A 5D gyrokinetic full-f global semi-lagrangian code for flux-driven ion turbulence simulations
Virginie Grandgirard, Jérémie Abiteboul, Julien Bigot, Thomas Cartier-Michaud, Nicolas Crouseilles, Guilhem Dif-Pradalier, Charles Erhlacher, Damien Esteve, Xavier Garbet, Philippe Ghendrih, et al.

To cite this version:

HAL Id: cea-01153011
https://hal-cea.archives-ouvertes.fr/cea-01153011v3
Submitted on 4 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A 5D gyrokinetic full-f global semi-Lagrangian code for flux-driven ion turbulence simulations

V. Grandgirard\textsuperscript{a}, J. Abiteboul\textsuperscript{b}, J. Bigot\textsuperscript{c}, T. Cartier-Michaud\textsuperscript{d}, N. Crouseilles\textsuperscript{e}, G. Dif-Pradalier\textsuperscript{a}, Ch. Ehrlicher\textsuperscript{a}, D. Esteve\textsuperscript{a}, X. Garbet\textsuperscript{a}, Ph. Ghendrih\textsuperscript{a}, G. Latu\textsuperscript{a}, M. Mehrenberger\textsuperscript{d}, C. Norscini\textsuperscript{a}, Ch. Passeron\textsuperscript{a}, F. Rozar\textsuperscript{a}, Y. Sarazin\textsuperscript{a}, E. Sonnendr"ucker\textsuperscript{b}, A. Strugarek\textsuperscript{f}, D. Zarzoso\textsuperscript{g}

\textsuperscript{a}CEA, IRFM, 13108 Saint-Paul-lez-Durance Cedex, France
\textsuperscript{b}Max-Planck-Institut fur Plasmaphysik, Boltzmannstr. 2, 85748 Garching, Germany
\textsuperscript{c}Maison de la Simulation, CEA Saclay, 91191 Gif-sur-Yvette Cedex, France
\textsuperscript{d}IRMA, 67084 Strasbourg Cedex, France
\textsuperscript{e}IRMAR, 35065 RENNES Cedex, France
\textsuperscript{f}GRPS, Physic department, Montreal university, H3C 3J7 Montreal, Canada
\textsuperscript{g}Aix-Marseille Universit´e, PHIM, UMR 7345 Marseille, France

Abstract

This paper addresses non-linear gyrokinetic simulations of ion temperature gradient (ITG) turbulence in tokamak plasmas. The electrostatic GYSELA code is one of the few international 5D gyrokinetic codes able to perform global, full-f and flux-driven simulations. Its has also the numerical originality of being based on a semi-Lagrangian (SL) method. This reference paper for the GYSELA code presents a complete description of its multi-ion species version including: (i) numerical scheme, (ii) high level of parallelism up to 500k cores and (iii) conservation law properties.

Keywords: plasma turbulence, gyrokinetic global full-f flux-driven simulations, semi-Lagrangian method, high-performance computing

1. Introduction

In magnetic fusion devices, the power gain increases non-linearly with the energy confinement time. The quality of the plasma energy confinement then largely determines the size and therefore the cost of a fusion reactor. This confinement time is mainly governed by the plasma turbulence as deserved in such devices – leading to fluctuations with relative magnitude of a few percents in the hot core – and the associated transport. Understanding its origin and properties in view of a possible control is one of the critical issues in fusion science [75]. The inhomogeneities in density, temperature, and magnetic field place the plasma naturally out of thermodynamical equilibrium, and tend to excite several micro-instabilities over a wide spectral range. These plasmas exhibit a low collisionality so that conventional fluid models are questionable and kinetic descriptions are more appropriate. A kinetic formalism is also needed to account for wave-particle interaction. In such first-principle descriptions of plasmas, the six dimensional evolution equation for the distribution function – Vlasov or Fokker-Planck equations – is solved for each species, coupled to the self-consistent equations for the electromagnetic fields, namely Maxwell’s equations. Fortunately, as far as turbulent fluctuations are concerned, they develop at much lower typical frequencies than the high frequency cyclotron motion. Therefore, this 6D problem can be reduced to a 5D one by removing, using phase space reduction, the gyromotion and other high-frequency dynamics. The useful part of the distribution function then evolves in a five dimensional phase space generated by four slow variables and an adiabatic invariant. This model is known as the gyrokinetic model. For detailed gyrokinetic theory see review papers by Brizard

\textsuperscript{a}Email address: virginie.grandgirard@cea.fr (V. Grandgirard)

Preprint submitted to Elsevier March 4, 2016
& Hahm (2007) [16] and Krommes (2012) [79]. But even with this dimensional reduction, building such a simulation tool remains quite challenging.

The first-principle gyrokinetic codes, developed for this purpose, make already an intensive use of massively parallel supercomputers and require state-of-the-art high performance computing (HPC). Nevertheless, solving 5D non-linear gyrokinetic equations for several ion species proves so challenging that to date no code is able to treat all the physics involved. Several strategies based on different simplifications have been developed to decrease these extreme numerical costs. We encourage the reader to refer to the paper by Garbet et al. (2010) [53] for an overview of these different strategies for numerical simulations and their comparisons with fusion experiments. A first strategy is to reduce the simulation domain to a small plasma volume aligned with the magnetic field lines and sufficiently small to neglect the radial variations of macroscopic quantities such as the density and temperature fields and their gradients. Such so called flux-tube codes (e.g. GS2 [44], GENE [76], GYRO [22], GKW [100]) have the advantages to drastically reduce CPU time and memory consumption compared to the global approach where computing over the whole plasma domain requires huge 3D meshes. Despite this constraint, a global approach must be considered whenever the “global” scales are comparable to that of the turbulence. This allows one in particular to address some aspects of the barrier physics. Gyrokinetic models can also be split in two distinct families with respect to the representation of the distribution function: either full-$f$ or $\delta f$ models. In the $\delta f$ model, only perturbations with respect to some prescribed background equilibrium (usually Maxwellian in velocity) are computed. Conversely, in full-$f$ models, the whole distribution function is evolved. Especially, the back reaction of turbulent transport is accounted for in the time evolution of the equilibrium. In global and full-$f$ simulations, the turbulent regime is evanescent if no free energy is injected in the system to prevent the inevitable relaxation of equilibrium profiles below the – linear or non-linear – threshold of the underlying instability. A heat source is mandatory in view of exploring the long time, typically on energy confinement times, behavior of turbulence and transport, leading to the so-called flux-driven simulations.

Finally, the existing codes differ also by their numerical schemes which have evolved all along the last twenty-five years, in direct link with the evolution of HPC resources. Historically, particle in cell methods (PIC) –pioneered for gyrokinetics by Lee [88]– have been most popular, and represent widely adopted approaches to numerical simulations of kinetic plasmas. They used to be considered as the most efficient tool to describe plasma dynamics, essentially because they are capable of describing many physical phenomena in the full dimensional case, at relatively small computational costs. Many gyrokinetic codes are PIC codes. Let us mention, Parker’s code [97], Sydora’s code [117], PG3EQ [42], GTC [90], ELMFIRE [63], GT3D [73], ORB5 [12, 77], GTS [123] and XGC1 [26]. However, it is well known that the relative numerical noise inherent to PIC methods constitutes a strong limiting factor to accurately describe the distribution function in phase space on long time scales. Moreover, the numerical noise only slowly decreases, like $1/\sqrt{N}$, when the number $N$ of particles is increased. The main problem for non-linear gyrokinetic simulations is that the noise level a priori accumulates in time [96]. Even small errors in the evaluation of these moments can cause a systematic corruption of the simulation results on relatively short periods of time. Consequently, the reduction of numerical noise has been right from the start a matter of intense research, and many improvements have taken place during the past ten years, making use of different statistical methods: (i) control variate method of variance reduction (with $\delta f$ scheme [5, 97] or “adjustable control variate” method [60]); (ii) importance sampling (with “optimized loading” [61]). These techniques have enabled to reduce the numerical noise by orders of magnitude. Let us specifically mention the achievements made on the ORB5 gyrokinetic PIC-code [77] on the noise issue, which are summarized in [120]. Another approach to avoid the issue of marker sampling noise is the Eulerian approach. It consists in discretizing the phase space on a fixed grid, and in applying finite differences, finite volumes and/or Fourier transforms to model the differential and integral operators (see e.g. [20] for a review). The main drawback stems from the fact that these numerical schemes are based on explicit time integration, so they are limited by the Courant-Friedrichs-Lewy (CFL) stability condition, which slaves the maximum time step to the grid space resolution. Several gyrokinetic codes are based on this approach as proved by this non-exhaustive list of Eulerian codes: GS2 [78, 44], GENE [76, 93], GYRO [22], GKV [124], GKW [101] and GT5D [70, 74].

As described in the following, one peculiarity of the GYSELA code is to be based on a semi-Lagrangian method, which is a mix between PIC and Eulerian approaches. The objective is to try to take advantages
of both methods, namely limited numerical dissipation with limited numerical noise. We had shown good
properties of energy conservation of semi-Lagrangian scheme for a 4D drift-kinetic model [55]. The Gysela
code exhibits now not only accurate radial force balance [38] but also good properties of local conservation
laws for charge density, energy and toroidal momentum [2]. These conservation properties are detailed
in the following in section 5. Such conservation properties are comparable to those obtained with the
GT5D code [68], which is based on a fourth-order non-dissipative conservative finite difference scheme
[71, 70] which conserves both $L^1$ and $L^2$ norms. Anyway, each of these three numerical approaches –
PIC, Eulerian or semi-Lagrangian– has advantages and drawbacks (see [56] for a review). But for sure,
the wide variety of gyrokinetic codes, coming from all these different choices, is a strength for the fusion
community. Indeed, due to the extremely challenging computing requirements, each gyrokinetic code runs at
the limit of its applicability range. For instance significant advances like taking into account kinetic electrons
and electromagnetic effects are extremely time consuming for global codes. This has been firstly achieved in
the GYRO code [21], then in the GENE code [54] and in the ORB5/NEMORB code\(^1\) [13] and more
recently in the GKY code [67]; this list is not exhaustive. Conversely, the full-$f$ and flux-driven regimes are
two necessary ingredients to investigate, among others, the mechanism underlying self-consistent transport
barrier creation. Benchmarks between the different approaches are therefore primordial (e.g. [118, 48]).
Difficulties reside generally in finding common domains of validity (set of parameters, initial and boundary
conditions,...). The choice of a global full-$f$ approach has been done for the electrostatic non-linear 5D
gyrokinetic code Gysela described in this paper. Electrons are at present considered adiabatic. Within the
community of the 5D gyrokinetic codes, Gysela is close to GT5D code [72] in the sense that they are both
global full-$f$ flux-driven codes. They mainly differ by their numerical schemes and their flux driven choice
where the source term is compensated by a sink term in GT5D [69].

In the following, all the components of the semi-Lagrangian Gysela code will be detailed both in terms
of physical equations and numerical methods as well as the verification and benchmarks that have been
performed. Such flux-driven ITG simulations are extremely challenging and would not be possible without
a high level of parallelism which will also be addressed. The code is currently using Petaflops HPC resources
and is actively preparing its evolution for the future exascale era. The paper is a comprehensive description
of the multi-ion species version of the code with a detailed description of all the numerical schemes employed
and with a precise presentation of the parallelisation of the code. The last same exercise has been performed
in 2006 [55] on the initial 4D drift-kinetic version of the code. The upgrade to a 5D gyrokinetic version
in 2007 [57] and the recent development of a multi-ion version [47] have completely changed the needs in
parallelization. Several important physical results have been obtained with the 5D version of the GYSELA
code in the flux-driven regime, although there was so far no reference paper for it. These results will be
highlighted in the various section of the paper. They deal with: (i) flux driven regime and non-local transport
[105, 37], (ii) neoclassical theory with the implementation of a collision operator [51, 36], (iii) poloidal and
toroidal rotation issues and [39, 106] (iv) the possible control of turbulence by means of sheared flows,
including transport barrier generation and stability [115, 116]. Each of these physical studies have been
made possible thanks to the constant upgrade of the code and critical numerical developments, which are
exhaustively detailed (and benchmarked whenever it is relevant) in the present paper. It is organized as
follows. The physical model is presented in section 2. The numerical methods and the hybrid OpenMP/MPI
parallelism are described in section 3. The code verification via invariance tests and benchmarks are detailed
in section 4. The conservation law properties are analyzed in section 5. Discussion and conclusion close the
paper in section 6.

2. GYSELA gyrokinetic global full-$f$ model

GYSELA is a global code presently used with a simplified concentric circular magnetic configuration
(section 2.2) similar to the Tore Supra equilibria. The new version of the code still considers adiabatic

\(^1\)NEMORB seems to have been the first electromagnetic branch of ORB5. Apparently, the ORB groups have meanwhile
decided to avoid distinction between the different branches. So for the rest of the paper, we will refer to the code with the
unique ORB5 name.
electrons but the possibility to address transport of impurities has been added. The time evolution of the full distribution function of each ion species (major species as e.g. Deuterium + one minor impurity) is governed by a 5D non-linear gyrokinetic Vlasov equation (section 2.3) self-consistently coupled to a 3D Poisson equation (section 2.4). The required gyroaverage operator (section 2.5) that used to be approximated by a Padé expansion in the past versions of Gysela can now be computed with a direct average on gyrocircles. Collisions are taken into account. With a linearized intra-species collision operator (section 2.7) neoclassical effects are recovered. Inter-species collisional transfers are now also considered (section 2.8).

The problem of initialization and radial boundary conditions inherent to global full-f codes like Gysela are addressed in section 2.6. Concerning boundary conditions, three modes are available in the code: (i) the fixed-gradient mode where the temperature profile is fixed at both radial boundaries, corresponding to decaying turbulence regimes (relaxation of equilibrium profiles below instability thresholds cannot be avoided); (ii) the gradient-driven mode, equivalent to the previous one but where gradient profiles are maintained by an artificial Krook-type operator and finally (iii) the flux-driven mode (the most often used) where temperature is still fixed at the outer boundary but can evolve freely at the inner one. In the gradient-driven mode the strength of the drag force of the Krook operator governs the dynamics of the mean (flux surface averaged) gradient profiles: they remain all the more sticked to their initial value if the Krook coefficient \( \nu_k \) is large. More precisely, significant departures of the mean profiles w.r.t. their initial value are only possible on short time scales \( (t \ll \nu_k^{-1}) \), while the long time behaviour ensures that they remain unchanged when time averaged. Concerning the flux-driven mode, the turbulence is forced with a constant-in-time incoming flux generated by a heat source independent of the distribution function (section 2.9) leading to possible long-time simulation. A simulation over several confinement times has been recently performed for comparison to Tore-Supra experiments [41].

### 2.1. Toroidal coordinate system

Let us introduce the notations used in the paper. We consider a set of coordinates labelled \( \{x^i\} \), the metric tensor \( \{g_{ij}\} \) is the product of the transposed Jacobian matrix \( J^T \) and the Jacobian matrix \( J \), i.e \( \{g_{ij}\} = J^T J \). For a set of cartesian coordinates \( X^i \), the elements \( J_{ij} \) of the Jacobian matrix are defined as \( J_{ij} = \partial x^i / \partial X^j \). Let \( g \) represents the determinant of the metric tensor \( \det \{g_{ij}\} \), then the Jacobian in space \( J_\chi \) is defined as \( J_\chi = \sqrt{g} \) and is equal to \( J_\chi = [(\nabla x^i \times \nabla x^j) \cdot \nabla x^3]^{-1} \), i.e. the volume element is \( J_\chi d^3 x \). The tensor \( \{g^{ij}\} \) is the inverse of the tensor \( \{g_{ij}\} \). The element of the contravariant metric tensor verifies the relation \( g^{ij} = \nabla x^i \cdot \nabla x^j \). With these notations, each vector \( A \) can be defined in terms of its covariant components \( A_i \) as \( A = A_i \nabla x^i \) and the equivalent norm is given by \( ||A|| = \sqrt{(A_i)^2 g^{11} + (A_2)^2 g^{22} + (A_3)^2 g^{33}} \). At present, in the code, the coordinate system used is the toroidal one, i.e. the set of coordinates \( (x^1, x^2, x^3) \) is equal to \( (r, \theta, \varphi) \) where \( r \) is the radial position, \( \theta \) is the poloidal angle and \( \varphi \) the toroidal angle. Therefore, \( g^{11} = g^{rr} = 1 \), \( g^{22} = g^{\theta\theta} = 1/r^2 \), \( g^{33} = g^{\varphi\varphi} = 1/R^2 \) and \( g^{ij} = 0 \) for all \( i \neq j \). \( R(r, \theta) = R_0 + r \cos \theta \) with \( R_0 \) the major radius of the torus at the magnetic axis. The Jacobian \( J_\chi \) is then equal to \( J_\chi = r R \).

### 2.2. Simplified magnetic configuration and current

Consistently with the chosen coordinate system, the magnetic topology is fixed and consists of concentric toroidal magnetic surfaces with circular poloidal cross-sections. The magnetic field \( B \) is defined as \( B = (B_0 R_0/R) [\zeta(r) \mathbf{e}_0 + \mathbf{e}_z] \) with \( \zeta(r) = r/(q R_0) \). \( B_0 \) corresponds to the magnetic field on the magnetic axis. The vectors \( \mathbf{e}_\theta = r \nabla \theta \) and \( \mathbf{e}_z = R \nabla \varphi \) are the unit vectors in the poloidal and toroidal periodic directions. The safety factor profile \( q(r) \) is defined by three parameters \( q_1, q_2, \) and \( q_3 \) as \( q(r) = q_1 + q_2 \exp(q_3 \log(r/a)) \). With this choice of angles, \( B \cdot \nabla \varphi / B \cdot \nabla \theta = q(r) R_0 / R \), namely the local field line pitch depends not only on \( r \) but also on \( \theta \). The current is decoupled from the field and the magnetic field is assumed to satisfy the Ampère equation, but not the force balance equation. Then the Ampère equation leads to a current of the form \( \mu_0 \mathbf{J} = \mu_0 J_T R \nabla \varphi \) with \( \mu_0 J_T = \frac{2 \mu_0 R_0 \zeta}{\zeta (1 + \frac{2 \mu_0 R_0 \zeta}{\mu_0 R_0 \zeta} - \frac{\mu_0 R_0 \zeta}{\mu_0 R_0 \zeta})} \).

### 2.3. Full-f gyrokinetic Vlasov equation

Let us consider the gyro-center coordinate system \( (x_G, v_{G||}, \mu_s) \) where \( x_G \) corresponds to 3D space coordinates, \( v_{G||} \) is the velocity parallel to the magnetic field line and \( \mu_s = m_s v_{G\perp}^2 / (2B) \) the magnetic
Let $F_s$ be the particle distribution function of species $s$ and $\bar{F}_s$ the one associated to the guiding-centers. The global gyrokinetic code GYSELA models for each species $s$, the time evolution of the guiding-center distribution function $\bar{F}_s$, with no separation between equilibrium and perturbation. The non-linear time evolution of $\bar{F}_s$ is governed by the 5D collisional gyrokinetic equation described by Brizard and Hahm [16]

$$\partial_t \bar{F}_s - [H, \bar{F}_s]_{GC} = Rhs(\bar{F}_s)$$

(1)

where $H$ the Hamiltonian of the system is defined as $H = \frac{1}{2}m_s v^2_{G||} + \mu_s B + q_s U$ and where $[.,.]_{GC}$ are the gyrokinetic Poisson brackets expressed as (see equation (150) in [16])

$$[X,Y]_{GC} = \frac{B^*_{s||}}{m_s B^*_{s||}} \cdot (\nabla X \partial_{v_G} Y - \partial_{v_G} X \nabla Y) - \frac{b}{q_s B^*_{s||}} \cdot (\nabla X \times \nabla Y)$$

(2)

with $b = B/\|B\|$ the unit vector along the magnetic field line at the guiding-center position. Here, $\bar{U} = \mathcal{F}_\mu U$ is the gyro-average of the fluctuating electrostatic potential $U$. It corresponds to an average over a cyclotron motion: $\mathcal{F}_\mu U = \frac{2\pi}{\mathcal{F}_\mu} U \frac{d\varphi_c}{\pi}$ where $\varphi_c$ stands for the cyclotron phase. This gyro-average operator $\mathcal{F}_\mu$ will be discussed in more detail in section 2.5. The scalar $B^*_{s||}$ corresponds to the volume element in guiding-center velocity space. It is simply $B^*_{s||} = B_s \cdot b$ with $B^*_s$ defined as

$$B^*_s \equiv B + \frac{m_s}{q_s} v_{G||} \nabla \times b$$

(3)

i.e

$$B^*_{s||} \equiv B + \frac{m_s}{q_s} v_{G||} b \cdot (\nabla \times b)$$

(4)

In our case, the right hand side $Rhs$ of the previous Boltzmann equation (1) is given by

$$Rhs(\bar{F}_s) = B^*_{s||} (\mathcal{D}_r(\bar{F}_s) + \mathcal{K}(\bar{F}_s) + \mathcal{C}(\bar{F}_s) + S)$$

where $\mathcal{D}_r$ and $\mathcal{K}$ are respectively a diffusion term and a Krook operator applied on a radial buffer region (see section 2.10), $\mathcal{C}$ corresponds to a collision operator (see section 2.7) and $\mathcal{S}$ refers to source terms which are detailed in section 2.9.

Let us define the Poisson bracket as $[F,G] = b^* \cdot (\nabla F \times \nabla G)$, and the parallel projection as $\nabla^* \equiv \nabla \cdot B/F$. Let $b_\parallel$ represent the covariant components of the unitary magnetic field vector $b$ and $J_\parallel$ the Jacobian in space of the system. Then it can be easily checked that $[F,G] = J_\parallel^{-1} \epsilon^{ijk} \partial_i F \partial_j G \partial_k b$ where $\epsilon^{ijk}$ is the Levi-Civita symbol and $b^*_\parallel \cdot \nabla F = b^*_\parallel \partial_i F$. By using this formalism, it can be proven that the Hamiltonian form (1) is equivalent to the following conservative form

$$B^*_{s||} \frac{\partial \bar{F}_s}{\partial t} + \nabla \cdot \left( B^*_{s||} \frac{d v_{G||}}{dt} \bar{F}_s \right) + \frac{\partial}{\partial v_{G||}} \left( B^*_{s||} \frac{d v_{G||}}{dt} \bar{F}_s \right) = Rhs(\bar{F}_s)$$

(5)

where the evolution of the gyro-center coordinates of species $s$ are described (within the electrostatic limit) by:

$$\frac{d x^i_{G||}}{dt} = v_G b^*_\parallel \cdot \nabla x^i_G + v_{E \times B_r} \cdot \nabla x^i_G + v_{D_r} \cdot \nabla x^i_G$$

(6)

$$m_s \frac{d v_{G||}}{dt} = -\mu_s b^*_\parallel \cdot \nabla B - q_s b^*_\parallel \cdot \nabla U + \frac{m_s v_{G||}}{B} v_{E \times B_r} \cdot \nabla B$$

(7)
where \( b_s^* \) is defined as \(^2\)

\[
b_s^* = \frac{B}{B_{\parallel s}} + \frac{m_s v_{G\parallel}}{q_s B_{\parallel s}^*} \nabla \times B
\]

(8)

The \( i \)-th contravariant components of the ‘\( E \times B \)’ drift are given by

\[
v_{E \times B_i} \cdot \nabla x_G^i = v_{E \times B_i}^i = \frac{1}{B_{\parallel s}} [U, x_G^i]
\]

(9)

and the \( i \)-th contravariant components of the ‘grad–B’ and ‘curvature’ drifts read (at low \( \beta = nT/(B^2/2\mu_0) \) limit)

\[
v_{D_s} \cdot \nabla x_G^i = v_{D_s}^i = \left( \frac{m_s v_{G\parallel}^2 + \mu_s B}{q_s B_{\parallel s}^* B} \right) [B, x_G^i]
\]

(10)

Besides, using the fact that the axi-symmetric equilibrium is determined by three of the ideal MHD equations \( \nabla \times \mathbf{B} = 0 \), \( \nabla \times \mathbf{B} = \mu_0 \mathbf{J} \) and \( \nabla \cdot \mathbf{B} = 0 \), the \( i \)-th contravariant components of \( b_s^* \) appearing in equation (6) read

\[
b_s^* \cdot \nabla x_G^i = b_s^{*i} = \frac{B \cdot \nabla x_G^i}{B_{\parallel s}^*} + \frac{m_s v_{G\parallel}}{q_s B_{\parallel s}^*} \mu_s \mathbf{J} \cdot \nabla x_G^i
\]

(11)

2.4. Self-consistent coupling with the quasi-neutrality equation

The electron density \( n_e \) is supposed to follow an adiabatic (Boltzmann) response on a flux surface, namely \( n_e(x, t) = n_{e_0}(r) \exp \left( c \left[ U(x, t) - \langle U \rangle_{FS}(r, t) \right]/T_e(r) \right) \) with \( T_e \) the electron temperature. \( \langle U \rangle_{FS} \) represents the flux surface average of the electrostatic potential \( U \), i.e.

\[
\langle U \rangle_{FS}(r) = \frac{\int U(r, \theta, \phi) J_\theta \, d\theta \, d\phi}{\int J_\theta \, d\theta \, d\phi}
\]

Let us define, for each species \( s \), its initial radial density profile \( n_{s_0} \) and its concentration \( c_s \equiv n_s/n_{e_0} \).

