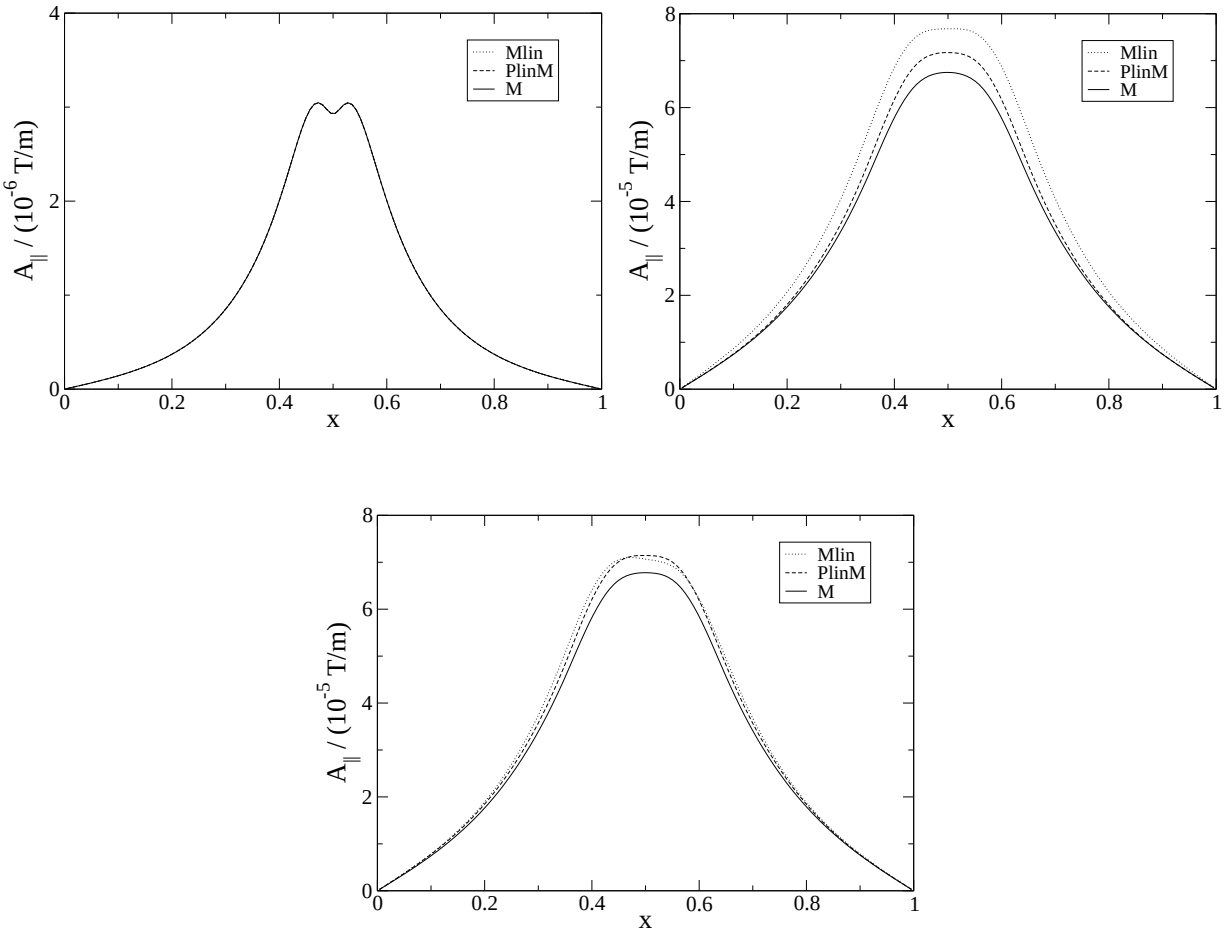

FIG. 6: Kleiber et al.