We also assume that \( n_{e_0} = \sum_s Z_s n_{s_0} \) so that \( \sum_s c_s Z_s = 1 \). Under these assumptions, the quasi-neutrality equation, self-consistently coupled to the gyrokinetic equation (5), reads

\[
- \frac{1}{n_{e_0}} \sum_s Z_s \nabla_{\perp} \cdot \left( \frac{n_{s_0}}{B_0 \Omega_s} \nabla_{\perp} U \right) + e \left( \frac{U - \langle U \rangle_{FS}}{T_e} \right) = \frac{1}{n_{e_0}} \sum_s Z_s (n_{G_s} - n_{G_s,eq})
\]

(12)

with \( \Omega_s = q_s B_0/m_s \) and \( \nabla_{\perp} = (\partial_r, \frac{1}{r} \partial_\theta, \partial_\phi) \). Here, the polarization density (first term of eq. (12)) is approximated by its expression in the limit of large wavelengths with respect to the Larmor radius (limit \( k_{\perp} \mu_s \ll 1 \)). The gyro-center density \( n_{G_s} \) of species \( s \) is defined by \( n_{G_s}(x, t) = \int J_\theta \, dv_{G\parallel} J_{\phi_s} F_s(x, \mathbf{v}, t) \) – where \( J_\theta = (2\pi B_{\parallel s}^*/m_s) \) stands for the Jacobian in the velocity space– similarly \( n_{G_s, eq} \) is the gyro-center density for \( \tilde{F}_s = \bar{F}_{s,eq} \). In practice, the right hand side of (12) the charge density of guiding-centers \( \rho \), is computed as

\[
\rho(x, t) = \frac{1}{n_{e_0}} \sum_s Z_s \int d\mu J_{\mu_s} \left[ \int J_\theta \, dv_{G\parallel} (\bar{F}_s - \tilde{F}_{s,eq}) \right]
\]

(13)

To avoid the problem of the singularity in \( 1/r \), the problem is solved within a ring \( r_{\min} \leq r \leq r_{\max} \) where \( r_{\min} \geq 10^{-5} \). One difficulty with equation (12) is to deal with the flux surface average term \( \langle U \rangle_{FS} \). This term is non-linear in \( \theta \), because the space Jacobian depends on \( \theta \). Therefore, it does not allow one to project simply in 2D Fourier space. To overcome this problem, solving (12) is performed in two steps as proposed

\(^2\)Let us mention that in this definition \( b_s^* \) is not equal to \( B_{\parallel s}^* / B_{\parallel s}^* \) but equal to \( b_s^* = \frac{B_s^*}{B_{\parallel s}^*} = \frac{m_s v_{G\parallel}}{q_s B_{\parallel s}^*} \left( \mathbf{b} \times \nabla B_{\parallel s}^* \right) \).
in [32]. Let us define the differential operator \( \mathcal{L} = -\frac{1}{\rho_s} \sum_{\ell} Z_{\ell} \nabla_\perp \cdot \left( \frac{\rho_s}{B} \nabla_\perp \right) \) and \( \langle \cdot \rangle_{\theta, \varphi} \) the average on \( \theta \) and \( \varphi \) directions. Then, the first step consists in finding \( \mathcal{U} \) solution of the following differential system,

\[
\left( \mathcal{L} + \frac{c}{T_e} \right) \mathcal{U} = \rho - \langle \rho \rangle_{\theta, \varphi} \quad \text{with} \quad \mathcal{U}(r, \theta, \varphi) = \mathcal{U}(r, \theta, \varphi) - \langle \mathcal{U} \rangle_{\theta, \varphi}(r)
\]

The second step consists in solving the following 1D radial differential equation

\[
\mathcal{L}(\mathcal{U})_{FS} = \langle \rho \rangle_{\theta, \varphi} + \left( \mathcal{L} + \frac{c}{T_e} \right) \langle \mathcal{U} \rangle_{FS}
\]

In this equation, \( \varphi \) plays the role of a parameter. A Fourier projection is performed in the \( \theta \) direction. In the radial direction, for both differential systems (14) and (15), finite differences are used. Finally, the electrostatic potential is reconstructed with the formula \( U = U - \langle U \rangle_{FS} - \langle \mathcal{U} \rangle_{FS} \). For both equations (14) and (15), Dirichlet boundary conditions are applied at the outer boundary at \( r_{max} \), while Dirichlet or Neumann can be chosen at the inner boundary at \( r_{min} \). Notice that these boundary conditions then apply to \( \mathcal{U} \), not to the actual electric potential \( U \). Given the relationship between \( U \) and \( \mathcal{U} \), imposing the Dirichlet condition \( U(r_{BC}) = 0 \), \( \{ r_{min}, r_{max} \} \), is then equivalent to the following conditions on \( U \) : \( U(r_{BC}) = U(r_{min}) = (U)_{FS}(r_{BC}) = 0 \). The Neumann condition proceeds in a similar way. In GYSELA, it can be only applied to the inner boundary provided \( r_{min} \leq 10^{-2} \). In this case, \( J_s(r_{min}, \theta) \) is fairly independent of \( \theta \), so that the flux surface average is almost equal to the average over both angles: \( \langle \cdot \rangle_{FS} \approx \langle \cdot \rangle_{\theta, \varphi} \). In the framework of this approximation, imposing the Neumann condition \( \partial_r U(r_{min}) = 0 \) is equivalent to \( \partial_r U(r_{min}) = \partial_r (U)_{\theta, \varphi}(r_{min}) \approx \partial_r (U)_{FS}(r_{min}) = 0 \). The fact that this solution is equivalent to solving directly equation (12) and the detailed numerical scheme are explained in Appendix A.

2.5. Gyro-average operator

The gyro-radius \( \rho_s \) is transverse to \( \mathbf{b} = \mathbf{B}/B \) and depends on the gyrophase angle \( \varphi_c \), i.e., \( \rho_s = (\mathbf{b} \times \mathbf{v})/\Omega_s = \rho_s (\cos \varphi_c \, \mathbf{e}_{\perp 1} + \sin \varphi_c \, \mathbf{e}_{\perp 2}) \). Here \( \mathbf{e}_{\perp 1} \) and \( \mathbf{e}_{\perp 2} \) are the unit vectors of a basis in the plane perpendicular to the magnetic field direction \( \mathbf{b} \). Let \( x_G \) be the guiding-center radial coordinate and \( x \) the position of the particle in real space. These two quantities differ by a Larmor radius \( \rho_s \), as \( x = x_G + \rho_s \). The gyro-average \( \hat{g} \) of any function \( g \) depending on the spatial coordinates corresponds to the following operation

\[
\hat{g}(x_G, \mathbf{v}_\perp) = \frac{1}{2\pi} \int_0^{2\pi} \frac{d\varphi_c}{2\pi} g(x) = \left\{ \int_0^{2\pi} \frac{d\varphi_c}{2\pi} \exp(\mathbf{p}_s \cdot \mathbf{\nabla}) \right\} g(x_G)
\]

The operator \( e^{\mathbf{p}_s \cdot \mathbf{\nabla}} \) corresponds to the change of coordinates \( (x, \mathbf{p}) \to (x_G, \mathbf{p}_G) \). This gyro-average process consists in computing an average on the Larmor circle. It weakens fluctuations that develop at sub-Larmor scales. Introducing \( \hat{g}(k) \) the Fourier transform of \( g \), with \( k \) the wave vector, it is possible to prove that the gyro-average operation reads

\[
\hat{g}(x_G, \mathbf{v}_\perp) = \int_{-\infty}^{+\infty} \frac{d^3k}{(2\pi)^3} J_0(k_{\perp} \rho_s) \hat{g}(k)e^{ikx_G}
\]

where \( k_{\perp} \) is the norm of the transverse component of the wave vector \( k_{\perp} = k - (\mathbf{b} \cdot k)\mathbf{b} \), and \( J_0 \) is the Bessel function \(^3\) of first order. Considering the expression (17), in Fourier space the gyro-average reduces to the multiplication by the Bessel function of argument \( k_{\perp} \rho_s \). This operation is straightforward in simple geometry with periodic boundary conditions, such as in local codes. Conversely, in the case of global codes, the use of Fourier transform is not applicable for two main reasons: (i) radial boundary conditions are non-periodic, and (ii) the radial dependence of the Larmor radius has to be accounted for. Several approaches have been developed to overcome this difficulty.

\(^3\)The Bessel function of the first kind are defined as \( J_n(z) = \frac{\sin(z) - iz^n}{\sin(z) + iz^n} \).
2.5.1. A Padé approximation for the gyro-average operator

The first one, currently used in the code, consists in simplifying the treatment of the gyro-average operator by approximating the Bessel function with a Padé expansion \( J_{\text{Padé}}(k_{\perp} \rho_s) = 1 / \left[ 1 + (k_{\perp} \rho_s)^2 / 4 \right] \) (e.g. see [108]). The advantage of this Padé representation is that it does no longer requires to use the Fourier space as required by the Bessel function. Indeed, since it involves \( k_{\perp}^2 \), it can easily be treated in the configuration space by using the relation \( \nabla^2_{\perp} \leftrightarrow -k_{\perp}^2 \). Using this Laplacian equivalence, the gyro-average operation of any \( g \) function is defined such that each \( m \) Fourier mode of \( \bar{g} \) is solution of the equation

\[
\left[ 1 - \frac{1}{2 \Omega_s^2} m_s \mu_s \left( \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} - \frac{m_s^2}{r^2} \right) \right] \bar{g}^m(r, \varphi) = g^m(r, \varphi)
\]

where at first approximation \( B(r, \theta) \) has been replaced by \( B_0 \) to be consistent with the quasi-neutrality equation. In this differential equation, first and second derivatives are computed using a Taylor expansion of second order leading to a tridiagonal matrix system. This Padé approximation is asymptotically correct in the large wavelength limit \( k_{\perp} \rho_s \ll 1 \) (indeed: \( J_0(k_{\perp} \rho_s) \approx 1 - k^2_{\perp} \rho^2_s / 4 \) for \( k_{\perp} \rho_s \ll 1 \)), while keeping \( J_{\text{Padé}} \) finite in the opposite limit \( k_{\perp} \rho_s \rightarrow \infty \). The drawback is a filtering of small scales: in the limit of large arguments \( x \rightarrow \infty \), \( J_{\text{Padé}}(x) \approx 4 / x^2 \), whereas \( J_0(x) \approx (2 / \pi x)^{1/2} \cos(x - \pi / 4) \).

2.5.2. Integration on gyro-circles using Hermite interpolation

A second widespread method for this gyro-averaging process is to use a quadrature formula. In this context, the integral over the gyro-ring is usually approximated by a sum over four points or more on the gyro-ring [88]. This is rigorously equivalent to considering the Taylor expansion of the Bessel function at order two in the small argument limit, namely \( J_0(k_{\perp} \rho_s) \simeq 1 - (k_{\perp} \rho_s)^2 / 4 \), and equivalent to computing the transverse Laplacian at second order using finite differences. This method has been extended to achieve accuracy for large Larmor radius [61], i.e. the number of points (starting with four) is linearly increased with the gyro-radius to guarantee the same number of points per arclength on the gyro-ring. In this approach –used e.g. in [73] and [77]– the points that are equidistantly distributed over the ring are rotated for each particle (or marker) by a random angle calculated every time step. This is performed on a finite element formalism and enables therefore high order accuracy by keeping the matrixial formulation. In [30] the influence of the interpolation operator (which is of great importance when the quadrature points do not coincide with the grid points) has been studied and it is shown that the cubic splines are appropriate. The direct integration on gyro-circles proposed in [30] has been recently generalized to arbitrary coordinates [113] and implemented in the code. For the distribution function the gyro-average operator is applied on \( \bar{F}_s - F_{s,\text{eq}} \) to deal with values close to 0 at the domain boundaries. Two interpolations –cubic splines and Hermite polynomial– have been tested on analytical cases and basic gyrokinetic simulations with a 4D drift-kinetic model, one Larmor radius and the standard linear Cyclone benchmark case (see [113] for more details and comparison with Padé approximation). Both appear to give the same results. However, the Hermite interpolation is slightly faster and its local character is more favourable for parallelization. The number of points per circle is an input parameter comprised between 4 and 16. There is no adaptive number of points depending on the Larmor radius value because the CPU time is in fact determined by the maximum number. As shown in [113] the method converges with the number of points, so there is no interest in decreasing the number of points at small radius. It is shown in [104] that 16 points is a good compromise between accuracy and CPU time consumption (2 times slower than the previous Padé approximation due to its higher algorithmic complexity). All numerical results presented in the following are performed with the Padé approximation. The reason for not choosing the Hermite interpolation is just chronological: this latter method has been developed only recently. First analyses of new simulations with gyro-average operators based on Hermite show similar results in terms of conservation laws. The impact on non-linear cases will be addressed in a future paper.

2.6. Initial and boundary conditions

Boundary conditions are periodic in \( \theta \) and \( \varphi \) directions. Concerning the radial direction, GYSELA is a global code, i.e. it considers a large fraction of the plasma radius. This is in contrast to flux-tube codes which
focus on a small volume around magnetic field lines by proceeding from a scale separation assumption, the 
fluctuation scale length being smaller than that of the equilibrium. In such codes, periodicity is almost 
always assumed along the radial direction. Conversely, global codes as GYSELA face the delicate problem of 
radial boundary conditions. Non-axisymmetric fluctuations of the electric potential and of the distribution 
function – i.e. $(m, n) \neq (0, 0)$ modes, with $m$ and $n$ the poloidal and toroidal wave numbers – are forced to zero 
at both radial boundaries of the simulated domain. 

As far as the axisymmetric component is concerned, the value of the potential is prescribed at the outer boundary, while the radial electric field is set to zero at the inner boundary. In addition, so as to avoid possible numerical instabilities, which might occur in the case 
where turbulent fluctuations reach the frontiers of the simulation domain, buffer regions have been added at 
both radial boundaries. They are characterized by non-vanishing ad-hoc dissipative coefficients, which aim 
at keeping all gradients finite and damping out all fluctuations (see section 2.10).

Initial conditions consist of an equilibrium distribution function $\bar{F}_{eq}$ perturbed by a sum of accessible 
$(m, n)$ Fourier modes ($m$ and $n$ being the poloidal and toroidal wave numbers, respectively). That means, 
$\bar{F}_s = \bar{F}_{eq} + \delta \bar{F}_s$ where the perturbation part $\delta \bar{F}_s$ reads $\delta \bar{F}_s = \bar{F}_{eq} \cdot g(r) \cdot h(\nu_{G\parallel}) \cdot \delta(p, \varphi)$ with $\delta(p, \varphi) = \sum_{m,n} \epsilon \cos(m\varphi + n\varphi + \delta_{mn})$ where the amplitude $\epsilon$ is fixed and the phases $\delta_{mn}$ have arbitrary values. The radial function $g(r)$ (resp. $h(\nu_{G\parallel})$) has a polynomial dependence and vanishes at both radial (resp. $\nu_{G\parallel}$) 
boundaries. Concerning the initialization of the equilibrium distribution function two choices are possible: 
(i) the first one is a local conventional Maxwellian, (ii) and the second one is a canonical Maxwellian, i.e 
deptends on the motion invariants. The Maxwellian distribution function is defined as

$$\bar{F}_{eq}(r, E) = n_{eq}(r) \times [2\pi T_s(r)/m_s]^{-\frac{3}{2}} \exp \left( - \frac{E}{T_s(r)} \right)$$

(19)

where $E$ stands for the kinetic energy $m_s v_{G\parallel}^2/2 + \mu B(r, \theta)$ which is the second invariant of the system (at 
vanishing electric potential). The initial radial profiles of the ion temperature and density (respectively $T_s(r)$ 
and $n_{eq}(r)$ are deduced by numerical integration of their gradient profiles given by the two parameters $\kappa$ and 
$\Delta r$: $\frac{d \log T_s(r)}{dr} = -\kappa T_s \cdot \cos^{-2}((r - r_p)/\Delta r_{T_s})$ with $r_p$ corresponding to the middle of the radial box. 
$\bar{F}_{eq}$ is constant on a magnetic surface labelled by the radial coordinate $r$. As shown in [38], such initial state 
does not constitute an equilibrium of the system defined by GYSELA at vanishing electric field. A stationary 
equilibrium of the collisionless equations of the code must depend on the three motion invariants, namely 
the adiabatic invariant $\mu$, the total energy $E$ and the toroidal kinetic momentum $P_{\varphi} = q_s \psi + m_s R v_{\varphi}$ with 
$\psi$ the poloidal flux and $v_{\varphi}$ the toroidal fluid velocity. In GYSELA, a convenient choice for this equilibrium is 
provided by the canonical Maxwellian (19) in which the radial coordinate $r$ is replaced – as proposed in [4] – 
by an effective radial coordinate $\bar{r}$, with the dimension of a length, derived from $P_{\varphi}$

$$\bar{r} = r_p - \frac{q_s}{r_p} \left[ \psi(r) - \psi(r_p) \right] - \frac{m_s q_p e B_0}{r_p} \left[ R v_{G\parallel} - R_0 v_{G\parallel} \right]$$

(20)

where $\psi(r) = -B_0 \int_0^r \frac{r'}{q} \, dr'$. The last term $v_{G\parallel}$ in (20) is defined as,

$$v_{G\parallel} = \text{sign}(v_{G\parallel}) \sqrt{2/m_s} \sqrt{E - \mu B_{\max} H(E - \mu B_{\max})}$$

with $H$ the Heaviside function and $B_{\max}$ the maximum of the magnetic field on the whole simulation box. 
It has been chosen to minimize parallel flows. With this expression the difference between $\bar{r}$ and $r$ is of 
order $\rho_s$, the ratio of $\rho_s$ the Larmor radius of species $s$ and the minor radius $a$. In the case of a decaying 
turbulence it is important to choose $\bar{F}_{eq}$ as a function of the motion invariants, especially for studying 
zonal flows. It had been observed in [73, 4] that breaking this rule leads to the development of large scale 
steady flows, which can prevent the onset of turbulence. This phenomena has also been observed in GYSELA 
simulations where a study of the difference between both equilibrium initialization has been performed in 
detail [38, 40]. In [40] it has also been shown that when the system is driven by an external source of free 
energy, the choice of a canonical equilibrium is less crucial. The turbulence onset is only delayed and its 
ultimate nature is unaltered; namely characterized by the same level of parallel and axisymmetric flows, the
the particle mass of species \( s \) in the transport coefficients. In the opposite limit, \( \nu \) plateau orbits are completed by a typical thermal particle. In this so-called banana regime, \( \nu \) frequency \( \nu \) where \( \nu \) is a diffusion term of the order of \( D \sim v_T^2 \nu_{ss'} \) with the collision frequency \( \nu_{ss'} \) being defined as

\[
\nu_{ss'} \sim \frac{D}{v_T^2} \sim \frac{n_{s'}}{m_s} \left( \frac{1}{m_s} + \frac{1}{m_{s'}} \right) \frac{q_i^2 q_{s'}^2}{(1 + \epsilon_0)^2} \left( \frac{v_T^2 + v_{T'}^2}{v_T^2} \right)^{-3/2}
\]

where \( n_{s'} \) is the density of species \( s' \), \( q_s \) (resp. \( q_{s'} \)) is the particle charge of species \( s \) (resp. \( s' \)), \( m_s \) is the particle mass of species \( s \), \( \epsilon_0 \) the permittivity of free space, \( \log \Lambda \approx 17 \) the Coulomb logarithm and \( v_T = \sqrt{T/m_s} \) is the thermal velocity of species \( s \). Let us first underline that the frequency \( \nu_{ss'} \) is different from \( \nu_{s's} \). Besides, let us compare the collision frequencies for the different following particle populations: (i) ion-ion collisions: \( \nu_{ii} \propto Z^4 n_i \sqrt{m_i} \epsilon_i \), (ii) electron-electron collisions: \( \nu_{ee} \propto n_e \sqrt{m_e} T_e^3 \epsilon_e \), (iii) electron-ion collisions: \( \nu_{ei} \approx Z \nu_{ee} \) and (iv) ion-electron collisions: \( \nu_{ie} \approx \frac{m_i}{m_e} Z^2 \nu_{ee} \ll \nu_{ii} \ll \nu_{ee} \). Therefore, the ion-electron collisions can be neglected. At the moment in the code, the electrons are considered adiabatic, therefore only the ion-ion collisions and impurity-ion collisions (see section 2.8.1) are taken into account. The full gyro-averaged and linearized Landau operator has been derived in [126] but such a full Coulomb collisional operator is difficult to implement in GYSELA without severe loss of parallelisation efficiency. As described in section 3.2, the GYSELA parallelization takes advantage of the fact that the magnetic moment \( \mu \) is an adiabatic invariant, that plays the role of a parameter in Boltzmann equation. A unique value of \( \mu \) is assigned to each processor. It was shown that the predictions of the neoclassical theory at low collisionality could be entirely recovered with a reduced collision operator acting in the \( v_i \) direction only [36]. In short, this results from the fact that the main ions of tokamak plasmas are weakly collisional. In this so-called banana regime, collisions essentially perturb the banana orbits at their turning points, where the parallel velocity of trapped particles vanishes. This corresponds to the trapped-passing boundary in the \((v_{\parallel i}, v_{\parallel e})\) plane. Accounting for diffusion in \( v_i \) then reveals sufficient to model such a transport, governed by the broadening of the trapped-passing boundary due to collisions. It also has the advantage of keeping \( \mu \) invariant, hence not degrading the efficiency of the code parallelization. The operator implemented in the code is a simplified version of the Lenard-Bernstein operator [89]. This simplified version has been derived in [51] where it is especially shown that it recovers the exact neoclassical transport in the banana and plateau regimes [4] (see Helander’s book [64] p.149 for complete neoclassical transport regime description). This generic energy and momentum-conserving collision operator has been implemented and successfully tested in the code (see [2], [36]). A new version also valid for the Pfirsch-Schüler regime is under development. The current collision

---

4 The physics of neoclassical transport depends on the collisionality \( \nu_\ast \). If the collisionality is low, \( \nu_\ast < 1 \), the particle orbits are completed by a typical thermal particle. In this so-called banana regime, trapped particles almost fully determine the transport coefficients. In the opposite limit, \( \nu_\ast > e^{-3/2} \), the particle orbit is not fully completed because its motion is disturbed by collisions before. This high collisionality regime is called the Pfirsch-Schüler regime or fluid regime. In-between, the plateau regime is characterized by a weak dependance of the transport on the collisionality of the plasma.
Let us express for each \( \mu \) on \( D \) the collision frequencies where the operators the collision term \( D \) where \( \bar{\nu} \) account:

\[
C_s(F_s) = \frac{1}{B_{s||s}^s} \frac{\partial}{\partial v_{G||}} \left( B_{s||s}^s D_{s||s}^s F_{s||s}^s \frac{\partial}{\partial v_{G||}} \left( \frac{F_s}{F_{M,s}} \right) \right)
\]

(23)

where \( F_{M,s} \) is a shifted Maxwellian distribution—chosen such that \( C_s(F_{M,s}) = 0 \) defined as:

\[
F_{M,s} = \frac{n_{s0}}{(2\pi T_{s,\text{coll}}/m_s)^{3/2}} \exp \left( -m_s^2 \frac{(v_{G||} - V_{s,\text{coll}})^2}{2T_{s,\text{coll}}} - \frac{\mu_s B}{T_{s,\text{coll}}} \right)
\]

(24)

The collision term \( D_{s||s} \) reads for each \( \mu_s \):

\[
D_{s||s}(r, v) = A_s(r) \left( \frac{\Phi(v) - G(v)}{2v} \right) \quad \text{with} \quad A_s(r) = 3\frac{\sqrt{\pi}}{2} \frac{v_{T_s,\text{coll}}^3 e^{3/2}}{qR_0} \nu_{ss}
\]

(25)

where \( v(r, v_{G||}) = \sqrt{E/T_{s,\text{coll}}} \). The scalar \( \nu_{ss} \), associated to the main \( s \) species, is a dimensionless ion- ion collisionality parameter depending on the ion-ion collision frequency \( \nu_{ss} \) such that:

\[
\nu_{ss} = \frac{qR_0}{v_{T_s,\text{coll}} e^{3/2}} \nu_{ss} \quad \text{with} \quad \nu_{ss} = \frac{4\sqrt{\pi}}{3} \frac{n_s e^4 \log \Lambda}{(4\pi\epsilon_0)^2 m_s v_{T_s,\text{coll}}^2}
\]

(26)

where \( v_{T_s,\text{coll}} \) denotes the initial thermal velocity \( v_{T_s,\text{coll}} = (T_{s,\text{coll}}/m_s)^{1/2} \), \( q \) the safety factor, \( R = R_0 + r \cos \theta \) the major radius, \( \epsilon = r/R_0 \) the inverse aspect ratio and \( n_s \) the density of ion species \( s \). The explicit expression (25) also involves the error function \( \Phi \) and the Chandrasekhar function \( G \) defined as:

\[
G(v) = \frac{\Phi(v) - \nu \Phi'(v)}{2v^2} \quad \text{with} \quad \Phi(v) = \frac{2}{\sqrt{\pi}} \int_0^v e^{-x^2} \, dx \quad \text{and} \quad \Phi'(v) = \frac{2}{\sqrt{\pi}} e^{-v^2}
\]

(27)

Considering that the \( \nu_{ss} \) scalar which is given (as input data) in the code corresponds to the main species, the collision frequencies \( \nu_{s's'} \) for minority species \( s' \) are deduced from this value as:

\[
\nu_{s's'} = \frac{v_{T_{s',\text{coll}}} e^{3/2}}{qR_0} \nu_{s's'} \quad \text{with} \quad \nu_{s's'} = \left( \frac{n_{s'}}{n_s} \right) \left( \frac{Z_{s'}}{Z_s} \right)^4 \left( \frac{T_{s,\text{coll}}}{T_{s',\text{coll}}} \right)^2 \nu_{ss}
\]

(28)

Let us express for each \( \mu_s \), the operator \( V_{s||s} \) as:

\[
V_{s||s}(r, v) = -\left( \frac{v_{G||} - V_{s,\text{coll}}}{v_{T_{s,\text{coll}}}^2} \right) D_{s||s}(r, v) \quad \text{with} \quad v_{T_{s,\text{coll}}} = \sqrt{T_{s,\text{coll}}/m_s}
\]

(29)

then the collision operator (23) can be expressed with a more classical Fokker-Planck structure as:

\[
C_s(F_s) = \frac{1}{B_{s||s}^s} \frac{\partial}{\partial v_{G||}} \left( B_{s||s}^s \left( D_{s||s}^s \frac{\partial F_s}{\partial v_{G||}} - V_{s||s} F_s \right) \right)
\]

(30)

where the operators \( D_{s||s} \) and \( V_{s||s} \) respectively model a diffusion and a drag in the parallel velocity direction. The conservation properties of parallel momentum and energy are ensured by constraining \( D_{s||s} \) to depend on \( \mu_s \) only and defining the local fluid velocity \( V_{s,\text{coll}} \) and ion temperature \( T_{s,\text{coll}} \) as follows (see Appendix
AppendixB for more details)

\[ V_{||s,\text{coll}} = P^{-1} \left[ \left( \frac{m_s}{B_{||s}} \partial_{v_{G||}} \left( B^*_{||s} D_{||s} v_{G||} \right) \right) \left( m_s D_{||s} v_{G||} \right) \right. \]

\[ - \left. \left( \frac{1}{B_{||s}} \partial_{v_{G||}} \left( B^*_{||s} D_{||s} \right) \right) \left( m_s^2 D_{||s}^2 v_{G||} \right) \right] \]

\[ T_{s,\text{coll}} = P^{-1} \left[ \left( m_s D_{||s} \right) \left( m_s^2 D_{||s}^2 v_{G||} \right) - \left( m_s D_{||s} v_{G||} \right) \left( m_s^2 D_{||s} v_{G||} \right) \right] \]

\[ P = \left( m_s D_{||s} \right) \left( m_s^2 D_{||s}^2 v_{G||} \right) - \left( m_s D_{||s} v_{G||} \right) \left( m_s^2 D_{||s} v_{G||} \right) \]

(31)

with the brackets \( \langle \cdot \rangle \) corresponding to the velocity space integral \( \langle \cdot \rangle = \int \cdot \, d\mu_{vG} F_s \). The impact of the collisions on the evolution of the distribution function \( F_s \) is taken into account by stepping the evolution of the distribution function \( \partial_t F_s = C_s(F_s) \) with a Crank-Nicolson scheme. This collision operator forces the system to relax towards the Maxwellian distribution function, calculated from the instantaneous and local parallel flow \( V_{||s,\text{coll}} \) and the isotropic temperature \( T_{s,\text{coll}} \sim T_s \).

2.8. Collisional transfer between two species

The inter-species operator currently implemented in the code is highly simplified. It only ensures the moment transfers and energy transfers between species. A more complete version, satisfying neoclassical results for impurity transport, is under development.

2.8.1. Collisional energy transfer

The energy exchange between two species is approached by the following reduced collision operator

\[ \frac{dF_s}{dt} = -\frac{\eta_{ss'}^{\varepsilon}}{(2\pi T_{\text{mean}}/m_s)^{3/2}} \frac{\Delta T_{ss'}}{T_{\text{mean}}} \left( \frac{E_{ss'} - \frac{3}{2} \exp \left( -\frac{E_{ss'}}{T_{\text{mean}}} \right) \equiv C_{ss'}(E_{ss'}) \right) \]

\[ \frac{dF_{s'}}{dt} = -\frac{\eta_{ss'}^{\varepsilon}}{(2\pi T_{\text{mean}}/m_{s'})^{3/2}} \frac{\Delta T_{s's'}}{T_{\text{mean}}} \left( \frac{E_{s's'} - \frac{3}{2} \exp \left( -\frac{E_{s's'}}{T_{\text{mean}}} \right) \equiv C_{s's'}(E_{s's'}) \right) \]

(34)

(35)

where \( d/dt \) stands for the phase space Lagrangian derivative and the following definitions have been adopted

\[ T_{\text{mean}} = T_s + T_{s'} \]

\[ V_{\text{mean}} = \frac{V_{||s} + V_{||s'}}{2} \]

\[ \Delta T_{ss'} = T_s - T_{s'} = -\Delta T_{s's} \]

\[ \Delta V_{ss'} = V_{||s} - V_{||s'} \]

\[ E_{ss'} = m_s (v_{G||} - V_{\text{mean}})^2 + \mu_s B \]

\[ E_{s's'} = m_{s'} (v_{G||} - V_{\text{mean}})^2 + \mu_{s'} B \]

(36)

(37)

(38)

The temperatures and fluid velocities which enter these definitions are flux surface averaged, i.e. by considering \( \langle \cdot \rangle_{FS} = \int \cdot J_s \, dv \, d\varphi / \int J_s \, dv \, d\varphi \). Then velocities correspond to \( V_{||s}(r) = \langle \int v_{G||} \, F_s \, d^3v \rangle_{FS} / N_s(r) \) with density \( N_s(r) = \langle \int F_s \, d^3v \rangle_{FS} \), the volume element being defined as \( d^3v = J_s \, dv_{G||} \, d\mu_s \). Temperatures are defined as \( T_s(r) = 2/3 \langle \int \varepsilon_s F_s \, d^3v \rangle_{FS} / N_s(r) \) where \( \varepsilon_s = \left( m_s/2 \right) \left( v_{G||} - V_{||s} \right)^2 + \mu_s B \). The parameter \( \eta_{ss'}^{\varepsilon} \) has been designed so that particles and parallel momentum are conserved

\[ \eta_{ss'}^{\varepsilon} = \frac{8}{3\sqrt{\pi}} \frac{m_s}{m_{s'}} \left( \frac{Z_{s'}}{Z_s} \right)^2 \frac{n_{s'}}{\nu_{s'}} \left( \frac{T_{\text{mean}}}{1 + \frac{\nu_{s'}^2}{\nu_s^2}} \right)^{-3/2} \]

(39)
where the normalized ion-ion collision frequency \( \nu_{s,s} \) is given by eq.(26) (for detailed calculation see Appendix AppendixC.1). Numerically, equation (34), is solved as \( F_s(t + \Delta t) = F_s(t) + \Delta t \nabla_{s,s}(E_{ss'} , t) \).

2.8.2. Parallel momentum transfer

Parallel momentum exchange between two species can be modelled by the following approximate collision operator:

\[
\frac{dF_s}{dt} = -\eta_{ss'} v_{ss'} \Delta V_{ss'} \exp \left(-\frac{E_s}{T_{\text{mean}}}\right) \equiv C_{ss'}^V (E_s) \quad \text{with} \quad E_s = \frac{1}{2} m_s v_{ss'}^2 + \mu_s B
\]

\[
\frac{dF_s'}{dt} = -\eta_{ss'} v_{ss'} \Delta V_{ss'} \exp \left(-\frac{E_s'}{T_{\text{mean}}}\right) \equiv C_{ss'}^V (E_s')
\]

where \( T_{\text{mean}} \) is defined by equation (36). \( \Delta V_{ss'} \) is equivalent to eq.(37). The parameters \( \eta_{ss'}^V \) and \( \eta_{ss'}^G \) are designed such that only parallel momentum exchanges are induced by this operator, thus leading to

\[
\eta_{ss'}^V = n_s m_s^2 \left( \frac{\varepsilon}{2\pi} \right)^{3/2} \frac{1}{q_R^3} \sqrt{T_{s,\text{coll}}} \left( \frac{Z_s'}{Z_s} \right)^2 \frac{v_{T_s}}{v_{T_s'}} \nu_{ss'} \quad \text{and} \quad \eta_{ss'}^G = \left( \frac{m_s}{m_s'} \right)^{3/2} \eta_{ss'}^V
\]

where the velocity \( v_{T_s} \) corresponds to the maximum value between \( v_{T_s} \) and \( v_{T_s'} \) (for more details see Appendix AppendixC.2).

2.9. Flux-driven code with source terms

2.9.1. Sources of heating, momentum and vorticity

As introduced previously GYSELA is a full-f code, namely the back reaction of turbulent transport is accounted for in the parallel momentum evolution of the equilibrium. In such a framework, the turbulence regime is evanescent if no free energy is injected in the system. Turbulent transport results in the flattening of the temperature profile, which would ultimately reach marginal stability in the absence of any forcing. A heat source is mandatory in view of exploring the long time–on energy confinement times– behaviour of turbulence and transport. In GYSELA, the possibility to drive the system by a prescribed source was added in 2009 [105]. This source consists of the sum of the product of Hermite and Laguerre polynomials in \( v_{ss}^G \) and \( \mu_s \), respectively, in the spirit of the pioneering work by Darmet et al [34]. It is versatile enough to allow for separate injection of heat, parallel momentum and vorticity. Such a versatility imposes serious constraints on the expression of the source in phase space. The separation between these three kinds of sources is achieved using projections onto the bases of orthogonal Hermite polynomials in \( v_{ss}^G \) and orthogonal Laguerre polynomials in \( \mu_s \). The retained expression for the source is the following (see Appendix AppendixD for detailed description)

\[
\frac{dF_s}{dt} = S_{\text{heat}}(r, \theta, v_{G||}, \mu_s) + S_{\text{moment}}(r, \theta, v_{G||}, \mu_s) + S_{\text{vorticity}}(r, \theta, v_{G||}, \mu_s)
\]

where the pure heating source is defined as:

\[
S_{\text{heat}} = \left[ v_{G||}^3 - \frac{1}{2} \left( J_{||} - \frac{J_{||} B}{2 - J_{||} B} (2 - \bar{\mu}_s) (2v_{G||} - J_{||} B) \right) \right] \frac{2S_0^\varepsilon}{(2\pi T_{s,\text{strce}}/m_s)^{3/2}} \frac{1}{T_{s,\text{strce}}^2} S_r^\varepsilon e^{-v_{G||}^2 - \bar{\mu}_s}
\]

with \( \bar{\mu}_s = \mu_s B / T_{s,strce}, v_{G||} = v_{G||} / \sqrt{2T_{s,strce}/m_s}, J_{||} = \sqrt{2m_s T_{s,strce}/(q_s B^2)} \mu_0 J_{||} \) and \( \alpha = (m_s/q_s^2) \times T_{s,\text{strce}}/(2B^2) \). \( S_{\text{moment}} \) is a pure momentum source expressed as

\[
S_{\text{moment}} = \left[ 2v_{G||} (2 - \bar{\mu}_s) - J_{||} B (1 + 2v_{G||} - \bar{\mu}_s) \right] \frac{S_0^\varepsilon}{4\pi^{3/2}} \frac{1}{T_{s,\text{strce}}^2} S_r^\varepsilon e^{-v_{G||}^2 - \bar{\mu}_s}
\]
while $S_{\text{vorticity}}$ is a pure source of vorticity

$$S_{\text{vorticity}} = -\left[2\tilde{v}_s^2 - \tilde{\mu}_s\right] \frac{S^0}{\alpha (2\pi T_{s,\text{src}}/m_s)^{3/2}} S^0 e^{-\tilde{v}_s^2 - \tilde{\mu}_s} \quad (46)$$

The prescribed radial envelopes $S^E_r$, $S^{\nu G}_r$ and $S^\Omega_r$ are chosen as the sum of two hyperbolic tangents but could be any radial function under the constraint that the integral over the minor radius is normalized to 1 while $S^E_0$, $S^{\nu G}_0$ and $S^\Omega_0$ correspond to the source amplitudes and $T_{s,\text{src}}$ to the temperature of the source. They are usually localized close to the inner boundary of the simulation domain [105]. The heat source is an isotropic source that takes particles at a vanishing velocity and accelerates them up to $v \approx 1.5 v_{Te}$. See schematic view with figure 2. As a remark, there is another choice possible for the heat source $S_{\text{heat}}$ in the code, which corresponds to

$$S_{\text{heat}} = \frac{S^E_0}{3\sqrt{2} \left( \frac{\pi T_{s,\text{src}}}{m_s} \right)^{3/2} T_{s,\text{src}} \exp \left(-\frac{E}{T_{s,\text{src}}}ight)} S^E_r \quad (47)$$

There are two main advantages in dealing with a prescribed heat source: (i) the forcing of turbulence can mimic that in experiments, in contrast to simulations where the mean gradient is prescribed and (ii) the sum of the spatially and time (on the $\tau_E$ time scale) averaged turbulent and neoclassical heat fluxes must balance the prescribed driving flux. In this case, the response is the temperature gradient, which ultimately governs the internal energy and therefore the performance of the discharge. An example of initial and final temperature profiles is shown in figure 1. Flux driven simulations then allow investigating the impact of heating power on the energy confinement time [106]. The source of vorticity described before has been efficiently used to polarize the plasma [115] inducing the development of sheared electric fields in the turbulent core. The creation of ion transport barriers by these externally induced sheared $E \times B$ flows has been studied in details in [116, 115].

![Figure 1: Schematic view of a heat source and buffer regions in the case of a flux driven simulation. Comparison between initial temperature profile (black line) and final temperature (red line).](image)

2.9.2. Energetic particle source

A source of energetic particles (EPs) has been also implemented in the code to study the interaction between EPs and turbulence. EPs are characterized by energies larger than the thermal energy. The excitation by EPs of the geodesic acoustic modes (GAMs) –corresponding to the oscillatory component of large scale $E \times B$ zonal flows– creates modes which are called energetic geodesic acoustic modes (EGAMs) [50, 95]. For more details on the impact of EGAMs on turbulence see [127, 128, 45]. In practice this source
is coupled to the heating source, such that \( d_t \bar{F}_s = S_{\text{heat}} + S_{\text{EP}} \) where \( S_{\text{EP}} \) is also built, as for the previous sources, by using projection onto the Laguerre and Hermite polynomial bases, with the constraint to inject only parallel energy. For symmetry reason, the energetic particle source, is built as

\[
S_{\text{EP}}(r, \theta, v_{G\parallel}, \mu_s, t) = S_0^{\text{EP}}(t) S^\parallel_{\text{EP}}(r) \quad \text{with} \quad S_\pm(\theta, v_{G\parallel}, \mu_s) = S_\pm(\theta, -v_{G\parallel}, \mu_s)
\]

where \( S_0^{\text{EP}} \) is the source amplitude and \( S^\parallel_{\text{EP}} \) the radial profile is normalized such that \( \int r \, dr S^\parallel_{\text{EP}} = 1 \). After the same kind of calculation as in the previous sources, the final expression reads

\[
S_\pm = \left[ (v_{G\parallel} \pm \bar{v}_0)^2 - \frac{1}{2} - Q_{\text{EP}}(2 - \mu_s)(2(v_{G\parallel} \pm \bar{v}_0) - J_{\parallel B}) \right] e^{-(v_{G\parallel} \pm \bar{v}_0)^2} e^{-\mu_s} \tag{48}
\]

where \( \mu_s = \mu_s B(r, \theta) / T_{G\parallel} \) and \( Q_{\text{EP}} = J_{\parallel B} / \left( 2 - J_{\parallel B}^2 (1 + 2\bar{v}_0^2) \right) \) with \( \bar{v}_0 = v_0 / \sqrt{2T_{G\parallel}} \) an arbitrary normalized velocity. The expressions of the parallel current \( J_{\parallel B} \) and the velocity \( v_{G\parallel} \) are the same as in equation (44) where \( T_{s,\text{srce}} \) is replaced by \( T_{G\parallel} \). \( T_{G\parallel} \) and \( T_{G\perp} \) correspond to the normalized parallel and perpendicular temperatures of the energetic particle source. Each of the terms \( S_+ \) and \( S_- \) does not inject neither mass –which is essential because electrons are considered adiabatic in the code-- nor vorticity. See figure 2 for a representative view of the source in parallel velocity direction. The source mimics the effects of two tangential neutral beam injectors, oriented in the co-and counter-current directions. It is localized around the mid position \( r = 0.5 (r_{\text{min}} + r_{\text{max}}) \) and brings the distribution function out of the equilibrium by creating a positive slope in energy. As explained in [128] \( v_0 \) and \( T_{G\parallel} \) are both critical parameters in view of exciting EGAMs. GYSELA results for EGAMs excitation have been successfully compared to analytical theory [128] and benchmarked [9] more recently with ORB5 code.

2.9.3. Source of impurities

Finally, it is also possible to add a source of impurities \( s' \) of the form

\[
S_{n,s'} = \frac{S_0^{n,s'} S_e}{(2\pi T_{s',\text{srce}} m_{s'})^{3/2}} \left( \frac{5}{2} - \frac{\mu_{s'} B_{s'}}{T_{s',\text{srce}}} - \frac{m_{s'} v_{G\parallel}^2}{2T_{s',\text{srce}}} \right) \exp \left( -\frac{m_{s'} v_{G\parallel}^2}{2T_{s',\text{srce}}} - \frac{\mu_{s'} B_{s'}}{T_{s',\text{srce}}} \right) \tag{49}
\]

Let us notice that this source of matter is not a pure source, due to the fact that it also injects some amount of vorticity. This injection is balanced by a sink for the major species \( s \), such that

\[
Z_s \int S_{n,s} J_{v_{G\parallel}} \, dv_{G\parallel} \, d\mu_s + Z_{s'} \int J_{v_{G\parallel}} S_{n,s'} \, dv_{G\parallel} \, d\mu_{s'} = 0
\]
2.10. Diffusion terms in the buffer regions

Finally, to avoid strong gradients at the boundaries, radial diffusion and artificial damping can be added in buffer regions. These buffer regions are defined at each side of the radial domain \( r \in [r_{\min}, r_{\max}] \) as a sum of hyperbolic tangents:

\[
H_{\text{buff}}(r) = 1 + \frac{1}{2} \left[ \tanh \left( \frac{r - r_{\max} + B_L L_r}{B_S L_r} \right) - \tanh \left( \frac{r - r_{\min} - B_L L_r}{B_S L_r} \right) \right]
\]

where \( L_r \) is the length of the radial domain, \( B_L \) and \( B_S \) are respectively the location and the stiffness of the buffer regions. The function \( H_{\text{buff}} \) plays the role of a mask which is equal to 1 in the buffer regions and 0 elsewhere. The diffusion term which is applied in the buffer regions is of the form

\[
\mathcal{D}_r(F_s) = \frac{1}{B_{||s}^*} \left[ \frac{1}{r} \frac{\partial}{\partial r} \left( r \chi(r) B_{||s}^* \frac{\partial}{\partial r} F_s \right) \right]
\]

with \( \chi(r) = \chi_0 H_{\text{buff}}(r) \) (\( \chi_0 \) being the diffusion coefficient). The equation \( \partial_t \dot{F}_s = \mathcal{D}_r(F_s) \) is solved by using a Crank-Nicolson scheme (see Appendix AppendixE). An artificial damping term \( \nu_0 \) is introduced in the buffer regions by defining a Krook operator \( \mathcal{K}_r(F_s) = -\nu(r)(\dot{F}_s - \dot{F}_{s,eq}) \) with \( \nu(r) = \nu_0 H_{\text{buff}}(r) \) and solving \( \partial_t \dot{F}_s = \mathcal{K}_r(F_s) \). Let \( \Delta t \) be the time step, then an analytic solution of the previous equation is given by

\[
\dot{F}_s(t + \Delta t) = \dot{F}_{s,eq}(t) + (\dot{F}_s(t) - \dot{F}_{s,eq}(t)) \exp(-\nu(r)\Delta t).
\]

This mechanism restores the distribution function to its initial equilibrium state \( \dot{F}_{s,eq} \), in the buffer regions, by slowly damping all the turbulent modes of the system. It also plays the role of a heat sink by effectively coupling the plasma with the outer thermal baths.

3. A highly parallel semi-Lagrangian code

Solving the set of gyrokinetic equations (5)-(12) is very challenging. It consists of (i) one evolution equation (5) of the distribution function for each ion species (so far, we are limited to 3 distribution functions in GYSELA. Considering that kinetic electrons will soon become operational, this leaves 2 slots for ions: either Deuterium and Tritium, or a main ion species plus any kind of impurity, either intrinsic like Helium or extrinsic, such as Tungsten for instance) in the 4D phase space parametrized by the adiabatic invariant \( \mu \); (ii) the set of four coupled ordinary differential equations (ODE) for the trajectories (6)-(7); and (iii) 3D integro-differential equations for the field, namely the quasi-neutrality equation (12). This set of equations is nonlinear, the dominant quadratic nonlinearity being due to the \( E \times B \) advection term. The quasi-neutrality equation is generally solved in most of the gyrokinetic codes by using Fourier projection in all the periodic directions and finite differences or finite elements in the others. Concerning the global algorithm, the challenge consists in finding numerical schemes which preserve the first principles such as the conservation of Casimir invariants, the phase space volume and the total energy. Various numerical schemes, classifiable as PIC, Eulerian or semi-Lagrangian, have been explored until now. In the following, we focus on the semi-Lagrangian approach which is specific of the GYSELA code.

3.1. Specificity of the GYSELA code: the semi-Lagrangian scheme

Semi-Lagrangian (SL) schemes have been first used for the advection of vorticity in simplified models of large scale flows. It has gained maturity when the discretization approach was introduced in the relevant context of atmospheric flows. A comprehensive review of semi-Lagrangian methods in this meteorological context until 1990 is due to Staniforth [112]. It is also applied to geophysical fluid dynamics (cf. [46]). In magnetized plasma turbulence area, the SL method has been first applied to calculate a turbulence driven by passing ions in 2D (1D in space, 1D in velocity) [111] and trapped ions in 3D (2D in space, 1D in velocity) [35, 108]. This method was then extended to the 4D model (3D in space and \( v_z \) (with \( \mu = 0 \)) of Ion Temperature Gradient (ITG) driven turbulence in cylindrical geometry with the development of the GYSELA code (for GYrokinetic SEmi-Lagrangian code) [55] and the CYGNE code [19]. The 4D drift-kinetic slab-ITG version of the GYSELA code has shown good properties of energy conservation in non-linear regime [55] as well as accurate description of fine spatial scales [107]. In the CYGNE code the standard Taylor
expansion is replaced by a B"urlisch-Stoer scheme (for the 2D advection) to increase the spatial accuracy and the logarithmic interpolation technique is used to ensure the positivity of the distribution function. Brunetti et al. [19] have shown that (i) the positivity can be preserved but at the cost of larger diffusion and (ii) that non-equidistant meshes in radial and parallel velocity directions are a key tool for obtaining accurate results. Due to the good conservation property obtained in 4D, the gyrokinetic 5D version of the GYSELA code [57] has been based on the same numerical scheme. The global algorithm for the new 5D gyrokinetic multi-ion species version of the code will be described in the following both in terms of numerics (section 3.1.3) and parallelisation (section 3.2). Advantages and drawbacks of such a 5D semi-Lagrangian code will be discussed. The purpose of the semi-Lagrangian method is to take advantage of both the Lagrangian and Eulerian approaches, with an accurate description of the phase space, in particular regions where the density is low, as well as an enhanced numerical stability. It is based on the fact that the most precise way to solve convection (or advection) hyperbolic PDE is to use their characteristics along which the distribution function remains constant. This method was primarily developed by Cheng and Knorr [27]. It has been cast in more general framework of SL by Sonnendr"ucker in 1998 [111] and Nakamura in 1999 [94]. In this approach, the phase-space mesh grid is kept fixed in time (Eulerian method) and the Vlasov equation is integrated along the trajectories (Lagrangian method) using the invariance of the distribution function along the trajectories. The GYSELA code is based on this standard semi-Lagrangian approach [111]. This approach has been recently renamed Backward semi-Lagrangian approach (BSL) by its author to make the distinction with the emergence of new schemes: (i) Forward semi-Lagrangian approach (FSL) firstly proposed in 2009 [33] and (ii) Conservative semi-Lagrangian approach (CSL) proposed in [31, 15] through Parabolic Spline Method [129]. The main difference between BSL and FSL approaches is that the advection equations are solved backward in time in the first case and forward in the second one. They are both based on solving the advective form of Vlasov equation while the CSL methods deal with the conservative form of the Vlasov equation. FSL and CSL schemes have both been tested in GYSELA [84, 14] but the actual version of the code is the standard BSL approach as described in the following.

3.1.1. Backward Semi-Lagrangian (BSL) concept

Let us consider the 5D gyrokinetic Vlasov equation

\[ B_{\parallel s} \frac{\partial F_s}{\partial t} + \nabla \cdot \left( B_{S,\parallel s} \frac{dx_G}{dt} F_s \right) + \frac{\partial}{\partial v_G} \left( B_{\parallel s} \frac{dv_G}{dt} F_s \right) = 0 \]  

which corresponds to equation (5) without collisions and source terms. By using the incompressibility property of the gyrocenter flow in the 4D phase space, i.e. \( \nabla \cdot \left( B_{\parallel s} d_1 x_G \right) + \partial_{v_G} \left( B_{\parallel s} d_1 v_G \right) = 0 \), the previous conservative Vlasov equation (52) is equivalent to its advective form

\[ \frac{\partial F_s}{\partial t} + \frac{dx_G}{dt} \nabla \cdot F_s + \frac{dv_G}{dt} \partial_{v_G} F_s = 0 \]  

Let \( \Gamma = (x_G, v_G, \mu_s) \) be a position vector in the phase space solution of the characteristic equations (6)-(7). Then equation (53) leads to \( d_1 F_s(\Gamma(t), t) = 0 \). The semi-Lagrangian method uses this invariance of the distribution function \( F_s \) along its characteristics. Let us consider the computational 5D domain \( [x_{G,\text{min}}^1, x_{G,\text{max}}^1] \times [x_{G,\text{min}}^2, x_{G,\text{max}}^2] \times [x_{G,\text{min}}^3, x_{G,\text{max}}^3] \times [v_{G,\text{min}}^l, v_{G,\text{max}}^l] \times [\mu_{s,\text{min}}, \mu_{s,\text{max}}^m] \) and the associated grid (fixed in time) defined by the finite set of mesh points \( \Gamma_{ijklm} = \left( x_{G,1}^1, x_{G,2}^1, x_{G,3}^1, v_{G,l}, \mu_{s,m} \right) \) with \( x_{G,q}^p = q \Delta x_{G}^p \). For all \( q = 0, \ldots, N_p \) with \( p = 1, 2, 3 \); \( v_{G,\parallel l} = l \Delta v_{G}^{\parallel} \) for all \( l = 0, N_v \) and \( \mu_{s,m} \) for all \( m = 0, N_{\mu} \). \( N_1, N_2, N_3 \) are the number of cells in each spatial directions and \( N_v \) the number of cells in \( v_{G,\parallel} \) direction. The \( (N_{\mu} + 1) \) values for \( \mu_s \) are not necessary equidistributed. Actually, the standard choice in GYSELA corresponds to an equidistributed grid in \( \sqrt{\mu} \). This choice leads to a better accuracy when computing integrals in \( \mu \). Let us also assume that \( F_s \) is known at each point \( \Gamma_{ijklm} \) of this grid at time \( t_n \). Therefore the distribution function can be computed at the next time \( t_n + \Delta t \), with \( \Delta t \) the time step, on the same grid by using the invariance property

\[ F_s(\Gamma_{ijklm}(t_n + \Delta t), t_n + \Delta t) = F_s(\Gamma_{ijklm}, t_n + \Delta t, t_n) \]
\( \Gamma(t_n, \Gamma_{ijklm}, t_n + \Delta t) \) corresponds to the solution of the characteristic at time step \( t_n \) which is equal to \( \Gamma_{ijklm} \) at time \( t_n + \Delta t \). The method consists in (i) finding the foot of the characteristic at the time \( t_n \): \( \Gamma(t_n, \Gamma_{ijklm}, t_n + \Delta t) \) by solving backward in time the advection equations (6)-(7) and (ii) computing \( F_s(\Gamma(t_n, \Gamma_{ijklm}, t_n + \Delta t), t_n) \) by interpolation, using the fact that at this time \( t_n \) the distribution function is known over the whole fixed grid. Cubic spline interpolations are used in the code, because it offers a good compromise between accuracy (small diffusivity) and simplicity (numerical cost) [49, 7].

3.1.2. Time-splitting

In low-dimensional systems, the semi-Lagrangian method is very efficient. When applied to higher dimensional problems, one faces the problem of multidimensional interpolation, which is extremely expensive for high dimensional problems. However, this problem has been partially cured by using the time-splitting idea of Cheng and Knorr [27]. Using the incompressibility property, Strang's operator decomposition into space and velocity can be applied, replacing equation (5) by a set of two conservative equations. Besides, to avoid dealing with a 3D space operator, the latter is also divided into two parts. Let us denote \( \mathcal{X}_G = (x_G^1, x_G^2) = (r, \theta) \) and remind that in our case \( x_G^1 = \varphi \) then the Boltzmann equation (5) is solved by applying a splitting of Strang [114] as

\[
\begin{align*}
B_{||}\frac{\partial F_s}{\partial t} + \nabla_{\perp} \cdot \left( B_{||} \frac{d\mathcal{X}_G}{dt} F_s \right) &= 0 \quad \text{at } (\varphi, v_G^\parallel) \text{ fixed} \quad (54) \\
B_{||} \frac{\partial F_s}{\partial t} + \frac{\partial}{\partial \varphi} \left( B_{||} \frac{d\mathcal{X}_G}{dt} F_s \right) &= 0 \quad \text{at } (\mathcal{X}_G, v_G^\parallel) \text{ fixed} \quad (55) \\
B_{||} \frac{\partial F_s}{\partial t} + \frac{\partial}{\partial v_G^\parallel} \left( B_{||} \frac{d\mathcal{X}_G}{dt} F_s \right) &= 0 \quad \text{at } x_G \text{ fixed} \quad (56)
\end{align*}
\]

This splitting into three equations was introduced in the 4D version of the code. As explained in [55], in the 4D drift-kinetic slab case the conservative and advective forms of the equations are equivalent in the \( (r, \theta) \) direction and \( z \) direction separately due to the independent vanishing divergence property, see equations (9)-(10) in [55] due to \( \nabla \cdot v_G = 0 \) and \( \partial_z v^\parallel = 0 \). In the 5D gyrokinetic case \( \nabla_{\perp} \cdot (B_{||}^s d_s \mathcal{X}_G) \neq 0 \) and \( \partial_z (B_{||}^s d_s v^\parallel) \neq 0 \) so that these terms should be taken into account as source terms of the advective form of equations (54) and (55). However, they are presently set to zero. This simplification may alter the accuracy of the conservation properties of the code (see section 5). Also, it likely has an impact on the maximal value acceptable for the discretization time step. A solution to overcome this problem could be to use a conservative scheme instead of the BSL one but the first tests we have performed [14] are not conclusive. This constraint on the numerical value of \( \Delta t \) is acceptable for ion turbulence simulations but could become problematic when addressing kinetic electrons. The development of more efficient semi-Lagrangian schemes is still an active axis of research. An idea currently under evaluation is to separate and to treat differently the linear and non-linear parts. Encouraging results have been presented in [85]. The splitting operation stays a drawback of the semi-Lagrangian method. An alternative method without splitting (based on a 4D advection and 4D cubic spline interpolation) is currently developed. The first drawback is an increase of the numerical diffusion due to the 4D interpolation which will require to be quantified.

In the current version, the advections in \( \varphi \) and \( v_G^\parallel \) directions are straightforward, but that in the \( \mathcal{X}_G \) direction requires more attention. If we consider the 2D advection in \( (r, \theta) \) direction between times \( t \) and \( t + \Delta t \), the value of the electric field \( E \) at time \( t + \Delta t/2 \) is required in second order time scheme. This value is calculated by using a predictor-corrector method. Besides, computing these 2D trajectories is equivalent to solving \( d\mathcal{X}_G/dt = V(\mathcal{X}_G, \varphi, t), V \) being the advection field. This system is solved by using the parabolic assumption developed in [111]. Let \( \mathcal{X}_{G,i} \) be the position of \( \mathcal{X}_G(t_n + \Delta t) \) at time \( t_n + \Delta t \), then there exists a displacement \( d_{ij} = (\alpha_{ij}, \beta_{ij}) \) tangent to the parabola such that \( \mathcal{X}_G(t_n) = \mathcal{X}_{G,i} - d_{ij} \) and \( \mathcal{X}_G(t_n - \Delta t) = \mathcal{X}_{G,i} - 2d_{ij} \). The displacement \( d_{ij} \) can be calculated by solving the implicit equation \( d_{ij} = \Delta t V(\mathcal{X}_{G,i} - d_{ij}, t_n) \) (see p. 129 in [56]). This is done with a Taylor expansion which is equivalent at second order to a Newton algorithm.
3.1.3. GYSELA global algorithm

Concerning now the complete Boltzmann equation (5) the right hand side is also split to treat separately the collision operator, the Krook operator, the diffusion and source terms. Let $\hat{X}_G$ denote the shift operator in the poloidal cross-section $(r, \theta)$ associated to equation (54) over a time step $\Delta t$. Similarly, $\hat{\phi}$ and $\hat{v}_{\parallel}$ denote the shift operators respectively in the $\varphi$ (equation (55)) and $v_{\parallel}$ directions (equation (56)). As described in the previous paragraph, each of these three shift operators are based on a backward semi-Lagrangian scheme which means two steps for each mesh point: (i) first the computation of the characteristic feet and (ii) second an interpolation by cubic splines. Let us denote $\hat{J}$ the collision operator corresponding to solving $\partial_t \hat{F}_s = \mathcal{C}(\hat{F}_s)$ and $\hat{D}$ the operator associated to the radial diffusion $\partial_r \hat{F}_s = D_r(\hat{F}_s)$. They are both solved by applying a Crank-Nicolson scheme (see Appendices AppendixB.2 and AppendixE). The Krook operator $\hat{K}$ corresponding to the Krook diffusion equation $\partial_t \hat{F}_s = \mathcal{K}(\hat{F}_s)$ is trivial while solving $\partial_t \hat{F}_s = \mathcal{S}$ associated to the source operator $\hat{S}$ is described in Appendix AppendixD.5. Then, using these notations the following sequence is used to solve the 5D Boltzmann equation (5)

$$\text{Boltz} \equiv \left[ \left( \frac{\hat{K}_r}{2}, \frac{\hat{D}_r}{2}, \frac{\hat{\mathcal{C}}}{2}, \frac{\hat{S}}{2} \right), \left( \frac{\hat{v}_{\parallel}}{2}, \hat{\phi}, \hat{\mathcal{C}}, \frac{\hat{v}_{\parallel}}{2} \right), \left( \frac{\hat{S}}{2}, \frac{\hat{\mathcal{C}}}{2}, \frac{\hat{D}_r}{2}, \frac{\hat{K}_r}{2} \right) \right]$$  (57)

where the factor $1/2$ means that the operator is applied on half a time step. The choice of the sequence (57) is not unique but some constraints are imposed in the code: (i) the first one is to impose a symmetry to keep second order accuracy in the splitting (Strang splitting [114]), (ii) the second one is to fix the 2D operator $\hat{X}_G$ which is the most costly at the middle of the algorithm; (iii) finally the operators coupled to the right hand side, respectively to the Vlasov equation, are contiguous. Let us also define the operator $\hat{Q}$ (corresponding to the Poisson solving) which denotes symbolically the four steps: (i) computation of right hand side of the quasi-neutrality equation (12) using expression (13), (ii) solving the QN equation (12) to deduce the electrostatic potential $\hat{U}$, (iii) computation of the gyro-averaged electric potential $\hat{U} = \hat{J}_{eq} \hat{U}$ with a Padé approximation (18) or with an integration on the gyro-circles as described in section 2.5 and (iv) computation of the electric field as $\mathbf{E} = -\nabla \hat{U}$.

Finally, the global numerical algorithm of the GYSELA code can be summarized (see schematic view in Figure 3) as follows,

1. Initialization

Considering a prescribed magnetic field $B(r, \theta)$ (see section 2.2) and equilibrium profiles of density $n_0(r)$, ion temperatures $T_{\text{eq}}(r)$ and safety factor $q(r)$ (see section 2.6), then

(a) Computation of the equilibrium distribution function $\hat{F}_{s,eq}$ as a local or canonical Maxwellian by using (19) and (20).

(b) Initialization of $\hat{F}_s(t = 0)$ as $\hat{F}_s(t = 0) = \hat{F}_{s,eq}(1 + \text{perturbation})$ as described in section 2.6.

2. For each time iteration,

Considering the distribution function $\hat{F}_{s}^n = \hat{F}_s(t = t_n)$ at time $t_n$ known on the 5D mesh grid, then the distribution function $\hat{F}_{s}^{n+1}$ at the next time $t_{n+1} = t_n + \Delta t$ on the same mesh grid is computed by using a predictor-algorithm as

(a) Computation of the electric field $\mathbf{E}(t_n)$ by using the $\hat{Q}$ sequence.

(b) Prediction on $\Delta t/2$:

- Computation of $\hat{F}_s(t = t_n + \Delta t/2)$ by solving $\text{Boltz}/2$ sequence with $\mathbf{E}(t_n)$.
- Computation of $\mathbf{E}(t_n + \Delta t/2)$ by solving $\hat{Q}$.

(c) Correction on $\Delta t$:

- Starting from $\hat{F}_s(t = t_n)$ –given that this 5D distribution function has been stored before prediction –, computation of $\hat{F}_s(t = t_n + \Delta t)$ by solving $\text{Boltz}$ on a time step $\Delta t$ with the electric field $\mathbf{E}(t_n + \Delta t/2)$ at time $t_{n+1}/2$. 

19
3.2. An efficient hybrid OpenMP/MPI parallelization

The code is developed in Fortran 90 with some I/O routines in C (47k lines of Fortran 90 and 2.3k lines of C code) using GIT as version control system. Doxygen is used to generate documentation from the source code. The only external library dependence is the HDF5 library. HDF5 is the chosen format for all output saving, both 0D to 3D diagnostics and 5D restart files. Diagnostic analyses are performed with Python. The parallelization is based on a hybrid MPI/OpenMP paradigm. This hybrid approach is suitable for cluster of SMP (symmetric shared memory multiprocessor) nodes where MPI provides communication across nodes and OpenMP exploits loop level parallelism within a node. Let us denote by $N_{\text{species}}$ the number of ion species and by $N_{\mu}$ the number of $\mu$ values. As the magnetic momentum $\mu$ is an adiabatic invariant it plays the role of a parameter. So for each species we have $N_{\mu}$ independent Boltzmann equations (5) to solve. Let $N_r$, $N_{\theta}$, $N_{\varphi}$ and $N_{v_G}$ be the number of points in each directions $r$, $\theta$, $\varphi$ and $v_G$. Large data structures are used in GYSELA: 5D data of size $N_r \times N_{\theta} \times N_{\varphi} \times N_{v_G} \times N_{\mu}$ for distribution functions and 3D data of size $N_r \times N_{\theta} \times N_{\varphi}$ for the electrostatic potential and its derivatives as for first moments of the distribution function (used for diagnostics). An MPI parallelization is mandatory to treat such large amount of data. Let us take the example of a typical 5D mesh used for our simulations\(^5\), namely $(N_r \times N_{\theta} \times N_{\varphi} \times N_{v_G} \times N_{\mu}) = (256 \times 128 \times 128 \times 128 \times 16)$. It corresponds to a mesh of almost 8.6 billion points. The size of one 5D array for the distribution function is of the order of 68 Gbytes, which is not tractable on a single node. For information, the biggest simulation run so far with GYSELA was an ITER simulation \cite{ITER} with 272 billion points. Taking into account the fact that two distribution functions are necessary for the numerical integration over time due to predictor-corrector scheme, more than 1 Tbytes of data (just for 5D arrays) were manipulated. So, as described in the following, we use a domain decomposition so that a MPI process never contains the complete 5D distribution function.

Concerning the MPI parallelization, an MPI communicator is defined per species. Inside each one of the MPI SPECIES communicators, an MPI communicator is defined for each value of the magnetic moment $\mu$.

\(^5\)For current GYSELA simulations, $N_{\mu}$ is chosen equal to 16 or 32 while the choice of $N_{v_G}$ is much larger (typically, $N_{v_G} \approx 4N_{\mu}$). Notice however that $N_{\mu}$ should be compared to $N_{v_G}/2$, since the grid in $\mu$ goes from 0 to $\mu_{\text{max}}$, while that in $v_G$ covers the range $-v_{\text{G},\text{max}}$ to $+v_{\text{G},\text{max}}$. The choice of $N_{v_G} > 128$ is necessary to take correctly into account the trapping and de-trapping of particles and also to solve accurately the collision operator (at this time, only in parallel velocity). The same number of points for $\mu$ direction will be required when collision operator effects will be added in perpendicular direction (mandatory step for kinetic electrons).
Within each MPI_MU communicator a 2D domain decomposition allows us to assign to each MPI process a sub-domain in $(r, \theta)$ dimensions. Let us consider $p_r$ (resp. $p_\theta$) the number of sub-domains in $r$ (resp. $\theta$) direction. The number of MPI processes used during one run is equal to $N_{MPI} = p_r \times p_\theta \times N_\mu \times N_{\text{species}}$. This MPI decomposition of the default MPI_COMM_WORLD communicator is summarized in Figure 4. Thus, each MPI process is then responsible for the sub-domain of the distribution function $\bar{F}_s(r = [i_{\text{start}}, i_{\text{end}}], \theta = [j_{\text{start}}, j_{\text{end}}], \varphi = *, v_G = *, \mu = \mu_{id})$ with the integer $\mu_{id} \in [0, N_\mu - 1]$. The local values $i_{\text{start}}, i_{\text{end}}, j_{\text{start}}$ and $j_{\text{end}}$ associated to the parallel decomposition are initially set by using a classical domain decomposition in $(p_r \times p_\theta)$ blocks. The OpenMP paradigm is used in addition to MPI. Let us denote by $N_{\text{thread}}$ the number of threads in each MPI process, then the number of cores for a simulation corresponds to $N_{\text{thread}} \times N_{MPI}$.

Figure 4: MPI_COMM_WORLD communicator decomposition for two species, 8 values of $\mu$, $p_r = 4$ radial sub-domains and $p_\theta = 2$ sub-domains in the poloidal direction. In this case, the number of MPI processes is equal to 128.

3.2.1. Parallel Vlasov solver: How to treat non-local interpolation aspects?

A difficulty when parallelizing the semi-Lagrangian Vlasov solver is due to the cubic spline interpolation. Cubic splines are a good compromise between simplicity and accuracy but a drawback is that they are non-local. Indeed, a lot of the values of the distribution function $\bar{F}_s$ are required to reconstruct the interpolated value of the function at any position in the domain. Two strategies are available in the code to overcome the problem: (i) local cubic spline interpolation or (ii) transposition. The Hermite spline interpolation on patches [28, 29, 83] has been specifically developed for dealing with 2D domains distributed on several MPI processes. The idea is to compute local cubic spline coefficients on each 2D $(r, \theta)$ sub-domains by solving reduced linear systems. Then one ensures a $C^1$ global interpolator similar to the sequential one by imposing Hermite boundary conditions at the interface of each patch [28]. The first limitation of this technique is that a minimum of 32 points per directions is needed per MPI process [83] (i.e $N_r/p_r \geq 32$ and $N_\theta/p_\theta \geq 32$) to provide good numerical stability and small communication overhead (in 2D, each processor has to exchange derivatives with its 8 neighboring processors). The second constraint is that the shift at one point on the border of a sub-domain, which results from the motion along the trajectories in the 4D phase space, must not exceed the elementary cell width. This constraint is linked to the choice made to limit the size of the interface transferred between processors. This CFL condition can be extremely restrictive specially in the $\theta$ direction where large shifts can occur but also in the radial direction when a source is imposed in the case of flux-driven simulations. For these reasons the choice of a 4D data transposition is now often preferred. This transposition consists in modifying in each MPI_COMM_MU communicator the parallel decomposition of $\bar{F}_s$.
such that each processor then contains only part of the data in \( \varphi \) and \( v_{G\parallel} \) direction but all the information in the poloidal cross-section. Standard cubic spline interpolation in \( (r, \theta) \) plane are then possible. Let us define the transposition operation \( T_F \) and its inverse \( T_F^{-1} \) as

\[
\tilde{F}_s(r_{\text{block}}, \theta_{\text{block}}, \varphi = *, v_{G\parallel} = *, \mu = \mu_{\text{id}}) \xleftarrow{T_F} \tilde{F}_s(r = *, \theta = *, \varphi_{\text{block}}, v_{G\parallel \text{block}}, \mu = \mu_{\text{id}})
\]

then the sequence (57) described above for solving the Boltzmann equation is replaced by the following one

\[
\text{Boltztransp} \equiv \left[ \left( \frac{\hat{K}_r}{2}, \frac{\hat{D}_r}{2}, \frac{\hat{C}}{2}, rac{\hat{S}}{2} \right), \left( \frac{\hat{G}_{||}}{2}, \frac{\hat{\varphi}}{2} \right), T_F \left( \hat{X}_G \right) T_F^{-1}, \left( \frac{\hat{\varphi}}{2}, \frac{\hat{G}_{||}}{2} \right), \left( \frac{\hat{S}}{2}, \frac{\hat{C}}{2}, \frac{\hat{D}_r}{2}, \frac{\hat{K}_r}{2} \right) \right]
\]

3.2.2. A scalable quasi-neutral solver

The parallel quasi-neutrality algorithm presently used in the code is summarized in algorithm 1. For more details on the different improvements which have been performed to obtain this parallel solver see \([86, 82]\). The presence of the non-local term \( \langle U \rangle_{FS}(r) \) couples the \( \theta \) and \( \varphi \) directions and penalizes the parallelization but the most important cost is the communication induced by the computation of the r.h.s, namely task 2 in algorithm 1. Indeed this calculation requires a collective communication step that involves all MPI processes. Instead of broadcasting \( U(r = *, \theta = *, \varphi = *) \) to all MPI processes, a refined strategy has been setup to reduce the large communication cost. It consists in sending to each process only a sub-domain in \( \varphi \) direction of \( U \). With this decomposition the gyro-average computation of \( U \) as the partial derivatives in \( r \) and \( \theta \) directions are straightforward. Afterwards, a transposition is performed to calculate \( \partial_\varphi U \). For a complete performance analysis of the different steps of the algorithm see \([87]\).

**Algorithm 1:** Quasi-neutrality algorithm in the GYSELA code

**Input:** \( \tilde{F}_s(r_{\text{block}}, \theta_{\text{block}}, \varphi = *, v_{G\parallel} = *, \mu = \mu_{\text{id}}) \)

1. \( v_{G\parallel} \) integration of \( \tilde{F}_s \) to compute within each MPI_COMM_MU communicator

\[
\text{intdvpar}_\text{Fs}(r_{\text{block}}, \theta_{\text{block}}, \varphi = *, \mu = \mu_{\text{id}}) = \int J_r \, dv_{G\parallel} \left( \tilde{F}_s - \tilde{F}_{s,\text{eq}} \right)
\]

2. Remapping within each MPI_COMM_MU communicator of \( \text{intdvpar}_\text{Fs} \) –because the gyroaverage operation requires to have all the data for each \( (r, \theta) \) plane– as

\[
\text{intdvpar}_\text{Fs}(r_{\text{block}}, \theta_{\text{block}}, \varphi = *, \mu = \mu_{\text{id}}) \Rightarrow \text{intdvpar}_\text{Fs}(r = *, \theta = *, \varphi_{\text{block}}, \mu_{\text{block}}).
\]

3. For all \( \varphi \), computation of the gyroaverage of \( \text{intdvpar}_\text{Fs} \) and integration over \( \mu \) to obtain

\[
\rho_s(r = *, \theta = *, \varphi_{\text{block}}) = \int d\mu_s J_\mu \cdot (\text{intdvpar}_\text{Fs}(r = *, \theta = *, \varphi_{\text{block}}))
\]

4. MPI reduction towards the \( N_{\text{species}} \times \text{MPI_COMM_SPECIES} \) communicators to finally compute \( \rho = \sum_s Z_s \rho_s \), the right hand side of the quasi-neutrality equation (12).

5. Solving (14) and (15) to deduce \( U(r = *, \theta = *, \varphi_{\text{block}}) \) and broadcast to the \( N_{\text{species}} \times N_{\mu} \) communicators.

**Output:** \( U(r = *, \theta = *, \varphi_{\text{block}}) \) on each MPI process.
3.2.3. Performing weak and strong scaling

The weak and strong scaling, presented in this section, have been performed on two different high performance computers. The strong scaling has been tested on the thin nodes of the Curie machine at TGCC\(^6\), Bruyères-le-Châtel, France which are based on bullx B510 architecture. Each node contains two 8 cores processors (INTEL Xeon E5-2680 Sandy-Bridge) running at 2.7 GHz with 64 GB of 1.6 GHz memory. The weak scaling has been performed on the IBM Blue Gene/Q machine JUQUEEN at JSC/IAS\(^7\), Juelich, Germany which is composed of 24 racks grouping 1024 nodes. Each node contains a single 17-cores processor running at 1.6 GHz with 16 GB of 1.33 GHz memory. These cores are twice slower compared to Sandy Bridge cores. The improvements of the code to adapt it efficiently to the BlueGene architecture are detailed in [11]. The scaling results are detailed in Figures 5 and 6 for the 4 main components of the code, namely (i) Vlasov solver, solving the Boltzmann equation as described in section 3.2.1; (ii) Field solver, solving the quasi-neutrality equation as summarized in algorithm 1; (iii) Derivatives computation, the computation of the first derivatives of the gyroaveraged electrostatic potential \(J_\mu U\) in the three \(r\), \(\theta\) and \(\varphi\) directions and (iv) Diagnostics corresponding to all physical quantities from 0D to 3D computed and saved in HDF5 format like densities, parallel and perpendicular temperatures, velocities, fluxes, energies et caetera. The strong scaling has been performed with the mesh size parameters \(N_r = 512\), \(N_\theta = 512\), \(N_\varphi = 128\), \(N_{vq1} = 128\) and \(N_\mu = 32\). The number of threads was kept constant \(N_{\text{thread}} = 8\) to assign two MPI process per node. The couple \((p_r, p_\theta)\) of processors in \(r\) and \(\theta\) directions take the following values \((2, 4), (8, 2), (8, 4), (16, 4), (16, 8)\) and \((16, 16)\) so that the scaling spans from 2048 to 65536 cores. For the weak scaling the testbed case was composed from 64k to 458k cores. The considered meshes vary from 17.18 to 481 billion points defined with \(N_r = 512\), \(N_\theta = 1024\), \(N_\varphi = 128\), \(N_{vq1} = 128\) and 7 different values of \(N_\mu = 2, 4, 8, 16, 32, 48\) and 56. The triplet \((p_r,p_\theta,N_{\text{thread}})\) is kept constant equal to \((16,32,64)\). The number of threads is chosen equal to 64 such that a single MPI process is mapped per node and 4 threads are dedicated per core as determined for BlueGene/Q optimization. The results of both scalings are summarized in Figures 5 and 6. Let us first remark that the weak scaling test (Fig. 6) exhibits an excellent scalability of the code with 90.9% of relative efficiency at 458752 cores—which corresponds to the totality of the JUQUEEN computer—compared to 16k. Concerning the strong scaling GYSELA globally scales with a relative efficiency of 89% at 16k cores and 60.6% on 65k cores compared to 2048 cores (Fig. 5). This is already a very good result for such a semi-Lagrangian code. Looking into more detail, we see that the deterioration from 89% at 16k to 60.6% at 65k is mainly due to the diagnostics and field solver. Indeed, the Vlasov solver which represents 60% of the application at 2k cores and 48% at 65k cores exhibits a good efficiency of 74.6% at 65k processors. Conversely, the diagnostics which correspond to 37.5% of the total time for the reference case, end taking as much time as the Vlasov solver due to the decreased efficiency at 55.8%. A lot of work has already been done to improve the field solver [82, 86, 87] but this work must continue because an efficiency of 41% impacts the global scalability. Let us finally notice that even if the computation time of the derivatives remains negligible until 16k, a further effort is needed to prepare the code to future exascale machines. The execution times are not comparable in the two scalings presented here because the considered meshes are not the same. Performance comparisons between BlueGene/Q and Bullx architectures are found in [11]. Production runs are commonly running on both architectures. In general, depending on simulation parameters, the code is a factor 2.5 to 3 times faster on bullx machines (as Curie or Helios at CSC, Rokkasho, Japan) than on BlueGene machines (as JUQUEEN or Turing at IDRIS, Orsay, France or Fermi at CINECA, Bologna, Italy). This behavior is consistent with that observed with many other codes.

3.2.4. Memory scalability

Due to the previous scalability results GYSELA uses frequently from 8k to 32k cores for one ion species with adiabatic electrons and the twice these values when an impurity is taken into account. Besides, a simulation often runs during several weeks. The annual time consumption on HPC machines is currently of 51 millions of core hours. So the code already benefits from petascale computational power of the current high

\(^6\)http://www-hpc.cea.fr/en/complexe/tgcc.htm
\(^7\)http://www.fz-juelich.de/ias/jsc/EN/Home/home_node.html
Figure 5: Strong scaling performed on the Curie machine from 2048 to 65536 cores: Execution time (a) and relative efficiency (b) for one Gysela run of 4 iterations for a mesh \((N_r \times N_\theta \times N_\phi \times N_{\mu G} \times N_{\mu}) = (512 \times 512 \times 128 \times 128 \times 32)\) with 8 threads and 92 values in \(\mu\) direction.

Figure 6: Weak scaling performed on the JUQUEEN machine from 64k to 458k cores: Execution time (a) and relative efficiency (b) for one Gysela run of 4 iterations for a mesh fixed in 4D as \((N_r \times N_\theta \times N_\phi \times N_{\mu G}) = (512 \times 1024 \times 128 \times 128)\) but for 7 different values of \(\mu = 2, 4, 8, 16, 32, 48, 56\). The number of parallel domains in \(r, \theta\) directions and the number of threads are fixed such that \((p_r, p_\theta, N_{\text{thread}}) = (16, 32, 64)\). The number of cores varies as \(p_r \times p_\theta \times N_\mu \times N_{\text{thread}}/4\) because 4 threads per core are used for BlueGene optimization.
performance computers. We also know that at short term when adding kinetic electrons, a simulation with ITER parameters on several confinement times will require exascale HPC capabilities. Among the Exascale challenges, the reduced memory per core has been identified as one of the most critical. This is particularly true for the GYSELA code due its global character which requires huge 5D meshes. The mesh discretization is already constrained by the memory required per node. The C/Fortran MTM (Modelization & Tracing Memory consumption) library has been developed [103] to investigate in detail the memory consumption of the code. This library provides an Application Programming Interface (API) which replaces the standard calls to allocation/de-allocation routines. This intrusive technique permits to retrieve precisely the peak of memory consumption and all the arrays involved. External Python scripts have been designed to analyze these results and provide memory prediction. Continued efforts are made to reduce the memory footprint of the code and to improve its memory scalability (see [103] for details) but Table 1 shows that lots of work remains to be done to be able to run ITER simulations on machines with only 16GB per node. Indeed, for a minimum ITER mesh, namely \((N_r \times N_\theta \times N_\phi \times N_{vG} \times N_\mu) = (1024 \times 1024 \times 256 \times 128 \times 16)\), Table 1 shows that such simulations can run on Curie Machine (64 GB/node) with 8192 cores but would require 524288 cores on the BlueGene current architecture (16 GB/node)- knowing that the optimal number of threads is 64. This number of cores is still out of reach on European high performance computers. Currently, as shown in Table 2 and in Figure 7, the global memory scalability of the code is of 47.8%. This strong scaling was performed with a constant 5D mesh of \((1024 \times 4096 \times 1024 \times 128 \times 2)\) points using MTM prediction mode and varying the number of MPI processes from 128 to 2048 by increasing the number of points in \(r\) and \(\theta\) directions. The 4D structures are very scalable (relative efficiency of 89.9%) compared to the others (see Fig.7). Indeed, 3D structures with 23.8% represent 32% of the global cost for 32k cores against 16% for 2048 cores. The 3D structures are then no more negligible. The scalability of the 1D structures is with 19% of the order of the 3D arrays but the cost is still not significant. At the opposite, the incompressibility of 2D structures between 256 and 2k MPI processes is extremely penalizing leading to an increase of the cost percentage by a factor 10. The fact that the memory for 2D structures remains constant (equal to 7.1 GB/second) is probably due to incompressible temporary arrays allocated for OpenMP loops. Improvement of memory footprint of 3D and 2D structures will be pursued in the future.

Another bottleneck for Exascale applications will be the possibly increased crash probability following that of the number of cores of future machines [24]. The Weibull law [109] gives an estimate of the time between two crashes which is of the order of the minute for a number on nodes larger than \(10^6\). Two approaches have been explored until now to try to improve the fault tolerance of the code. The first one consists in employing an asynchronous method to increase the frequency of the restart file writing. A GYSELA simulation runs typically several days or even weeks. So the complete simulation is split into a series of jobs of approximately 10 hours with automatic re-submission. The restart files are not only saved at the end of each job but several times during the job in parallel to calculations (see [119] for complete description). Another checkpointing using the FTI library [6] (high performance Fault Tolerance Interface for hybrid systems) is under investigation. The main idea is to benefit from fast access local SSD disks available on some HPC architectures.

<table>
<thead>
<tr>
<th>MPI procs</th>
<th>OpenMP threads</th>
<th>16</th>
<th>32</th>
<th>64</th>
<th>128</th>
</tr>
</thead>
<tbody>
<tr>
<td>128</td>
<td>126.1 GB</td>
<td>126.9 GB</td>
<td>128.7 GB</td>
<td>132.2 GB</td>
<td></td>
</tr>
<tr>
<td>512</td>
<td>35.2 GB</td>
<td>36.1 GB</td>
<td>37.8 GB</td>
<td>41.3 GB</td>
<td></td>
</tr>
<tr>
<td>2048</td>
<td>16.4 GB</td>
<td>16.4 GB</td>
<td>16.5 GB</td>
<td>23.2 GB</td>
<td></td>
</tr>
<tr>
<td>8192</td>
<td>12.5 GB</td>
<td>12.5 GB</td>
<td>12.5 GB</td>
<td>19.1 GB</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Memory peak (in GBytes) depending on the number of MPI procs and of OpenMP threads for a 5D mesh \((N_r \times N_\theta \times N_\phi \times N_{vG} \times N_\mu) = (1024 \times 1024 \times 256 \times 128 \times 16)\).
Table 2: Strong scaling for each kind of data for a 5D mesh (1024 × 4096 × 1024 × 128 × 2): (first lines) memory allocation size in GBytes and (second lines) percentage with respect to the total memory at the peak of the memory consumption. (Table from [103]).

<table>
<thead>
<tr>
<th></th>
<th>Number of cores</th>
<th>2k</th>
<th>4k</th>
<th>8k</th>
<th>16k</th>
<th>32k</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of MPI processes</td>
<td>128</td>
<td>256</td>
<td>512</td>
<td>1024</td>
<td>2048</td>
</tr>
<tr>
<td>4D structures</td>
<td>Memory size</td>
<td>207.2 GB</td>
<td>104.4 GB</td>
<td>53.7 GB</td>
<td>27.3 GB</td>
<td>14.4 GB</td>
</tr>
<tr>
<td></td>
<td>Memory percentage</td>
<td>79.2%</td>
<td>71.5%</td>
<td>65.6%</td>
<td>52.2%</td>
<td>42.0%</td>
</tr>
<tr>
<td>3D structures</td>
<td>Memory size</td>
<td>42.0 GB</td>
<td>31.1 GB</td>
<td>18.6 GB</td>
<td>15.9 GB</td>
<td>11.0 GB</td>
</tr>
<tr>
<td></td>
<td>Memory percentage</td>
<td>16.1%</td>
<td>21.3%</td>
<td>22.7%</td>
<td>30.4%</td>
<td>32.1%</td>
</tr>
<tr>
<td>2D structures</td>
<td>Memory size</td>
<td>7.1 GB</td>
<td>7.1 GB</td>
<td>7.1 GB</td>
<td>7.1 GB</td>
<td>7.1 GB</td>
</tr>
<tr>
<td></td>
<td>Memory percentage</td>
<td>2.7%</td>
<td>4.9%</td>
<td>8.7%</td>
<td>13.6%</td>
<td>20.8%</td>
</tr>
<tr>
<td>1D structures</td>
<td>Memory size</td>
<td>5.2 GB</td>
<td>3.3 GB</td>
<td>2.4 GB</td>
<td>2.0 GB</td>
<td>1.7 GB</td>
</tr>
<tr>
<td></td>
<td>Memory percentage</td>
<td>2.0%</td>
<td>2.3%</td>
<td>3.0%</td>
<td>3.8%</td>
<td>5.1%</td>
</tr>
<tr>
<td>Total per MPI process in GBytes</td>
<td>261.5</td>
<td>145.9</td>
<td>81.8</td>
<td>52.3</td>
<td>34.2</td>
<td></td>
</tr>
</tbody>
</table>

Figure 7: Memory relative efficiency for a GYSELA simulation at the memory peak of time consumption for the four kind of structures used in the code (1D arrays to 4D arrays). The results are extracted from Table 2. The reference point corresponds to 128 MPI processes.
4. GYSELA verification

Since the first steps 15 years ago, the code has evolved significantly, including more and more physics with more and more numerical complexity and high level of parallelism. Portability of the code on several HPC architectures has been achieved. In such a complex code it becomes very difficult to track or to propagate changes by being sure to reproduce previous results and simulations. Over the past two years, a large effort has been made to improve the development process with due regard to modularity, reproducibility and efficiency. This approach is based on the joint use of a version control system (GIT\(^8\)) together with that of a continuous integration platform such as JENKINS\(^9\). At each commit on the GIT reference branch, automatic compiling and executing are submitted. For more details on our strategy the reader can refer to [10]. A database of non-regression tests is also under construction with the objective to be run less frequently (every night or week) but to ensure that new changes have no impact on well-established results. All the tests proposed in this section, which were used for GYSELA verification\(^10\), are detailed with the objective to become part of the non-regression database. They all correspond to 5D gyrokinetic simulations for one species. Other verification tests can be founded on simpler 4D models in [55, 85].

4.1. Normalization

The GYSELA code is written in normalized units based on the following normalization choices. We use SI units and a thermal energy scale in electron volts (\(1\,eV = 1.6022\times 10^{-19}\,J\)). The four fundamental dimensional normalizing quantities are: a reference ionic mass \(m_0 = A_0 m_p\) (Kilogram), a reference ionic charge \(q_0 = Z_0 e\) (Coulomb), a reference magnetic induction \(B_0\) (Tesla) and a reference thermal energy \(T_0\) (\(eV\)). Here, \(A_0\) and \(Z_0\) are the (dimensionless) mass number and charge state of the main ion species and \(e\) the modulus of the electron charge. These quantities are used to define the reference ion cyclotron frequency \(\Omega_{\ast 0}\), the reference thermal speed \(v_{T\ast 0}\) and the reference Larmor-radius \(\rho_0\) as \(\Omega_{\ast 0} = \frac{Z_0 e B_0}{m_0}\), \(v_{T\ast 0} = \sqrt{\frac{T_0}{m_0}}\) and \(\rho_0 = \frac{v_{T\ast 0}}{\Omega_{\ast 0}} = \frac{\sqrt{T_0 m_0}}{Z_0 e B_0}\).

Finally, we choose the equilibrium electron density at mid radius \(n_0\) as reference density. Physical quantities (mass, length, time, charge and density) can be recovered from the normalized quantities used in the code (denoted with a hat symbol) by choosing values for \([A_0, Z_0, B_0, T_0, n_0]\) and applying \(m_s = A_s m_p = m_0 \hat{A}_s\) with \(A_s = A_0 \hat{A}_s\), \(q_s = Z_0 e \hat{Z}_s\), \(l = \rho_0 \hat{l}\), \(t = \frac{\mu_0}{T_0}\) and \(n_s = n_0 \hat{n}_s\). The velocities are normalized to the corresponding thermal velocities \(v_{T\ast 0} = \sqrt{T_0/m_s}\), \(i.e\. v = v_{T\ast 0} \hat{v}_s = \frac{\mu_0}{\sqrt{\hat{A}_s}} \hat{v}_s\). The main normalizations are thus \(U = \frac{T_0}{Z_0 e} \hat{\phi}, B = B_0 \hat{B}, T_s = T_0 \hat{T}_s\) while \(\mu_s = \frac{T_0}{B_0} \hat{\mu}_s\) with \(\hat{\mu}_s = \frac{\mu_0}{2 B}\) and \(\mu_0 J = \frac{B_0}{\rho_0} \hat{J}\) with \(\hat{J} = \hat{\nabla} \times \hat{B}\).

By deduction, \(\mathbf{E} = E_0 \hat{\mathbf{E}}\) with \(E_0 = v_{T\ast 0} B_0\) and the normalized distribution function \(\hat{F}_s\), which evolves in the code, is defined as \(\hat{F}_s = F_s v_{T\ast 0}^3 / n_0\). Finally, the energy is normalized to the reference thermal energy \(T_0\). The subsequent normalized equations used in the code are presented in Appendix AppendixF. In what follows, all the quantities considered are normalized coordinates, but hat symbols are omitted for the sake of readability.

4.2. Invariance test

In the present work the source terms are not taken into account. Let us call \texttt{simul} a first fixed gradient simulation (\(i.e\. no source term\)) for one unique species of charge \(Z_1\) and mass \(A_1\) and \texttt{simu2} a second one for a species of charge \(Z_2 = \alpha Z_1\) and mass \(A_2 = \alpha A_1\). Then it is possible to define the other parameters of the second simulation, only depending on the multiplying factors \(\alpha Z \in \mathbb{N}\) and \(\alpha A \in \mathbb{N}\) such that both simulations \texttt{simul} and \texttt{simu2} are identical. The idea is to define an invariance test which permits to verify that charge and mass are correctly taken into account in the code. Notice that a transformation departs from a simple \(\rho_s\) scaling, where only the \(\rho_s\) parameter is modified (\(v.e\. a change of the minor radius at constant aspect ratio\)) [23, 92]. Such a test can be defined as a non-regression test of any gyrokinetic...
code but it is not generic. It depends on the equations treated in each code. For any code, the solved equations (gyrokinetic + quasi-neutrality) are invariant under certain groups of transformations. Within these groups, the invariance is exact provided that the code actually solves the equations it is supposed to. We have derived in Appendix AppendixG such a group of transformations which leaves the solution unchanged for the GYSELA code. Seven control parameters are required \( \alpha_x, \alpha_t, \alpha_T, \alpha_\phi, \alpha_{coll}, \alpha_{diff} \) and \( \alpha_{Krook} \) that respectively rescale space, time, electron temperature, electrostatic potential, collision operator, diffusion term and Krook operator. The only issue which cannot be accounted for (i.e. which cannot be rescaled) in this operation is the impact of the boundary conditions. The invariance constraints (i.e. the relationships between the various scaling factors which leave the Boltzmann equation invariant under these transformations), as summarized in Table 3, are \( \alpha_x = \sqrt{\alpha_A/\alpha_Z}, \alpha_t = \alpha_A/\alpha_Z, \alpha_T = 1/\alpha_Z, \alpha_{coll} = 1, \alpha_{diff} = 1/\alpha_Z \) and \( \alpha_{Krook} = \alpha_Z/\alpha_A \). Then, the relation between the electrostatic potential \( \phi_1 \) solution of the first simulation and \( \phi_2 \) the electrostatic potential of the second one is \( \phi_2 = \alpha_\phi \phi_1 \) with \( \alpha_\phi = 1/\alpha_Z \).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Simulation 1</th>
<th>Simulation 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charge</td>
<td>( Z_1 )</td>
<td>( Z_2 = \alpha_Z Z_1 )</td>
</tr>
<tr>
<td>Mass</td>
<td>( A_1 )</td>
<td>( A_2 = \alpha_A A_1 )</td>
</tr>
<tr>
<td>Mesh discretization</td>
<td>( \Delta x_1 )</td>
<td>( \Delta x_2 = \sqrt{\alpha_Z} \Delta x_1 )</td>
</tr>
<tr>
<td>Time step ( \Delta t )</td>
<td>( \Delta t_1 )</td>
<td>( \Delta t_2 = \frac{\alpha_Z}{\alpha_Z} \Delta t_1 )</td>
</tr>
<tr>
<td>Temperature ( T_e )</td>
<td>( T_{e1} )</td>
<td>( T_{e2} = \frac{1}{\alpha_Z} T_{e1} )</td>
</tr>
<tr>
<td>Collision coefficient</td>
<td>( K_{coll1} )</td>
<td>( K_{coll2} = K_{coll1} )</td>
</tr>
<tr>
<td>Krook coefficient</td>
<td>( K_{krook1} )</td>
<td>( K_{krook2} = \frac{\alpha_Z}{\alpha_A} K_{krook1} )</td>
</tr>
<tr>
<td>Diffusion coefficient</td>
<td>( K_{diff1} )</td>
<td>( K_{diff2} = \frac{1}{\alpha_Z} K_{diff1} )</td>
</tr>
</tbody>
</table>

Table 3: Parameter dependence between two equivalent simulations. \( \alpha_Z \) and \( \alpha_A \) are the multiplying factors between both cases respectively for charge and mass.

Three different simulations were run to test the invariance property of the code, considering first a reference plasma with Hydrogen (mass \( A_s = 1 \) and charge \( Z_s = 1 \)), then Helium (\( A_s = 4 \) and \( Z_s = 2 \)) and finally Tungsten (\( A_s = 150 \) and \( Z_s = 50 \)). For the following, let us respectively call simu1A121, simu2A422 and simu3A150250 these three simulations. These cases are not relevant in terms of physics but have been designed for numerical tests. The idea was to define small cases tractable as non-regression tests. So the reference simulation is based on a small plasma (\( \rho_s = 1/150 \)) for a radial domain between \( 0.15p \) and \( 0.85p \) (with \( \rho = r/a \)) and for half a torus. The idea was to define small cases tractable as non-regression tests. So the reference simulation is based on a small plasma (\( \rho_s = 1/150 \)) for a radial domain between \( 0.15p \) and \( 0.85p \) (with \( \rho = r/a \)) and for half a torus. The radial profiles of density \( n_s(r) \), temperature \( T_{s0}(r) \) and safety factor \( q(r) \) are analytically prescribed as \( d \log n_s(r)/dr = -2.2 \cos^2((\rho-0.5)/0.04) \), \( d \log T_{s0}(r)/dr = -8 \cos^2((\rho-0.5)/0.04) \) with \( \rho \approx r - r_{\text{min}}/(r_{\text{max}} - r_{\text{min}}) \) and \( q(r) = 1.5 + 1.7 \exp(2.8 \log(r/a)) \). A Krook operator (see section 2.10) of amplitude 0.01 and a diffusion (eq.(51)) of 0.015 are applied in a buffer region defined by eq.(50) with a location \( B_L = 0.06 \) and a stiffness of 0.017635. The collision operator (eq.(30)) is applied every iterations while \( T_{s,\text{coll}} \) (eq.(32)) and \( V_{\|,\text{coll}} \) (eq.(31)) are refreshed every 10 iterations. All the numerical parameters of this simu1A121 simulation are summarized in Tables 4 and 5. For the complete description of the case see the associated input data file of the code (Figure L.19) in Appendix AppendixL.

So taking into account the equivalences defined in Table 3, the second simulation simu2A422 is defined with the same parameters as simu1A121 except the fact that Helium is considered instead of Hydrogen and (i) \( \tau = 2 \), (ii) \( \Omega_{co} \Delta t = 10 \), (iii) \( D_{\text{coeff}} = 0.0075 \), (iv) \( K_{\text{krook coeff}} = 0.005 \) while (v) the diagnostic time step \( \Omega_{co} \Delta t_{\text{diag}} = 100 \) instead of 50. The species concentration \( c_s \) has been divided by 2 to satisfy the constraint \( \sum_s c_s Z_s = 1 \). The third simulation simu3A150250 is performed for a Tungsten species \( A_s = 150 \) and \( Z_s = 50 \) and differs from simu1A121 by the following parameters: (i) \( 1/\rho_s = 36.7423 \) (\( 1/\rho_{\text{simu1}} \times \sqrt{150/50} \)), (ii) \( c_s = 0.02 \), (iii) \( \tau = 50 \), (iv) \( \Omega_{co} \Delta t = 15 \), (v) \( \Omega_{co} \Delta t_{\text{diag}} = 150 \), (vi) \( D_{\text{coeff}} = 0.0003 \)
Collision operator

\[ ν \]

factors have been correctly implemented in the code equations. The same order of magnitude is obtained for the small size of the plasma considered (\( a \approx 10^{−3} \) simulations. It exhibits a relative error of \( 10^{−3} \) to the location and the stiffness of the buffer regions where both diffusion and Krook operator are applied with respective amplification coefficients and Krook coefficient. The collision operator is applied every \( \text{nbstep}_{\text{coll}} \) iterations but refreshed every \( \text{nbrefresh}_{\text{coll}} \) iterations.

\[ t = \frac{1}{150} \]

\[ R_\text{int}/a \]

\[ r_{\text{ext}}/a \]

\[ \text{Torus} \]

\[ Z_s \]

\[ A_s \]

\[ Ω_{\text{cs}} \Delta t \]

\[ q_1 \]

\[ q_2 \]

\[ q_3 \]

\[ κ_{n_s} \]

\[ Δ_{n_s} \]

\[ κ_{T_s} \]

\[ Δ_{T_s} \]

\[ T_1/T_e \]

Table 4: Main parameters for reference simulation simul_1A1Z1. The velocity phase space is defined by \( −n_{\text{b1}0}\nu_{\text{T,th}} \leq v_{\text{Gi}} \leq n_{\text{b1}0}\nu_{\text{T,th}} \) and \( 0 \leq μ \leq μ_{\text{max}} T_0/B_0 \). Torus indicates the fraction of the torus simulated. The safety factor radial profile is defined as \( q(r) = q_1 + q_2 \exp(q_3 \log(r/a)) \). The radial density profile is defined by its gradient as \( \frac{d \log n_{s0}(r)}{dr} = −κ_{n_{s0}} \cosh^{−2}((r−0.5)/Δr_{n_{s0}}) \). The same analytical expression is used for the temperature with \( κ_{T_{s0}} \) and \( Δ_{T_{s0}} \).

and (vii) Krook coefficient = 0.01/3. The differences between the three considered simulations are highlighted in Table 6. The time evolution of the zonal flow component \( (m = 0, n = 0) \), where \( m \) and \( n \) are respectively poloidal and toroidal mode number in the middle of the radial domain is presented in Figure 8 for the three simulations. It exhibits a relative error of \( 10^{−15} \) when Hydrogen case is compared to Helium case and of \( 10^{−7} \) when compared with Tungsten simulation. The loss of accuracy for Tungsten is probably due to the small size of the plasma considered \( (a = 36.7423/\rho_0) \). In this case boundary conditions can play a more important role. For each test the results are sufficiently accurate to consider that the charge and mass factors have been correctly implemented in the code equations. The same order of magnitude is obtained when the distribution function values are compared.

\[ B_L \]

\[ B_S \]

\[ D_{\text{coeff}} \]

\[ \text{Krook}_{\text{coeff}} \]

\[ ν_{s,\text{buff}} \]

\[ ν_{s,\text{coeff}} \]

\[ \text{nbstep}_{\text{coll}} \]

\[ \text{nbrefresh}_{\text{coll}} \]

Table 5: Main parameters for simul_1A1Z1 simulation concerning buffer region and collisions. \( B_L \) and \( B_S \) correspond respectively to the location and the stiffness of the buffer regions where both diffusion and Krook operator are applied with respective amplitude \( D_{\text{coeff}} \) and Krook coefficient. The collision operator is applied every \( \text{nbstep}_{\text{coll}} \) iterations but refreshed every \( \text{nbrefresh}_{\text{coll}} \) iterations.

\[ A_s \]

\[ Z_s \]

\[ 1/\rho_s \]

\[ τ \]

\[ Ω_{cs} \Delta t \]

\[ Ω_{cs} \Delta t_{\text{diag}} \]

\[ D_{\text{coeff}} \]

\[ \text{Krook}_{\text{coeff}} \]

Table 6: Modified parameters according to the scaling law for the three simulations used in invariance tests. The same invariance exercise was performed for a smaller reference case: \( ρ_s = 1/75, N_r = 128, N_θ = 128 \) and a full torus (all the other simulation parameters being identical to the ones described in Tables 4 and 5). The two others simulations defined for comparison follow the same rules as previously which corresponds to a \( ρ_s = 1/18.37 \) for the Tungsten case. The relative error is also of the order of \( 10^{−15} \) when comparing Hydrogen and Helium but of \( 10^{−6} \) when considering the Tungsten. The two cases \( ρ_s = 1/75 \) for Hydrogen and Helium will be good candidates as non-regression tests. Regarding the Tungsten simulations more work could be done. The first idea would be to consider a \( ρ_s = 1/75 \) plasma for Tungsten, but this would imply a reference case with \( ρ_s \sim 1/306.205 \) for the Hydrogen (i.e a mesh of the order of 68 billion of points with \( (N_r, N_θ, N_φ, N_{n_{Gi}}, N_{n}) = (512, 512, 128, 128, 16) \)). Such a simulation is feasible but
extremely time consuming for a verification test (several hours on 8192 cores). It would not be possible on the development cluster Poincare (92 nodes of 8 Sandy Bridge E5-2670 bi-processors (2.60GHz)) at IDRIS french supercomputing center that has been used for all tests presented above (performed on 512 cores).

4.3. Usual linear and non-linear gyrokinetic benchmarks

Together with the verification invariance test described above, several other benchmarks have already been performed in the past to validate the code. The so-called “Rosenbluth-Hinton test” (RH) [102] –which became an essential test for gyrokinetic codes to check the validity of zonal flows and Geodesic-Acoustic-Modes (GAMs) treatment– was verified in 2008 [58]. Linear and non-linear benchmarks were successfully achieved [58] with the classical Cyclone DIII-D base case (CBC) [43]. Non-linear benchmarks were also performed [57] with the global PIC code ORB5 [77]. We also participated to the European turbulence code benchmarking effort [48]. Finally, flux-driven simulations have been compared [106] between GYSELA, ORB5 and XGC1 [25, 26]. In the following we present only the tests which have been recently investigated again with the new multi-species version of the code to be added as non-regression tests, namely the RH test and the linear CBC test. These tests have been re-designed with the objective of being sufficiently small to be automatically launched on the continuous integration platform (JENKINS) at a reasonable frequency with the aim of recovering the predicted results. Here are described the results of our investigation. This work is still in progress and we plan to continue decreasing the CPU time consumption of such tests in the future.

4.3.1. Rosenbluth-Hinton test as a non-regression test

This test consists in computing the linear evolution of the zonal flow component \( \phi_{00} \) for an initial electrostatic perturbation. In practice it corresponds to initializing the code with the distribution function \( \bar{F}_s = \bar{F}_{s,eq}(1 + \epsilon \sin(2\pi(r - r_{\text{min}})/2L_r)) \) with a perturbation amplitude \( \epsilon \) equal to \( 10^{-3} \). This initial state leads to the development of GAMs which correspond to the \( (m,n) = (0,0) \) mode coupled to sidebands \( (m,n) = (\pm 1,0) \) as a result of toroidal coupling. These GAMs are Landau-damped because of the finite
poloidal wavenumber of the sideband while the zonal flows relax towards a residual value which has been analytically predicted in the case of large aspect ratio and small \( \rho_s \) as \( \phi_{th}(t) = \phi_{th}(t_0) \times A_r \) with \( A_r = 1/(1 + 1.6q^2/\sqrt{r/R}) \), \( r \) being the minor radius of the considered magnetic surface, \( R \) and \( q \) corresponding respectively to the major radius and the safety factor on this surface of interest. Eight simulations (detailed in Table 8) were performed varying both time and phase space discretizations. For all simulations (see Table 7 for common parameters), the safety factor profile is constant \( q(r) = 1.9 \forall r \in [r_{min}, r_{max}] \). The density and temperature profiles are quasi-constant with \( \kappa_{T_s} = \kappa_{n_s} = 10^{-7} \). For a complete description of the case see the associated input data file of the code (Figure I.20) in Appendix AppendixI. Time evolution of the \((0,0)\) mode is plotted in Figure 9 for the highest discretized simulation (simm 1 in Table 8). The theoretically predicted residual value \( A_r \) is found to be recovered up to 20\% in this case. A much better agreement can be found by decreasing the \( \rho_s \) value of the simulation as noticed by Biancalani (see figure 4 in [8]). One of the possible reasons is that the radial boundaries may have a stronger impact at large \( \rho_s \) values. A detailed analysis of both the transient GAM oscillation and damping on the one hand, and of the residual agreement can be found by decreasing the \( \rho_s \) value of the simulation as noticed by Biancalani (see figure 4 in [8]).

\[
\begin{align*}
\rho_s & \quad A_r & \quad Z_s & \quad R_0/a & \quad r_{int}/a & \quad r_{ext}/a & \quad \text{Torus} & \quad \text{nb}_{vth0} & \quad \mu_{\text{max}} \\
1/100 & \quad 1 & \quad 1 & \quad 2.78 & \quad 0.2 & \quad 0.8 & \quad 1 & \quad 7 & \quad 12 \\
q_1 & \quad q_2 & \quad q_3 & \quad \tau_{\text{peak}}/a & \quad \kappa_{n_s} & \quad \Delta_{n_s} & \quad \kappa_{T_s} & \quad \Delta_{T_s} & \quad T_1/T_c \\
1.9 & \quad 0 & \quad 0 & \quad 0.5 & \quad 1e^{-7} & \quad 0.2 & \quad 1e^{-7} & \quad 0.2 & \quad 1 \\
\end{align*}
\]

Table 7: Common parameters for Rosenbluth-Hinton test. The velocity phase space is defined by \( -nb_{vth0}v_{T_s} \leq \nu_{\text{GAM}} \leq nb_{vth0}v_{T_s} \) and \( 0 \leq \mu \leq \mu_{\text{max}} T_0/B_0 \). Torus indicates the fraction of the torus simulated. The safety factor radial profile is defined as \( q(r) = q_1 + q_2 \exp(q_3 \log(r/a)) \). The radial density profile is defined by its gradient as \( d \log n_{s_0}(r)/dr = -\kappa_{n_s} \cosh^{-2} \left( (r - \tau_{\text{peak}}/a)/\Delta_{n_s} \right) \). The same analytical expression is used for the temperature with \( \kappa_{T_s} \) and \( \Delta_{T_s} \).
Figure 9: Comparison of the residual value of the $(0, 0)$ mode $\phi_0(r_{\text{peak}}, t)/\phi_0(r_{\text{peak}}, t = 0) = A_r$ where $A_r = 1/(1 + 1.6 q^2/\sqrt{r/R}) = 0.06825$ is given by Rosenbluth-Hinton theory. This result corresponds to simulation 1 described in Table 8.

### Table 8: Rosenbluth-Hinton test: Values of numerical damping $\gamma_{\text{GAM}}$ and frequency $\omega_{\text{GAM}}$ for 8 GyseLa simulations varying according to 5D mesh size ($N_r, N_\theta, N_\phi, N_{vG}, N_{\mu}$) and time discretization.

<table>
<thead>
<tr>
<th>simu.</th>
<th>$N_r$</th>
<th>$N_\theta$</th>
<th>$N_\phi$</th>
<th>$N_{vG}$</th>
<th>$N_{\mu}$</th>
<th>$\Omega_{G0}\Delta t$</th>
<th>$\gamma_{\text{GAM}}$</th>
<th>$\omega_{\text{GAM}}$</th>
<th>CPU time (monoproc)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>128</td>
<td>128</td>
<td>32</td>
<td>128</td>
<td>16</td>
<td>2.</td>
<td>2.46459$10^{-4}$</td>
<td>9.11061$10^{-3}$</td>
<td>7219. h.</td>
</tr>
<tr>
<td>2</td>
<td>128</td>
<td>128</td>
<td>32</td>
<td>128</td>
<td>16</td>
<td>5.</td>
<td>2.46459$10^{-4}$</td>
<td>9.11061$10^{-3}$</td>
<td>2921. h.</td>
</tr>
<tr>
<td>3</td>
<td>64</td>
<td>64</td>
<td>16</td>
<td>128</td>
<td>8</td>
<td>5.</td>
<td>2.46459$10^{-4}$</td>
<td>9.11061$10^{-3}$</td>
<td>159. h.</td>
</tr>
<tr>
<td>4</td>
<td>64</td>
<td>64</td>
<td>16</td>
<td>128</td>
<td>4</td>
<td>5.</td>
<td>2.63456$10^{-4}$</td>
<td>9.73893$10^{-3}$</td>
<td>79. h.</td>
</tr>
<tr>
<td>5</td>
<td>64</td>
<td>64</td>
<td>16</td>
<td>128</td>
<td>8</td>
<td>10.</td>
<td>2.46459$10^{-4}$</td>
<td>9.11061$10^{-3}$</td>
<td>41. h.</td>
</tr>
<tr>
<td>6</td>
<td>64</td>
<td>64</td>
<td>16</td>
<td>64</td>
<td>8</td>
<td>10.</td>
<td>2.54957$10^{-4}$</td>
<td>9.42477$10^{-3}$</td>
<td>83. h.</td>
</tr>
<tr>
<td>7</td>
<td>64</td>
<td>64</td>
<td>16</td>
<td>128</td>
<td>8</td>
<td>25.</td>
<td>2.46459$10^{-4}$</td>
<td>9.11061$10^{-3}$</td>
<td>37. h.</td>
</tr>
<tr>
<td>8</td>
<td>64</td>
<td>64</td>
<td>16</td>
<td>128</td>
<td>8</td>
<td>50.</td>
<td>2.54957$10^{-4}$</td>
<td>9.42477$10^{-3}$</td>
<td>21. h.</td>
</tr>
</tbody>
</table>

4.3.2. Linear benchmark based on Cyclone DIII-D case for global codes

For the present benchmark, the considered physical parameters are the same as the ones defined in Lapillone’s paper [81] corresponding to the standard linear Cyclone base case (CBC) [43]. The circular concentric magnetic equilibrium is defined with an aspect ratio of $R/a = 2.78$ and a safety factor profile $q(r) = 0.86 - 0.16r/a + 2.52(r/a)^2$. This corresponds to a local safety factor $q(r_{\text{peak}}) = 1.4$ and a local magnetic shear $s(r_{\text{peak}}) = 0.8$ with $r_{\text{peak}} = r_{\text{min}} + 0.5(r_{\text{max}} - r_{\text{min}})$ and $s(r) = (r/q)dq/dr$. The initial density and temperature profiles are defined with the radial form function $f(r) = \exp(-\kappa_\rho \Delta r_\rho \tanh((r - r_{\text{peak}})\Delta r_\rho))$. The $\kappa_\rho$ and $\Delta r_\rho$ parameters are chosen to obtain peaked profiles at the middle of the radial box $r_{\text{peak}}$ with
The radial density profile is defined by its gradient as \( \frac{d \log n_s}{dr} \). For the 6 cases which were run for |\( n_s \)| ranging from 5 to 30 (see Table 10) the discretization of the velocity space was kept constant (\( N_s = 128 \) but doubled in both poloidal and toroidal directions for |\( n_s \)| ≥ 25. This corresponds to a mesh of 8 billions of points for the smallest poloidal mode numbers but of 32 billions of points for the largest. The time step \( \Delta t \) is equal to 40./\( \Omega_G \), for |\( n_s \)| ≥ 15 and divided by a factor 4. The linear growth rate is estimated by a linear fit of the exponential growth of \( \int \phi^2 \, d\mathbf{x} \, d\mathbf{v} \) during the linear phase. An example of this exponential growth is plotted at the top of Figure 11 for the smallest \( k_p \rho_s = |\mathbf{n}| q(r_{peak}) / \rho_s / r_{peak} = 0.078 \). So the linear fit is performed on a time interval \( \Omega_G [t_{init}, t_{end}] \) depending on the duration of this linear phase. This interval is taken equal to [5000, 50000] for \( k_p \rho_s = 0.078 \) and [4000, 18000] for the others. The frequency is estimated on the same time intervals by spectra analysis. Both have been compared with GENE results [81] and show a very good agreement (see Figure 10). This work is in progress in the framework of the Eurofusion project [9]. The smallest runs were performed on Poincaré machine on 512 cores (with \( (p_r, p_p, p_\mu, N_{\text{thread}}) = (2, 4, 16, 4) \)) with a CPU time cost of 9k hours/monoprocessor (for 1300 iterations). The two largest cases ran on Turing (IDRIS-France) BlueGene machine using 32768 processors (\( (p_r, p_p, p_\mu, N_{\text{thread}}) = (8, 4, 16, 64) \)) with an expensive CPU time cost of 1.7 millions of monoprocessor hours for 2000 iterations (equivalent to 1.3 10^9 h. on an INTEL machine as Curie (CCRT-France)). Such simulations are definitely more expensive for a global full-f code than for \( \delta f \) flux-tube codes. This explains why the two last points performed by GENE for \( k_p \rho_s > 0.5 \) (Fig. 10) have not been simulated with GYSELA. Nevertheless, the global aspect can give access to useful information like the time evolution of the radial structure of poloidal modes (for toroidal mode |\( n_s \)|=5). This radial structure is plotted at 4 different times on Figure 11, going from initialization to the end of the linear phase. It shows that the linear phase, i.e. the exponential growth, starts as soon as the global eigenmode –characterized by a single \( n \) and several \( m \) mode numbers– acquires its radial structure. The process of mode reorganization and the main physical parameters involved (as \( q \) profile, 1/\( L_T \) shape or ...) would be interesting to investigate in more detail.
Table 10: CBC results: Linear growth rate $\gamma$ and frequencies $\omega$ estimated in the time interval $\Omega_{C_0}[t_{\text{init}},t_{\text{end}}]$ for 6 different toroidal mode numbers $n$. The poloidal wave number $k_{\theta}$ is computed as $k_{\theta} = \frac{|n| q(r_{\text{peak}})}{r_{\text{peak}}} \rho_{\ast} = \frac{|n|}{2 \times 180}$, $r_{\text{peak}}$ being the middle of the radial box. These results are compared to GENE results in Figure 10.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$m_{\text{max}}$</th>
<th>$N_0$</th>
<th>$N_\rho$</th>
<th>$k_{\theta} \rho_{\ast}$</th>
<th>$\Omega_{C_0} \Delta t$</th>
<th>$\Omega_{C_0}[t_{\text{init}},t_{\text{end}}]$</th>
<th>$\gamma$</th>
<th>$\text{err}_{\gamma}$</th>
<th>$\omega$</th>
</tr>
</thead>
<tbody>
<tr>
<td>−5</td>
<td>7</td>
<td>256</td>
<td>128</td>
<td>0.078</td>
<td>40.</td>
<td>[5000, 50000]</td>
<td>0.045</td>
<td>0.00258</td>
<td>0.106</td>
</tr>
<tr>
<td>−10</td>
<td>14</td>
<td>256</td>
<td>128</td>
<td>0.156</td>
<td>40.</td>
<td>[4000, 18000]</td>
<td>0.158</td>
<td>0.0069</td>
<td>0.238</td>
</tr>
<tr>
<td>−15</td>
<td>21</td>
<td>256</td>
<td>128</td>
<td>0.233</td>
<td>40.</td>
<td>[4000, 18000]</td>
<td>0.217</td>
<td>0.00123</td>
<td>0.476</td>
</tr>
<tr>
<td>−20</td>
<td>25</td>
<td>512</td>
<td>256</td>
<td>0.311</td>
<td>10.</td>
<td>[4000, 18000]</td>
<td>0.257</td>
<td>0.00127</td>
<td>0.636</td>
</tr>
<tr>
<td>−25</td>
<td>35</td>
<td>512</td>
<td>256</td>
<td>0.389</td>
<td>10.</td>
<td>[4000, 18000]</td>
<td>0.271</td>
<td>0.00468</td>
<td>0.827</td>
</tr>
<tr>
<td>−30</td>
<td>42</td>
<td>512</td>
<td>256</td>
<td>0.467</td>
<td>10.</td>
<td>[4000, 18000]</td>
<td>0.253</td>
<td>0.002</td>
<td>0.978</td>
</tr>
</tbody>
</table>

Figure 10: Benchmark between GYSELA and GENE codes for the Cyclone base Case for 6 different $k_{\theta} \rho_{\ast}$ values: (Left) Linear growth rate (plotted with the error bar defined in Table 10), (Right) Linear frequency. Private communication [9].

Figure 11: Radial structure of the electrostatic potential for $(m,n) = (7,5)$ and a bandwidth of $m \pm 3$ modes, at different times: (i) initial time, (ii) beginning of the linear phase $t = 7200 \Omega_{C_0}^{-1}$, (iii) $t = 36000 \Omega_{C_0}^{-1}$ and (iv) end of the linear phase $t = 50400 \Omega_{C_0}^{-1}$. 

34
5. Local conservation laws

Verifying adequate conservation laws is an essential step in providing a correct description of mean flows. Since the controversy on the suitability of gyrokinetic codes for describing toroidal momentum transport [98, 99] a specific effort has been devoted to address conservation equations both in the context of a reduced gyrofluid model [17] or gyrokinetic field theory [110, 18]. As shown in this section, it is also possible [2] to derive local conservation equations for density, energy and toroidal momentum from the gyrokinetic electrostatic model (5)-(12) implemented in the GYSELA code (more detailed calculations can be found in Abiteboul’s PhD [1]). The radial force balance is presented in section 5.1 while the conservation laws are detailed for: (i) charge density in section 5.2, (ii) energy in section 5.3 and toroidal momentum in section 5.4.

As explained in [2], the force balance equation added to the energy and toroidal momentum conservation equations ensures a self-consistent treatment of the radial electric field and flows. We present for each property an example of numerical results, (see Figures 12 to 15) for a simulation close to the non-linear CBC simulation proposed in [81] for $\rho_*$ = 1/180. It corresponds to a collisionless simulation, with Dirichlet boundary conditions and without diffusion or Krook operators. Contrary to the linear case, there is no filtering on the toroidal modes. The initial perturbation is defined as $p(r, \theta, v_G) = \epsilon \sum_{m=1}^{m_{\text{max}}} \sum_{n=1}^{n_{\text{max}}} \cos(m\theta + n\varphi + \delta_{mn}) f_{\text{eq}}(r, \theta, v_G)/(m_{\text{max}} n_{\text{max}})$ for $m_{\text{max}} = 28$ and $n_{\text{max}} = 20$ and random values for $\delta_{mn}$ phases. The perturbation amplitude $\epsilon$ is equal to $10^{-6}$. The equilibrium distribution function $f_{\text{eq}}$ is a local Maxwellian function given by eq.(19). The other numerical parameters of this simulation are summarized in Table 11.

The conservation equations are derived at second order in $\rho_*$. Accurate results (error of 2%) were already shown with the code for force balance and toroidal momentum for a simulation with $\rho_* = 1/512$ (see Figure 2 in [2]). For a larger $\rho_* = 1/180$ the results are still accurate with a relative error of 5 to 10%. As expected, boundary conditions play a more important role leading to a degradation of the accuracy outside an internal region $0.4 \leq r / a \leq 0.7$. Comparing to relative errors for all conservation properties (Figures 12 to 15), the local energy conservation (Fig. 14) is the least accurate. It was not possible to analyze the energy behaviour for smaller $\rho_*$ (as 1/300 and 1/512) because such simulations are extremely expensive (several millions of mono-processor hours) and the diagnostic was not fully implemented in large simulations until present. This will be investigated in more details on the upcoming big simulations. Complete flux driven simulations with source terms and collisions have also been analyzed (but not presented here). Even for large $\rho_* = 1/150$, the force balance equation, local charge density and toroidal momentum are conserved with less than 10% error even at times when turbulence is fully developed. The fact that this requirement is not met for energy is still under investigation.

5.1. Radial force balance equation

In the fluid description, the radial electric field and the flows are related via the force balance equation. It was verified that this relation holds also in gyrokinetics [38, 51]. Indeed the radial force balance can

<table>
<thead>
<tr>
<th>$N_e$</th>
<th>$N_\varphi$</th>
<th>$N_{\varpi}$</th>
<th>$N_\mu$</th>
<th>$n_{\text{bth0}}$</th>
<th>$\mu_{\text{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>256</td>
<td>256</td>
<td>128</td>
<td>64</td>
<td>16</td>
<td>7</td>
</tr>
<tr>
<td>$\rho_*$</td>
<td>$R_0/a$</td>
<td>$r_{\text{int}}/a$</td>
<td>$r_{\text{ext}}/a$</td>
<td>Torus</td>
<td>$Z_s$</td>
</tr>
<tr>
<td>1/180</td>
<td>2.78</td>
<td>0.1</td>
<td>0.9</td>
<td>1/2</td>
<td>1</td>
</tr>
<tr>
<td>$q_1$</td>
<td>$q_2$</td>
<td>$q_3$</td>
<td>$\kappa_{\text{ns0}}$</td>
<td>$\Delta_{\text{ns0}}$</td>
<td>$\kappa_{\text{Tso}}$</td>
</tr>
<tr>
<td>0.86</td>
<td>-0.16</td>
<td>2.52</td>
<td>2.23</td>
<td>0.3</td>
<td>6.96</td>
</tr>
</tbody>
</table>

Table 11: Main parameters of the $\rho_*$ = 1/180 simulation used to check conservation law properties. The velocity phase space is defined by $-n_{\text{bth0}}v_{\text{Tso}} \leq v_G \leq n_{\text{bth0}}v_{\text{Tso}}$ and $0 \leq \mu \leq \mu_{\text{max}}T_0/B_0$. Torus indicates the fraction of the torus simulated. The safety factor radial profile is defined as $q(r) = q_1 + q_2(r/a) + q_3(r/a)^2$. The radial density profile is defined by its gradient as $d \log n_{\text{so}}(r)/dr = -\kappa_{\text{ns0}} \cosh^{-2}((r-r_0)/\Delta_{\text{ns0}})$. The same analytical expression is used for the temperature with $\kappa_{\text{Tso}}$ and $\Delta_{\text{Tso}}$. 

\[ \text{Accurate results (error of 2%) were already shown with the code for force balance and toroidal momentum for a simulation with } \rho_* = 1/512 \text{ (see Figure 2 in [2]). For a larger } \rho_* = 1/180 \text{ the results are still accurate with a relative error of 5 to 10%. As expected, boundary conditions play a more important role leading to a degradation of the accuracy outside an internal region } 0.4 \leq r / a \leq 0.7. \text{ Comparing to relative errors for all conservation properties (Figures 12 to 15), the local energy conservation (Fig. 14) is the least accurate. It was not possible to analyze the energy behaviour for smaller } \rho_* \text{ (as 1/300 and 1/512) because such simulations are extremely expensive (several millions of mono-processor hours) and the diagnostic was not fully implemented in large simulations until present. This will be investigated in more details on the upcoming big simulations. Complete flux driven simulations with source terms and collisions have also been analyzed (but not presented here). Even for large } \rho_* = 1/150, \text{ the force balance equation, local charge density and toroidal momentum are conserved with less than 10% error even at times when turbulence is fully developed. The fact that this requirement is not met for energy is still under investigation.} \]
be recovered analytically from the conventional first order gyrokinetic equations (see Appendix E in [2]), yielding the standard fluid expression

$$\frac{\partial \chi}{\partial t} + \frac{\partial P_{eq}}{\partial n_{eq} e} + \frac{B}{T} V_{||eq} = \frac{q B^2 R^2}{T^2} (\mathbf{V} \cdot \nabla \vartheta)$$

(58)

where the equilibrium density $n_{eq}$, pressure $P_{eq}$, velocity $V_{||eq}$ and potential $\phi_{eq}$ are functions of $-P_{\varphi}/e$ (which is approximately equal to $\chi$ at first order in $\rho_s$). The label of flux surfaces $\chi$ is chosen as the opposite of the poloidal flux of the magnetic field, i.e $\chi \equiv -\psi_{pol}$ with $\psi_{pol} = -(2\pi)^{-1} \int_{S} B \cdot \nabla \vartheta/|\nabla \vartheta|$. Considering the form of the magnetic field used presently in the code $B = (B_0 R_0 / (r, \theta)) [\zeta(r) \mathbf{e}_\varphi + \mathbf{e}_\theta]$ with $\zeta(r) = r / (q R_0)$, the label $\chi$ is determined by $\chi = B_0 \int_{r'} r' / q(r') \, dr'$ which implies $d\chi/dr = B_0 q/r$.

In order to check that the force balance equation is verified numerically, we compare $v_{GYS}^{\varphi}$ the poloidal velocity directly computed within Gysela –from the distribution function as the sum of the poloidal $E \times B$, curvature, gradient-B and magnetization flows– with the expected velocity from the force balance equation, namely $v_{FB}^{\varphi}$ corresponding to the left-hand side of eq.(58) (see Figure 12 (left)). This agreement was robustly verified in Gysela simulations for a wide range of parameters for temperature gradient, collisionality and normalized gyroradius $\rho_s$ [39, 36]. A precision of 2% was obtained for a ITER parameter case with $\rho_s = 1/512$ [2]. In Figure 12 (right), we recover this good agreement for the case $\rho_s = 1/180$ described in Table 11. Indeed, even for a larger $\rho_s = 1/180$, the relative error –defined here as the difference normalized to the quadratic mean– is smaller than 0.08 at time $t = 20040 \Omega_{ci}$ which corresponds to the beginning of the non-linear phase.

![Figure 12: Left: Numerical test of the radial force balance equation (58) at time $t = 20040 \Omega_{ci}$, comparing the poloidal velocity $v_{GYS}^{\varphi}$ directly computed in the code and $v_{FB}^{\varphi}$ the sum of the three contributions $-E \times B$, $\nabla \vartheta/n_e$ and $v_{\varphi} B_\varphi$. Right: Relative error between both. The parameters used for this simulation are summarized in Table 11.](image)

5.2. Local charge density conservation

To obtain local conservation equations, we perform integrations of the conservative form of the gyrokinetic equation (5) over the velocity space and over the magnetic flux surfaces, i.e over $d\tau^* \equiv J_x d\theta d\varphi d^3\mathbf{v}$ with $d^3\mathbf{v} = J_\varphi \, d\varphi d^2\mathbf{v}$ where $J_\varphi = 1/(\mathbf{B} \cdot \nabla \vartheta)$ is the Jacobian in real space and $J_\varphi = 2\pi B_\parallel^*/m_s$ is the Jacobian in gyro-center velocity-space.

We consider a radial region outside the buffer region and without source terms, i.e we consider the following simplified version of the conservative Boltzmann equation (5)

$$\frac{dF_s}{dt} = \frac{\partial F_s}{\partial t} + \frac{1}{B_\parallel^*} \nabla \mathbf{\cdot} \left( \mathbf{\nabla} \cdot (\nabla \vartheta) \right) \left( \frac{2\pi B_\parallel^* F_s}{m_s} \right) = C(F_s)$$

(59)
where \( z = (\chi, \theta, \varphi, v_G, \mu_s) \) and \( \dot{z} = d_t z \). We integrate eq.(59) in a small phase-space volume between two surfaces \( \chi \) and \( \chi + \delta \chi \) and apply a divergence theorem. Using the collision operator conservation property \( \int C(F_s) d\tau^* = 0 \) and summing over all species, this leads directly to the usual expression for local transport of charge density

\[
\partial_t \bar{\rho} + \partial_\chi J^\chi = 0
\]

where \( \bar{\rho} \) is the charge density and \( J^\chi \) is the radial current of gyrocenters:

\[
\bar{\rho} = \sum_{\text{species}} q_s \int d\tau^* \bar{F}_s \quad \text{and} \quad J^\chi = \sum_{\text{species}} q_s \int d\tau^* (\dot{z} \cdot \nabla \chi) \bar{F}_s
\]

Figure 13 shows that equation (61) is numerically satisfied with an error of less than 1%. Part of this error is due to the fact that the time derivative \( \partial_t \bar{\rho} \) is computed from post-processed data (\( \bar{\rho} \) is not saved at each time step \( \Delta t \) but at each diagnostic time step \( \Delta t_{\text{diag}} = 12 \Delta t \) for this simulation). Another error source is the fact that the terms \( \nabla_\perp \cdot (B^*_s d_t \chi_G) \) and \( \partial_\varphi (B^*_s d_t \varphi) \) are neglected in the splitting algorithm (see section 3.1).

**5.3. Local energy conservation**

Similarly, a conservation equation for the total energy can be derived by multiplying the gyrokinetic equation (59) by the gyrocenter Hamiltonian, which reads

\[
\bar{H}_s = \frac{1}{2} m_s v^2_{G\parallel} + \mu_s B + q_s \bar{U} \quad \text{with} \quad \bar{U} = J_0 \cdot U
\]

Using the fact that our collision operator is constructed to conserve the total Hamiltonian and integrating over \( d\tau^* \) yields

\[
\partial_t \int d\tau^* \left( \frac{1}{2} m_s v^2_{G\parallel} + \mu_s B \right) \bar{F}_s + \int d\tau^* q_s \bar{U} \partial_\chi \bar{F}_s + \int d\tau^* \bar{H}_s \frac{1}{B^*_s} \nabla_\perp \cdot (\dot{z} B^*_s \bar{F}_s) = 0
\]

Then for each species \( s \), equation (63) can be rewritten as an equation for the radial energy transport as

\[
\partial_t E_{K_s} + \partial_\chi Q_s = W_s
\]
where $E_{K_s}$ corresponds to the kinetic energy of the gyrocenters for the species; the term $Q_s$ is equivalent to a radial flux of energy and the right-hand side $W_s$ appears as a kinetic energy source

$$E_{K_s} \equiv \int d\tau^* \left( \frac{1}{2} m_s v_{||}^2 + \mu_s B \right) \bar{F}_s \ ; \ Q_s = \int d\tau^* \bar{H}_s (\hat{z} \cdot \nabla \chi) \bar{F}_s \ ; \ W_s = -q_s \int d\tau^* \bar{U} \partial_t \bar{F}_s$$

The term $W_s$ corresponds to an exchange of energy between a given species and the turbulence, and is generally referred to as turbulent heating [91, 122, 66, 121, 52]. It corresponds to a transfer of energy between particles and the electromagnetic field. The numerical computation of this term is expensive because it requires saving the 5D distribution function of each species at two successive time steps which corresponds to a large amount of memory. So to obtain a local conservation equation with no source term, we consider the total energy by summing (64) over all species. Then,

$$\partial_t E_K + \partial_\chi Q = W \quad \text{with} \quad E_K = \sum_s E_{K_s} \ , \ Q = \sum_s Q_s \quad \text{and} \quad W = \sum_s W_s \quad (65)$$

The term $W$ is decomposed into two parts as

$$\sum_s q_s \int d\tau^* \bar{U} \partial_t \bar{F}_s = \sum_s q_s \int d\tau^* U \left( J_{0_s} \cdot \partial_t \bar{F}_s \right) + \sum_s q_s \int d\tau^* \left\{ \left( J_{0_s} \cdot \bar{U} \right) \partial_t \bar{F}_s - U \left( J_{0_s} \cdot \partial_t \bar{F}_s \right) \right\} \quad (66)$$

As detailed in Appendix AppendixJ.2 the first term in eq.(J.5) can be expressed in function of the potential energy $E_p$ as

$$\sum_{\text{species}} q_s \int d\tau^* U \left( J_{0_s} \cdot \partial_t \bar{F}_s \right) \equiv \partial_t E_p \quad \text{with} \quad E_p \equiv \frac{1}{2} \sum_{\text{species}} q_s \int d\tau^* U \left( J_{0_s} \cdot \bar{F}_s \right)$$

Besides, the second term $T_E = \sum_s q_s \int d\tau^* \left\{ \left( J_{0_s} \cdot \bar{U} \right) \partial_t \bar{F}_s - U \left( J_{0_s} \cdot \partial_t \bar{F}_s \right) \right\}$ in eq.(J.5) corresponds to a polarization term, due to the difference between particles and gyro-center densities. As a remark, considering that the gyroaverage operator $J_{0_s}$ is a self-adjoint operator, this term vanishes when integrating over the whole phase-space volume. It is indeed the divergence of a flux in the local conservation equation. To express this term explicitly as a flux contribution, let us consider the low wavenumber approximation of the gyroaverage operator used in the code, i.e. the Padé approximation $J_{0_s} \simeq 1 + \frac{1}{2} \nabla \cdot \left( \frac{m_s \mu_s}{q_s B} \nabla_{\perp} \right)$ (see section 2.5.1 for more details). Using this approximation, $T_E$ can be expressed for each species as

$$T_E = \frac{m_s}{2q_s} \partial_\chi \left[ \int d\tau^* \partial_t \bar{F}_s \frac{\mu_s}{B} \nabla \chi \cdot \nabla_{\perp} \bar{U} - \int d\tau^* U \frac{\mu_s}{B} \nabla \chi \cdot \nabla_{\perp} (\partial_t \bar{F}_s) \right]$$

It can also be expressed in a more compact form using the gyrocenter perpendicular stress $P_s \perp = \int d^3 v \bar{F}_s \mu_s B$. This leads to the following conservation equation summed over all species

$$\partial_t (E_K + E_p) + \partial_\chi (Q + Q_{pot} + Q_{pol}) = 0 \quad (67)$$

38
where $E_K = \sum_s E_{Ks}$, $E_p = \sum_s E_{ps}$, $Q = \sum_s Q_s$, $Q_{pot} = \sum_s Q_{spot}$ and $Q_{pol} = \sum_s Q_{s,pol}$ with

$$E_{Ks} = \int d\tau^* \mathcal{E}_s \bar{F}_s \quad \text{with} \quad \mathcal{E}_s = \frac{1}{2} m_s v_s^2 + \mu_s B$$

$$E_{ps} = \frac{q_s}{2} \int d\tau^* U(J_0_s \cdot \bar{F}_s)$$

$$Q_s = \int d\tau^* \mathcal{E}_s (\bar{z} \cdot \nabla \chi) \bar{F}_s$$

$$Q_{s,pot} = \int d\tau^* (J_0_s \cdot U)(\bar{z} \cdot \nabla \chi) \bar{F}_s$$

$$Q_{s,pol} = \frac{m_s}{2 q_s} \int J_x d\theta d\phi 1 B^2 \left\{ \delta P_s,\perp \nabla U \cdot \nabla \chi - \nabla \left( \delta P_s,\perp \nabla \chi \right) \right\}$$

We recall that, although the polarization term is necessarily the divergence of a flux term, the expression for $Q_{pol}$ given here is not exact as it relies on an approximation of the gyroaverage operator. $Q_s$ corresponds to the energy flux of species $s$ while $Q_{s,pot}$ corresponds to the flux due to the electric potential. For the numerical results presented in Figure 14 the term $Q_{s,pol}$ is not taken into account. Indeed, it requires 3D values of $U$ and $P_{s,\perp}$ which were not saved for this simulation because this possibility has been more recently implemented in the code. Analyzing recent non-linear flux-driven simulations where this computation is available shows that the assumption $Q_{s,pot} \approx 0$ is justified. Even in regimes where turbulence is well developed the term $Q_{s,pot}$ stays sufficiently small to have no impact on the local energy conservation law.

The numerical validation of equation (67) is performed on flux surface average quantities. The comparison between $\langle \partial_t Q_s \rangle_{FS}$ and $d\langle E \rangle_{FS}/dt$ with $Q = Q_s + Q_{s,pot} + Q_{s,pol}$ and $E = E_{Ks} + E_{ps}$ defined by eqs.(67)-(72) are plotted at Figure 14 (right) showing that this energy conservation derives from the compensation of different complex radial profiles.

### 5.4. Local toroidal momentum conservation

Formally, the derivation of a conservation law for toroidal angular momentum is very similar to that for energy in the previous section. The general idea is to multiply the conservative gyrokinetic equation (59) by an invariant of motion. For the energy, this invariant was the gyro-center Hamiltonian $H_s$ given by eq.(62). For this conservation law, let us consider the gyrocenter toroidal canonical momentum $P_{\phi}$ defined as

$$P_{\phi} = m_s u_{\phi} - q_s \chi$$
with the definition \( u_\varphi = (I/B)\mathbf{v}_{\varphi||} = R^2 \mathbf{b} \cdot \nabla \varphi \mathbf{v}_{\varphi||} \). \( P_\varphi \) is an exact invariant of the \textit{unperturbed gyrocenter motion} described by the Hamiltonian \( H_{\text{eq}} = \frac{1}{2} m_s v_{\varphi||}^2 + \mu_s B + U_{\text{eq}} \), which corresponds to collisionless motion in a fully axisymmetric system. \( U_{\text{eq}} \) is the gyroaverage of the equilibrium electric potential \( U_{\text{eq}} \) independent on the toroidal angle. When axisymmetry is broken, which can occur for instance due to turbulence or magnetic field ripple, \( P_\varphi \) is no longer a motion invariant. In particular, when the electric potential becomes non-axisymmetric, the evolution of \( P_\varphi \) is governed by the equation \( \partial_t P_\varphi = - q_s \partial_\varphi \dot{U} \). This result can be obtained by using the expression of the gyrokinetic Poisson brackets in the gyro-center coordinates \([\cdot, \cdot]_{\text{GC}} \) (see [16], equation (150))

\[
[X, Y]_{\text{GC}} = \frac{B^*_s}{m_s B^*_||} \cdot (\nabla X \partial_{v_{\varphi||}} Y - \partial_{v_{\varphi||}} X \nabla Y) - \frac{b}{q_s B^*_||} \cdot (\nabla X \times \nabla Y)
\]

(74)

Details of the calculation are presented in Appendix AppendixJ.3. From the definition of \( P_\varphi \) given by eq.(73), we define the local toroidal angular momentum as

\[
\mathcal{L}_\varphi = \sum_{\text{species}} m_s \int \mathrm{d} \tau^* v_{\varphi} \dot{F}_s
\]

(75)

Note that \( \mathcal{L}_\varphi \) is the momentum for \textit{gyrocenters}, which differs from the particle momentum by terms of order \( \mathcal{O}(\rho^2) \). As said before, in order to derive a local conservation equation for \( \mathcal{L}_\varphi \), we multiply the conservative gyrokinetic equation (59) by \( P_\varphi \) and integrate over all variables other than \( \chi \), leading to

\[
\int \mathrm{d} \tau^* P_\varphi \frac{\partial \dot{F}_s}{\partial t} + \int \mathrm{d} \tau^* P_\varphi \frac{1}{B^*_||} \nabla \mathbf{z} \cdot (\dot{\mathbf{z}} B^*_|| \dot{F}_s) = \int \mathrm{d} \tau^* P_\varphi \mathcal{C}(\dot{F}_s)
\]

Using the conservation properties of the collision operator \( \int \mathrm{d} \tau^* P_\varphi \mathcal{C}(\dot{F}_s) = 0 \) and integrating by parts the second term, then

\[
\partial_t \left( \int \mathrm{d} \tau^* P_\varphi \dot{F}_s \right) - \int \mathrm{d} \tau^* \dot{F}_s \frac{\partial P_\varphi}{\partial t} - \int \mathrm{d} \tau^* \dot{F}_s \frac{1}{B^*_||} \nabla \mathbf{z} \cdot (\dot{\mathbf{z}} B^*_|| P_\varphi) + \partial_\chi \int \mathrm{d} \tau^* (\dot{\mathbf{z}} \cdot \nabla \chi) \dot{F}_s P_\varphi = 0
\]

Finally, using the fact that \( \partial_t P_\varphi = - q_s \partial_\varphi \dot{U} \), we obtain for each species \( s \),

\[
m_s \partial_t \int \mathrm{d} \tau^* v_{\varphi} \dot{F}_s - q_s \int \mathrm{d} \tau^* \chi \partial_\varphi \dot{F}_s + q_s \int \mathrm{d} \tau^* \dot{F}_s \partial_\varphi \dot{U} + \partial_\chi \int \mathrm{d} \tau^* (\dot{\mathbf{z}} \cdot \nabla \chi) \dot{F}_s P_\varphi = 0
\]

(76)

Summing over all species, this leads to

\[
\partial_t \left( \sum_s m_s \int \mathrm{d} \tau^* v_{\varphi} \dot{F}_s \right) - \chi \sum_s q_s \int \mathrm{d} \tau^* \partial_\varphi \dot{F}_s + \sum_s q_s \int \mathrm{d} \tau^* \dot{F}_s \partial_\varphi \dot{U} + \partial_\chi \left( q_s \sum_s \chi \int \mathrm{d} \tau^* (\dot{\mathbf{z}} \cdot \nabla \chi) \dot{F}_s \right) = 0
\]

Using the local particle conservation eq.(60), the second term can be written as \( \chi \partial_\chi J^x \). We also identify the last term as \( - \partial_\chi (\chi J^x) \). Then the conservation equation of the local toroidal momentum \( \mathcal{L}_\varphi = \sum_s m_s \int \mathrm{d} \tau^* v_{\varphi} \dot{F}_s \) reads

\[
\partial_t \mathcal{L}_\varphi + \partial_\chi \Pi^x_\varphi + T^x_{\text{pol}} = J^x
\]

(77)

where

\[
\Pi^x_\varphi = \sum_{\text{species}} m_s \int \mathrm{d} \tau^* v_{\varphi} u_{\varphi} v^x_{\varphi||} ; \quad T^x_{\text{pol}} = \sum_{\text{species}} q_s \int \mathrm{d} \tau^* F_s \partial_\varphi \dot{U} ; \quad J^x = \sum_{\text{species}} q_s \int \mathrm{d} \tau^* v^x_{\varphi||} \dot{F}_s
\]
Equation (77) is an exact equation for the transport of gyrocenter toroidal momentum, in the sense that it was obtained directly from the gyrokinetic model, with no specific assumptions regarding the ordering. The term $\tau_{\text{pol}}^\chi$ corresponds to a polarization flux of momentum term.

Figure 15: Numerical test for local toroidal momentum conservation for simulation detailed in Table 11. All the quantities are flux-surface averaged. Left: Comparison between $\partial_t L_\phi$ and $J^\chi - \partial_t \Pi^\chi + \tau_{\text{pol}}^\chi$ which must be equal according to eq.(77). Middle: Relative error (defined here as the difference normalized to the quadratic mean). Right: Detailed contributions of neoclassical and turbulent parts.

The numerical results obtained with the code are presented in Figure (15) with the contribution of the different terms of equation (77) (at the right) averaged on the flux surface. The relative error plotted at the middle shows an accuracy better than 10%.

6. Conclusion

A complete description of the multi-ion species version of the 5D non-linear gyrokinetic code GYSELA has been presented. Adding to its global full-$f$ character its original semi-Lagrangian scheme, the GYSELA code is unique. It can tackle the complex problem of ion turbulence self-organization (in adiabatic electrons and electrostatic limits), a special care having been taken to fully consistently model the interplay between turbulence and collisions as well as the interplay between all possible scales in the problem, with no assumption of scale separation. Code verification has been successfully performed with a permanent concern for improving its reliability, namely (i) comparison with analytical theory; (ii) benchmarks with other gyrokinetic codes; (iii) invariance tests to check the correct implementation of mass and charge parameters; and (iv) conservation law properties of the semi-Lagrangian method. Particular attention was recently paid to ensure the traceability of the code for the two main developments under progress; the implementation of kinetic electrons and taking into account of a more realistic magnetic configuration. The code is scalable up to actual parallel machine sizes.

Acknowledgments

The authors acknowledge the fruitful and lively discussions held during the Festival de Théorie in Aix-en-Provence. This work was supported by ANR grant GYPSI ANR-10-BLAN-941 and G8-Exascale action NuFUSE project. The simulations were performed not only with national HPC resources of CCRT, CINES and IDRIS under GENCi (Grand Equipement National de Calcul Intensif) allocations but also with Helios/IFERC machine dedicated to Fusion at Rokkasho in Japan and with PRACE (Partnership for Advanced Computing in Europe) Research infrastructure. This work, supported by the European Communities under the contract of Association between EURATOM and CEA, was carried out within the framework of the European Fusion Development Agreement. The views and opinions expressed herein do not necessarily reflect those of the European Commission. One of the author (D. Z.) is supported by the A*MIDEX project (no. 41
It explains the particular treatment performed to overcome the problem of the flux surface average term \( \langle \hat{\phi} \rangle_{FS} \) in Fourier space. The previous equation (F.11) can be written as
\[
\mathcal{L} \hat{\phi} + \frac{1}{Z_0^2 T_e(\hat{r})} \left[ \hat{\phi} - \lambda(\hat{\phi})_{FS} \right] = \rho(\hat{r}, \theta, \varphi) \tag{A.2}
\]
where the differential operator of second order \( \mathcal{L} \) is defined as
\[
\mathcal{L} = -\frac{1}{n_{e0}(\hat{r})} \sum_s \hat{A}_s \hat{n}_s(\hat{r}) \left\{ \frac{\partial^2}{\partial r^2} + \left[ \frac{1}{\hat{r}} + \frac{1}{n_{e0}(\hat{r})} \frac{d\hat{n}_s(\hat{r})}{d\hat{r}} \right] \frac{\partial}{\partial \hat{r}} + \frac{1}{\hat{r}^2} \frac{\partial^2}{\partial \theta^2} \right\}
\]
and the right hand side reads
\[
\rho(\hat{r}, \theta, \varphi) = \frac{1}{n_{e0}(\hat{r})} \sum_s \hat{Z}_s \left[ \hat{n}_{G_s}(\hat{r}, \theta, \varphi) - \hat{n}_{G_s, eq}(\hat{r}, \theta) \right]
\]
The constant \( \lambda \) can be chosen equal to 1 (by default) or equal to 0. Let for all function \( g \), \( \langle g \rangle_{\theta, \varphi} \) being the radial function equal to
\[
( g(\hat{r}))_{\theta, \varphi} = \frac{1}{L_\theta L_\varphi} \iint g(\hat{r}, \theta, \varphi) \, d\theta \, d\varphi
\]
By applying the integration \( \frac{1}{L_\theta L_\varphi} \int \int d\theta \, d\varphi \) to the previous equation (A.2) and by using the fact that \( \langle \langle \hat{\phi} \rangle_{\theta, \varphi} \rangle_{FS} = \langle \hat{\phi} \rangle_{\theta, \varphi} \) then:
\[
\mathcal{L} \langle \hat{\phi} \rangle_{\theta, \varphi} + \frac{1}{Z_0^2 T_e} \left[ \langle \hat{\phi} \rangle_{\theta, \varphi} - \lambda \langle \hat{\phi} \rangle_{FS} \right] = \langle \rho \rangle_{\theta, \varphi} \tag{A.3}
\]
Let \( \hat{\Phi} \) being \( \hat{\Phi} = \hat{\Phi} - \langle \hat{\phi} \rangle_{\theta, \varphi} \) then, by subtracting (A.3) to (A.2), and by using Dirichlet boundary conditions we obtain \( \forall \lambda \in \mathbb{R}, \forall \theta \in [0, L_\theta] \) and \( \forall \varphi \in [0, L_\varphi] \):
\[
\left\{ \begin{array}{ll}
( \mathcal{L} + \frac{1}{Z_0^2 T_e} ) \hat{\Phi}(\hat{r}, \theta, \varphi) = g(\hat{r}, \theta, \varphi) & \text{with} \quad g = \rho - \langle \rho \rangle_{\theta, \varphi} & \forall \hat{r} \in [\hat{r}_{\min}, \hat{r}_{\max}] \\
\hat{\Phi}(\hat{r}_{\min}, \theta, \varphi) = \hat{\Phi}(\hat{r}_{\max}, \theta, \varphi) = 0
\end{array} \right. \tag{A.4}
\]
Besides, (A.3) can be written as
\[
\mathcal{L} \left( \langle \hat{\phi} \rangle_{\theta, \varphi} - \langle \hat{\phi} \rangle_{FS} \right) + \mathcal{L} \langle \hat{\phi} \rangle_{FS} + \frac{1}{Z_0^2 T_e} \left[ \langle \hat{\phi} \rangle_{\theta, \varphi} - \langle \hat{\phi} \rangle_{FS} \right] + \mathcal{L} \langle \hat{\phi} \rangle_{FS} - \lambda \langle \hat{\phi} \rangle_{FS} = \langle \rho \rangle_{\theta, \varphi}
\]
Then using the fact that \( \langle \hat{\phi} \rangle_{FS} = \langle \hat{\phi} \rangle_{FS} - \langle \langle \hat{\phi} \rangle_{\theta, \varphi} \rangle_{FS} \) and using Dirichlet boundary conditions, the previous equation leads to the following system:

\[
\begin{cases}
\left( \mathcal{L} + (1 - \lambda) \frac{1}{Z^2 \Gamma_c} \right) \langle \hat{\phi} \rangle_{FS} = \langle \rho \rangle_{\theta, \varphi} + \left( \mathcal{L} + \frac{1}{Z^2 \Gamma_c} \right) \langle \hat{\Phi} \rangle_{FS} \\
\langle \hat{\phi} \rangle_{FS}(\hat{r}_{\text{min}}) = \langle \hat{\phi} \rangle_{FS}(\hat{r}_{\text{max}}) = 0
\end{cases}
\]  
(A.5)

Using the definition of \( \hat{\Phi} \), then

\[
\langle \hat{\Phi} \rangle_{FS} = \langle \hat{\phi} \rangle_{FS} - \langle \langle \hat{\phi} \rangle_{\theta, \varphi} \rangle_{FS}
\]

\[
\langle \hat{\phi} \rangle_{FS} - \langle \langle \hat{\phi} \rangle_{\theta, \varphi} \rangle_{FS}
\]

because \( \langle \langle \hat{\phi} \rangle_{\theta, \varphi} \rangle_{FS} = \langle \hat{\phi} \rangle_{\theta, \varphi} \)

so \( \langle \hat{\phi} \rangle_{\theta, \varphi} = \langle \hat{\phi} \rangle_{FS} - \langle \hat{\Phi} \rangle_{FS} \)

and using the fact that \( \hat{\phi} = \hat{\Phi} + \langle \hat{\phi} \rangle_{\theta, \varphi} \), we obtain the expression of the electric potential \( \hat{\phi} \) as:

\[
\hat{\phi}(\hat{r}, \theta, \varphi) = \hat{\Phi}(\hat{r}, \theta, \varphi) - \langle \hat{\phi} \rangle_{FS}(\hat{r}) + \langle \hat{\phi} \rangle_{FS}(\hat{r})
\]

(A.6)

To summarize, the solving of the equation (F.11) can be replaced by the solving of two simpler equations (A.4) and (A.5). Indeed, the equation (A.5) is a differential equation only depending on the radial direction. Besides, in (A.4) the variable \( \varphi \) plays the role of a parameter, then the discretization of the equation can be performed by projecting in Fourier space in \( \theta \) direction and by using finite differences in the radial direction as described in the following paragraph. However, it is important to realize that the boundary conditions are not directly applied on \( \phi \) but on \( \hat{\Phi} = \hat{\phi} - \langle \hat{\phi} \rangle_{\theta, \varphi} \). So the fact to impose \( \Phi(\hat{r}_{\text{min}}) = 0 \) does not imply \( \phi(\hat{r}_{\text{min}}) = 0 \) but \( \Phi(\hat{r}_{\text{min}}) = \langle \hat{\phi} \rangle_{\theta, \varphi}(\hat{r}_{\text{min}}) \) (same remark can be done at \( \hat{r} = \hat{r}_{\text{max}} \)). Another treatment is available when \( \hat{r}_{\text{min}} \) is sufficiently close to 0 (\( \hat{r}_{\text{min}} < 10^{-2} \) in the code). Indeed, let us assume that in this case \( J_s(\hat{r}_{\text{min}}, \theta) \) is equal to a constant. Then, for all function \( g \),

\[
\langle g \rangle_{FS}(\hat{r}_{\text{min}}) = \int \frac{g J_s(\hat{r}_{\text{min}}, \theta)}{J_s(\hat{r}_{\text{min}}, \theta)} \, d\theta = \frac{1}{4\pi} \langle g \rangle_{\theta, \varphi} \quad \forall |\hat{r}_{\text{min}}| \ll 1
\]

In this case employing a Neumann boundary condition on the (0, 0) mode at the axis \( \langle \hat{\phi} \rangle_{\theta, \varphi}(\hat{r}_{\text{min}}) = 0 \) is equivalent to applying \( \partial_\theta \langle \hat{\phi} \rangle_{FS}(\hat{r}_{\text{min}}) = 0 \) in the matrix system (A.5).

Appendix A.2. Finite differences in radial direction and Fourier projections in \( \theta \) and \( \phi \)

Appendix A.2.1. Solving of the equation system (A.4)

Let \( \Phi \) and \( \rho \) be represented in terms of the Fourier expansion as \( \hat{\Phi}(\hat{r}, \theta, \varphi) = \sum_m \hat{\Phi}^m(\hat{r}, \varphi) \exp(im\theta) \) and \( \rho(\hat{r}, \theta, \varphi) = \sum_m \rho^m(\hat{r}, \varphi) \exp(im\theta) \) then the equation (A.4) can be rewritten in the wave number representation, for each poloidal mode \( m \) and for each independent value of \( \varphi \), as the following differential equation:

\[
\left( \mathcal{L}^m + \frac{1}{Z^2 \Gamma_c} \right) \hat{\Phi}^m(\hat{r}, \varphi) = \hat{\rho}^m(\hat{r}, \varphi)
\]  
(A.7)

with the operator \( \mathcal{L}^m \) defined as

\[
\mathcal{L}^m = -\frac{1}{\hat{n}_{\text{sc}}(\hat{r})} \sum_s \hat{A}_s \hat{n}_{\text{sc}}(\hat{r}) \left\{ \frac{\partial^2}{\partial \hat{r}^2} + \left[ 1 + \frac{1}{\hat{n}_{\text{sc}}(\hat{r})} \frac{d \hat{n}_{\text{sc}}(\hat{r})}{d \hat{r}} \right] \frac{\partial}{\partial \hat{r}} - \frac{m^2}{\hat{r}^2} \right\}
\]  
(A.8)

and where \( \hat{\Phi}^m \) (resp. \( \hat{\rho}^m \)) is the Fourier transform in \( \theta \) of \( \hat{\Phi} \) (resp. \( \hat{\rho} \)). Let \( N_r \) be the number of radial points and let assumes that the radial domain is defined inside \([\hat{r}_1, \hat{r}_{N_r}] \) (i.e \( \hat{r}_{\text{min}} = \hat{r}_1 \) and \( \hat{r}_{\text{max}} = \hat{r}_{N_r} \)), then up to
second order in \( \Delta \hat{r} \), the system of equations (A.4) leads to the tridiagonal \( (N_r - 2) \times (N_r - 2) \) system:

\[
\begin{pmatrix}
    l_{\hat{r}_1} & u_{\hat{r}_1} & 0 \\
    d_{\hat{r}_1} & d_{\hat{r}_1} & u_{\hat{r}_1} \\
    \vdots & \vdots & \ddots \\
    0 & l_{\hat{r}_{N_r-2}} & d_{\hat{r}_{N_r-2}} & u_{\hat{r}_{N_r-2}} \\
    0 & 0 & l_{\hat{r}_{N_r-1}} & d_{\hat{r}_{N_r-1}}
\end{pmatrix}
\begin{pmatrix}
    \Phi_1^m \\
    \Phi_2^m \\
    \Phi_3^m \\
    \vdots \\
    \Phi_{N_r-2}^m \\
    \Phi_{N_r-1}^m
\end{pmatrix} =
\begin{pmatrix}
    g_1^m \\
    g_2^m \\
    g_3^m \\
    \vdots \\
    g_{N_r-2}^m \\
    g_{N_r-1}^m
\end{pmatrix}
\]  

(A.9)

with for each \( \hat{r}_i \in [\hat{r}_1, \hat{r}_{N_r}] \)

\[
\begin{align*}
    l_{\hat{r}_i} &= -\left( \frac{\beta(\hat{r}_i)}{2\pi^2} - \frac{\alpha(\hat{r}_i)}{2\pi^2} \right) \\
    d_{\hat{r}_i} &= \beta(\hat{r}_i) \left( \frac{2}{2\pi^2} + \frac{m^2}{\pi^2} \right) + \frac{1}{Z_0 T_0(\hat{r}_i)} \quad \text{with} \quad \alpha(\hat{r}_i) = \sum_s A_s \frac{\delta_n(\hat{r}_i)}{\pi_n(\hat{r}_i)} \left( \frac{1}{\hat{r}_i} + \frac{1}{\pi_n(\hat{r}_i)} \frac{d\pi_n(\hat{r}_i)}{d\hat{r}_i} \right) \\
    u_{\hat{r}_i} &= -\left( \frac{\beta(\hat{r}_i)}{2\pi^2} + \frac{\alpha(\hat{r}_i)}{2\pi^2} \right) \\
    \Phi_i^m &= \psi_m(\hat{r}_i)
\end{align*}
\]  

(A.10)

and where \( \Phi_1^m = \Phi_{N_r}^m = 0 \).

Solving the previous matrix system (A.9) is equivalent to solve a matrix system of the form \( Ax = b \) where \( A \) is tridiagonal, with \( [d_1, \ldots, d_N] = [d_{\hat{r}_1}, d_{\hat{r}_2}, \ldots, d_{\hat{r}_{N_r-1}}] \) the diagonal, \( [l_1, \ldots, l_N] = [l_{\hat{r}_1}, \ldots, l_{\hat{r}_{N_r-1}}] \) the lower diagonal and \( [u_1, \ldots, u_N] = [u_{\hat{r}_1}, \ldots, u_{\hat{r}_{N_r-2}}] \) the upper one while the right side vector \( b \) corresponds to \( [b_1, \ldots, b_N] = [\psi_m(\hat{r}_2), \ldots, \psi_m(\hat{r}_{N_r-1})] \). The result vector \( \Phi^m \) is given by \( [\Phi_1^m, \ldots, \Phi_{N_r}^m] = [0, x_1, x_2, \ldots, x_{N_r}, 0] \).

Appendix A.2.2. Solving the equation system (A.5)

The system (A.5) can be rewritten as:

\[
\begin{pmatrix}
    \lambda + (1 - \lambda) \frac{1}{Z_0^2 T_0(\hat{r}_i)} & 0 \\
    0 & \lambda + (1 - \lambda) \frac{1}{Z_0^2 T_0(\hat{r}_i)} \\
    \vdots & \vdots \\
    0 & \lambda + (1 - \lambda) \frac{1}{Z_0^2 T_0(\hat{r}_i)}
\end{pmatrix}
\begin{pmatrix}
    \dot{\Phi} \big|_{\hat{r}_i} \\
    \dot{\Phi} \big|_{\hat{r}_{\text{min}}} \\
    \vdots \\
    \dot{\Phi} \big|_{\hat{r}_{\text{max}}}
\end{pmatrix} = \Gamma(\hat{r}_i) \quad \text{for each } \hat{r}_i \in [\hat{r}_1, \hat{r}_{N_r}]
\]

with \( \Gamma(\hat{r}_i) = (\rho)_{\theta, \varphi}(\hat{r}_i) + \left( \lambda + \frac{1}{Z_0^2 T_0(\hat{r}_i)} \right) \langle \hat{\Phi} \rangle_{\text{FS}} \) where \( (\rho)_{\theta, \varphi}(\hat{r}_i) = \frac{1}{\pi_{\text{max}}^2} \int \rho(\hat{r}_i, \theta, \varphi) \, d\theta \, d\varphi \) which is equivalent (by using the same notation than for the previous matrix system (A.9)) to:

\[
\begin{pmatrix}
    l_{\hat{r}_2} & u_{\hat{r}_2} & 0 \\
    d_{\hat{r}_2} & d_{\hat{r}_2} & u_{\hat{r}_2} \\
    \vdots & \vdots & \ddots \\
    0 & l_{\hat{r}_{N_r-2}} & d_{\hat{r}_{N_r-2}} & u_{\hat{r}_{N_r-2}} \\
    0 & 0 & l_{\hat{r}_{N_r-1}} & d_{\hat{r}_{N_r-1}}
\end{pmatrix}
\begin{pmatrix}
    \langle \phi \rangle_{\text{FS}}(\hat{r}_2) \\
    \langle \phi \rangle_{\text{FS}}(\hat{r}_3) \\
    \vdots \\
    \langle \phi \rangle_{\text{FS}}(\hat{r}_{N_r-2}) \\
    \langle \phi \rangle_{\text{FS}}(\hat{r}_{N_r-1})
\end{pmatrix} =
\begin{pmatrix}
    \Gamma(\hat{r}_2) \\
    \Gamma(\hat{r}_3) \\
    \vdots \\
    \Gamma(\hat{r}_{N_r-2}) \\
    \Gamma(\hat{r}_{N_r-1})
\end{pmatrix}
\]  

(A.11)

with for each \( \hat{r}_i \in [\hat{r}_2, \hat{r}_{N_r-1}] \)

\[
\begin{align*}
    l_{\hat{r}_i} &= -\left( \frac{\beta(\hat{r}_i)}{2\pi^2} - \frac{\alpha(\hat{r}_i)}{2\pi^2} \right) \\
    d_{\hat{r}_i} &= \frac{2}{2\pi^2} \beta(\hat{r}_i) \left( 1 - \lambda \right) \frac{1}{Z_0^2 T_0(\hat{r}_i)} \\
    u_{\hat{r}_i} &= -\left( \frac{\beta(\hat{r}_i)}{2\pi^2} + \frac{\alpha(\hat{r}_i)}{2\pi^2} \right) \\
    \Gamma(\hat{r}_i) &= (\rho)_{\theta, \varphi}(\hat{r}_i) + \left( \lambda + \frac{1}{Z_0^2 T_0(\hat{r}_i)} \right)
\end{align*}
\]
where $\alpha(\hat{r}_i)$ and $\beta(\hat{r}_i)$ are defined by (A.10). Let us notice that the super-diagonal $(u_{\hat{r}_i})_{i=2,\ldots,N_r-2}$ and the lower-diagonal $(l_{\hat{r}_i})_{i=3,\ldots,N_r-1}$ are the same than the ones in the previous matrix system (A.9), while the diagonal can be deduced from the previous diagonal of the poloidal mode $m = 0$ (i.e. $(d_{\hat{r}_i}^m)_{i=2,\ldots,N_r-1}$ for $m = 0$) by the relation $d_{\hat{r}_i}^0 = d_{\hat{r}_i}^0 - \lambda/(2Z_0T_e(\hat{r}_i))$.

**Appendix A.2.3. Global algorithm for the quasi-neutrality solver**

Then the different steps for solving (A.2) and obtaining $\hat{\phi}$ are the following:

(i) Compute and save $\rho$.
(ii) Solve (A.4) to obtain $\hat{\Phi}$ and save the 3D array $\hat{\Phi}$.
(iii) Compute $\langle \hat{\Phi} \rangle_{FS}$ and save this 1D array.
(iv) Compute the RHS of (A.5), i.e $\langle \hat{\rho} \rangle_\theta \varphi + \left( \mathcal{L} + \frac{1}{Z_0^2T_e} \right) \langle \hat{\phi} \rangle_{FS}$ and save this 1D array.
(v) Solve (A.5) to obtain $\langle \hat{\phi} \rangle_{FS}$ and save it, and finally
(vi) Compute $\hat{\phi}(\hat{r}, \theta, \varphi)$ by using (A.6).

**Appendix B. Numerical implementation of the collision operator in GYSELA**

In this appendix, the simplified expression of the Lorentz-type operator which is used in GYSELA is detailed. The expression of this collision operator (including the perpendicular direction, which is not yet implemented in GYSELA) is

$$\mathcal{C}(\hat{F}_s) = \frac{1}{B_{\|s}} \partial_{v_G\|} \left\{ B_{\|s} D_{\parallel s} \hat{F}_{M_s} \partial_{v_G\|} \left( \frac{\hat{F}_s}{\hat{F}_{M_s}} \right) \right\} + \frac{1}{B_{\|s}} \partial_\mu \left\{ B_{\|s} D_{\parallel s} \hat{F}_{M_s} \frac{1}{B_\|} \partial_\mu \left( \frac{\hat{F}_s}{\hat{F}_{M_s}} \right) \right\}$$

(B.1)

where $B_{\|s}^*(r, \theta, v_G\|) = B(r, \theta) + v_G\| b \cdot \nabla \times b$ is the Jacobian of the guiding-center coordinates, and $\hat{F}_{M_s}$ is the following shifted Maxwellian distribution

$$\hat{F}_{M_s} = \frac{n_{\|s}}{(2\pi T_{s,\text{coll}}/m_\|)^{3/2}} \exp \left( -\frac{m_\| (v_G\| - V_{\|,s,\text{coll}})^2}{2T_{s,\text{coll}}} - \frac{\mu_s B}{T_{s,\text{coll}}} \right)$$

(B.2)

with the mean temperature $T_{s,\text{coll}} = T_{s,\text{coll}}(r, \theta, \varphi)$ and the mean velocity $V_{\|,s,\text{coll}} = V_{\|,s,\text{coll}}(r, \theta, \varphi)$. $\hat{F}_{M_s}$ is such that $\mathcal{C}(\hat{F}_s) = 0$. The collision term $D_{\parallel s}$ is defined by equation (25). The expression of $T_{s,\text{coll}}$ and $V_{\|,s,\text{coll}}$ are constrained such that the collision operator is momentum and energy preserving. These calculations and expressions are detailed in a first paragraph **Appendix B.1** while the numerical implementation based on a semi-implicit second order Crank-Nicolson scheme is described in the second one **Appendix B.2**.

**Appendix B.1. Expressions of the mean temperature $T_{s,\text{coll}}$ and mean velocity $V_{\|,s,\text{coll}}$ for the collision operator**

While this simplified collision operator obviously conserves the number of particles, the profiles $T_{s,\text{coll}}(r, \theta, \varphi)$ and $V_{\|,s,\text{coll}}(r, \theta, \varphi)$ must be chosen so that the operator is also compatible with the conservations of momentum and energy

$$\int d^3v m_\| v_G\| \mathcal{C}(\hat{F}_s) = 0$$

(B.3)

$$\int d^3v \left( \mu_s B + \frac{1}{2}m_\| v_G\|^2 \right) \mathcal{C}(\hat{F}_s) = 0$$

(B.4)

where $J_c = 2\pi B_{\|s}^*/m_\|$ is the jacobian in velocity space and $d^3v = J_c d\mu_s dv_{G\|}$ is the space velocity element. Let us first consider only the contribution of collisions in the parallel direction. We use the expression of
the collision operator \( (B.1) \) and integrate eq.\((B.3)\) by parts twice in the variable \( vG_\parallel \)

\[
(B.3) = 2\pi \int d\mu_s \, dvG_\parallel \left\{ \frac{F_s}{F_{M_s}} B_{\parallel s}^* D_{\parallel s} \partial_{vG_\parallel} \tilde{F}_{M_s} + \frac{F_s}{F_{M_s}} \partial_{vG_\parallel} (B_{\parallel s}^* D_{\parallel s}) \right\}
\]

Let us define the average \( \langle \cdot \rangle \) as \( \langle \cdot \rangle = \int d^3v \cdot F_s \) and use the fact that \( \partial_{vG_\parallel} \tilde{F}_{M_s} = -\tilde{F}_{M_s} \frac{m_s (vG_\parallel - vG_{\parallel,\text{coll}})}{T_{s,\text{coll}}} \), then

\[
(B.3) = \frac{m_s}{T_{s,\text{coll}}} \left\{ V_{[s,\text{coll}]} \langle m_s D_{\parallel s} \rangle - \langle m_s D_{\parallel s} vG_\parallel \rangle + T_{s,\text{coll}} \left\{ \frac{1}{B_{\parallel s}^*} \partial_{vG_\parallel} (B_{\parallel s}^* D_{\parallel s}) \right\} \right\}
\]  \( \text{(B.5)} \)

We perform similar operations on the contribution to eq.\((B.4)\) of collisions in the parallel direction:

\[
(B.4) = \frac{m_s}{T_{s,\text{coll}}} \left\{ V_{[s,\text{coll}]} \langle m_s D_{\parallel s} vG_\parallel \rangle - \langle m_s D_{\parallel s} vG_\parallel ^2 \rangle + T_{s,\text{coll}} \left\{ \frac{1}{B_{\parallel s}^*} \partial_{vG_\parallel} (B_{\parallel s}^* D_{\parallel s}) \right\} \right\}
\]  \( \text{(B.6)} \)

For collisions in the perpendicular direction, eq.\((B.3)\) is trivially verified. We perform two integrations by parts in the variable \( \mu_s \) for eq.\((B.4)\),

\[
(B.4)_{\perp} = \frac{2\pi}{m_s} \int dvG_\parallel \, d\mu_s \frac{F_s}{F_{M_s}} \frac{1}{B} \partial_\mu (B_{\parallel s}^* D_{\perp s}) \tilde{F}_{M_s}
\]

Then using the fact that \( \partial_\mu \tilde{F}_{M_s} = -\tilde{F}_{M_s} \frac{B}{T_{s,\text{coll}}} \),

\[
(B.4)_{\perp} = \frac{1}{T_{s,\text{coll}}} \left\{ \frac{1}{BB_{\parallel s}^*} \partial_\mu (B_{\parallel s}^* D_{\perp s}) \right\} T_{s,\text{coll}} - \langle D_{\perp s} \rangle \right\}
\]  \( \text{(B.7)} \)

Using equations \((B.5)\), \((B.6)\) and \((B.7)\), the conservation equations \((B.3)\) and \((B.4)\) form a linear system in \( T_{s,\text{coll}}(r, \theta, \varphi) \) and \( V_{[s,\text{coll}]}(r, \theta, \varphi) \) as follows

\[
V_{[s,\text{coll}]} \langle m_s D_{\parallel s} \rangle + T_{s,\text{coll}} \left\{ \frac{1}{B_{\parallel s}^*} \partial_{vG_\parallel} (B_{\parallel s}^* D_{\parallel s}) \right\} = \langle m_s D_{\parallel s} vG_\parallel \rangle
\]

\[
V_{[s,\text{coll}]} \langle m_s^2 D_{\parallel s} vG_\parallel ^2 \rangle + T_{s,\text{coll}} \left\{ \frac{m_s}{B_{\parallel s}^*} \partial_{vG_\parallel} (B_{\parallel s}^* D_{\parallel s}) \right\} + \langle m_s D_{\parallel s} vG_\parallel ^2 \rangle + \langle D_{\perp s} \rangle = \langle m_s^2 D_{\parallel s} vG_\parallel ^2 \rangle + \langle D_{\perp s} \rangle
\]

Solving this system, we find that the conservation constraints are verified if the profiles \( V_{[s,\text{coll}]}(r, \theta, \varphi) \) and \( T_{s,\text{coll}}(r, \theta, \varphi) \) for the collision operator are defined as follows

\[
m_s PV_{[s,\text{coll}]} = \left\{ \frac{m_s}{B_{\parallel s}^*} \partial_{vG_\parallel} (B_{\parallel s}^* D_{\parallel s} vG_\parallel) \right\} + \left\{ \frac{1}{B_{\parallel s}^*} \partial_\mu (B_{\parallel s}^* D_{\parallel s}) \right\} \langle m_s D_{\parallel s} vG_\parallel \rangle
\]

\[
PT_{s,\text{coll}} = \langle D_{\parallel s} \rangle \left\{ \frac{m_s^2}{B_{\parallel s}^*} \partial_{vG_\parallel} (B_{\parallel s}^* D_{\parallel s}) \right\} + \langle m_s^2 D_{\parallel s} vG_\parallel ^2 \rangle + \langle D_{\perp s} \rangle - \langle m_s D_{\parallel s} vG_\parallel \rangle ^2
\]

where

\[
P = \langle D_{\parallel s} \rangle \left\{ \frac{m_s}{B_{\parallel s}^*} \partial_{vG_\parallel} (B_{\parallel s}^* D_{\parallel s} vG_\parallel) \right\} + \left\{ \frac{1}{B_{\parallel s}^*} \partial_\mu (B_{\parallel s}^* D_{\parallel s}) \right\} - \langle m_s D_{\parallel s} vG_\parallel \rangle \left\{ \frac{1}{B_{\parallel s}^*} \partial_{vG_\parallel} (B_{\parallel s}^* D_{\parallel s}) \right\}
\]

46
and as a reminder $\langle \cdot \rangle = \int J_v \, d\mu_s \, dv_G \cdot F_s$ with $J_v = 2\pi B^*_v/m_v$.

Let us remind that only collisions in the parallel direction are taken into account at the moment in the code, i.e $D_{\perp s} = 0$ in above expressions. Then, let us define the 5 following integrals:

\[ I_0 = \langle D_{||s} \rangle \quad ; \quad I_1 = \langle m_s D_{||s} v_G^2 \rangle \quad ; \quad I_2 = \langle m_s^2 D_{||s} v_G^4 \rangle \]

\[ I_3 = \left( \frac{1}{B^*_s} \partial_{v_G} \left( B^*_s \, D_{||s} \right) \right) \quad \text{and} \quad I_4 = \left( \frac{m_s}{B^*_s} \partial_{v_G} \left( B^*_s \, v_G \, D_{||s} \right) \right) \]

Therefore the mean velocity and mean temperature can be simply expressed as:

\[ m_s V_{||s,\text{coll}} = P^{-1} \left( I_1 - I_2 \times I_3 \right), \quad T_{s,\text{coll}} = P^{-1} \left( I_0 - I_2 - I_1^2 \right) \quad \text{and} \quad P = I_0 \times I_4 - I_1 \times I_3 \]

**Appendix B.2. Crank-Nicolson scheme for collision operator solving**

In the following, the semi-implicit second-order Crank-Nicolson scheme which is implemented in GYSELA to take into account the collisional effects as

\[ \partial_t F_s = C_{ss}(F_s) = \frac{1}{B^*_s} \partial_{v_G} \left( B^*_s \, D_{||s} \, F_s - B^*_s \, V \, F \right) \quad (B.8) \]

is detailed. The diffusion term $D$ is defined as $D = D_{||s}(r,v) = A(r) \left( \frac{2(vG)}{2vG} \right)$ with $A(r) = 3\frac{v^2}{2} \frac{\nu}{q} \nu_s$, while the expression of the drag term $V$ is given by $V = V_{||s}(r,v) = -\frac{v_{||s}}{v_{||s,\text{coll}}} \frac{D_{||s}(r,v)}{\nu_s}$. Classically, let us write: $f^n_j \equiv F_s(r, \theta, \varphi, v_G, j, t_n)$, where subscript $j \in \{0, \ldots, N\}$ refers to the discretised index in parallel velocity space and superscript $n$ refers to the time. For each value of $\mu_s$, each term in eq.(B.8) reads:

\[ B^*_s \, \partial_t f \rightarrow B^*_s \frac{f_j^{n+1} - f_j^n}{\Delta t} \quad (B.9) \]

\[ \partial_{v_G} \left( B^*_s \, D_{||s} \, f \right) = \frac{1}{2\Delta v_G^2} \left\{ B^*_{||s,j+\frac{1}{2}} D_{||j + \frac{1}{2}} \left[ \frac{f_{j+1}^{n+1} - f_{j+1}^n + f_{j+1}^{n+1} - f_{j+1}^n}{\Delta t} \right] ight. \]

\[ \left. - B^*_{||s,j-\frac{1}{2}} D_{||j - \frac{1}{2}} \left[ f_{j+1}^{n+1} - f_{j}^{n+1} + f_{j}^n - f_{j-1}^{n+1} \right] \right\} \quad (B.10) \]

\[ \partial_{v_G} \left( B^*_s \, v f \right) = \frac{-B^*_{||s,j+1} V_{||j+1} \left[ f_{||j+1}^{n+1} + f_{||j+1}^n \right] - B^*_{||s,j-1} V_{||j-1} \left[ f_{||j-1}^{n+1} + f_{||j-1}^n \right]}{4\Delta v_G^2} \quad (B.11) \]

where we use the fact that $B^*_v$ is linear in $v_G$, i.e the expression $B^*_{||s,j+\frac{1}{2}} = \frac{1}{2} \left( B^*_{||s,j} + B^*_{||s,j+1} \right)$ is exact. Then, using expressions (B.9), (B.10) and (B.11) in the equation (B.8) gives, for all $j = 0, \ldots, N$:

\[ A_j f_j^{n+1} + B_j f_j^{n+1} + C_j f_j^{n+1} = -A_j f_j^{n-1} + \left( \frac{2B^*_{||s,j}}{\Delta t} - B_j \right) f_j^n - C_j f_j^{n+1} \quad (B.12) \]

where the coefficients $A_j, B_j$ and $C_j$ are defined as:

\[ A_j = -\frac{\alpha_j - \frac{1}{2}}{2\Delta v_G^2} - \frac{\beta_j - \frac{1}{2}}{4\Delta v_G^2} ; \quad B_j = B^*_{||s,j} \frac{\Delta t}{\Delta v_G^2} + \frac{\alpha_j + \frac{1}{2} + \alpha_j - \frac{1}{2}}{2\Delta v_G^2} ; \quad C_j = -\frac{\alpha_j + \frac{1}{2}}{2\Delta v_G^2} + \frac{\beta_j + \frac{1}{2}}{4\Delta v_G^2} \quad (B.13) \]

with

\[ \alpha_j = B^*_{||s,j} D_{||j} \quad \text{and} \quad \beta_j = B^*_{||s,j} V_{||j} \quad (B.14) \]

47
To compute the terms $\alpha_{-1/2}$, $\beta_{-1}$, $\alpha_{N+1/2}$ and $\beta_{N+1}$ required for $A_0$, $B_0$, $B_N$ and $C_N$ calculations, the boundary conditions $\partial^2 D/\partial v_G^2 = 0$ and $\partial^2 V/\partial v_G^2 = 0$ are imposed, i.e.:

\[
D_{-\frac{1}{2}} = 2D_0 - D_{\frac{1}{2}} \quad \text{and} \quad \mathcal{V}_{-1} = 2\mathcal{V}_0 - \mathcal{V}_1 \quad \text{(B.15)}
\]

\[
D_{N+\frac{1}{2}} = 2D_N - D_{N-\frac{1}{2}} \quad \text{and} \quad \mathcal{V}_{N+1} = 2\mathcal{V}_N - \mathcal{V}_{N-1} \quad \text{(B.16)}
\]

and the fact that $B_{||}^*$ is linear in $v_G$ is used, i.e.:

\[
B_{||,s,-\frac{1}{2}}^* = \frac{1}{2} \left( 3B_{||,s,0}^* - B_{||,s,\frac{1}{2}}^* \right) \quad \text{;} \quad B_{||,s,-1}^* = 2B_{||,s,0}^* - B_{||,s,1}^* \quad \text{(B.17)}
\]

\[
B_{||,s,N+\frac{1}{2}}^* = \frac{1}{2} \left( 3B_{||,s,N}^* - B_{||,s,N-\frac{1}{2}}^* \right) \quad \text{;} \quad B_{||,s,N+1}^* = 2B_{||,s,N}^* - B_{||,s,N-1}^* \quad \text{(B.18)}
\]

Finally, let us consider that the second derivative of $f$ is vanishing at domain boundaries, i.e. $f_{-1} = 2f_0 - f_1$ and $f_{N+1} = 2f_N - f_{N-1}$ then

\[
(B_0 + 2A_0)f_{0}^{n+1} + (C_0 - A_0)f_{1}^{n+1} = \frac{2B_{||,s,0}^*}{\Delta t}f_{0}^n - (C_0 - A_0)f_{1}^n
\]

\[
(A_N - C_N)f_{N-1}^{n+1} + (B_N + 2A_N)f_{N}^{n+1} = -(A_N - C_N)f_{N-1}^n + \frac{2B_{||,s,N}^*}{\Delta t} - B_N - 2C_N\big]f_{N}^n
\]

The system can be rewritten in the compact tridiagonal form:

\[
\begin{pmatrix}
B_0 + 2A_0 & C_0 - A_0 \\
A_1 & B_1 & \ddots \\
& \ddots & \ddots & C_{N-1} \\
A_N - C_N & B_N + 2A_N & \end{pmatrix}
\begin{pmatrix}
f_{0}^{n+1} \\
f_{1}^{n+1} \\
\vdots \\
f_{N}^{n+1}
\end{pmatrix}
= \begin{pmatrix}
R_{0}^n \\
\vdots \\
\vdots \\
R_{N}^n
\end{pmatrix}
\quad \text{(B.19)}
\]

and $(R_{0}^n, \ldots, R_{N}^n)^t = \mathcal{R} (f_{0}^n, \ldots, f_{N}^n)^t$ with the matrix $\mathcal{R}$ defined as

\[
\mathcal{R} =
\begin{pmatrix}
-2A_0 - B_0 + \frac{2B_{||,s,0}^*}{\Delta t} & -(C_0 - A_0) \\
-A_1 & -B_1 + \frac{2B_{||,s,1}^*}{\Delta t} & \ddots \\
& \ddots & \ddots & -C_{N-1} \\
-(A_N - C_N) & -2C_N - B_N + \frac{2B_{||,s,N}^*}{\Delta t}
\end{pmatrix}
\quad \text{(B.20)}
\]

The tridiagonal system (B.19)-(B.20) is solved by using a modified Thomas algorithm.

**Appendix C. Expressions for simplified collisional transfer between two species**

**Appendix C.1. Conservation properties of collisional energy transfer**

Let us consider the energy exchange between two species defined by equations (34)-(39), i.e. $\frac{d\bar{F}_s}{dt} = C_{ss}(\bar{\mathcal{E}}_{ss})$ and $\frac{d\bar{F}_{ss}^{e}}{dt} = C_{s's's}(\bar{\mathcal{E}}_{s's})$ with

\[
C_{ss}(\bar{\mathcal{E}}_{ss}) \equiv -\frac{\xi_{ss}^{s}}{(2\pi m_s T_{\text{mean}})^{3/2}} \frac{\Delta T_{ss}}{T_{\text{mean}}} \left( \frac{\mathcal{E}_{ss}}{T_{\text{mean}}} - \frac{3}{2} \right) \exp \left( -\frac{\mathcal{E}_{ss}}{T_{\text{mean}}} \right)
\]

\[
C_{s's's}(\bar{\mathcal{E}}_{s's}) \equiv -\frac{\xi_{s's}^{s'}}{(2\pi m_{s'} T_{\text{mean}})^{3/2}} \frac{\Delta T_{s's}}{T_{\text{mean}}} \left( \frac{\mathcal{E}_{s's}}{T_{\text{mean}}} - \frac{3}{2} \right) \exp \left( -\frac{\mathcal{E}_{s's}}{T_{\text{mean}}} \right)
\]
with \( E_{ss} \), \( E_{s} \) defined by equation (38) and \( T_{\text{mean}} \), \( \Delta T_{ss} \) defined by equation (36). Let us consider the brackets \( \langle \cdot \rangle_{v,T} \) defined as

\[
\langle \cdot \rangle_{v,T} = \frac{1}{(2\pi T/m_s)^{3/2}} \int \int \exp \left( -\frac{v^2}{2T} - \frac{\mu B}{T} \right) \frac{2\pi}{m_s} B_\parallel^{1/2} \, \mu_s \, dv \, \mu_s \, dv \quad (C.1)
\]

Let us define \( v_\parallel = v_{G\parallel} - V_{\text{mean}} \). Then, using the fact that \( \langle 1 \rangle_{v_\parallel,T_{\text{mean}}} = 1 \) and \( \langle m_s v^2 \rangle_{v_\parallel,T_{\text{mean}}} = \langle \mu_s B \rangle_{v_\parallel,T_{\text{mean}}} = T_{\text{mean}} \), it is straightforward to show that such a collision operator conserves both particles and parallel momentum, i.e. that

\[
\langle C_{E_{ss}}^E(E_s) \rangle = \langle C_{E_{s}'}^E(E_s) \rangle = 0 \quad \text{and} \quad \langle v_{G\parallel} C_{E_{ss}}^E(E_s) \rangle + \langle v_{G\parallel} C_{E_{s}'}^E(E_s) \rangle = 0 \quad \text{where} \quad \langle \cdot \rangle = \int \cdot d^3 v
\]

The parameters \( \eta_{E_{ss}}^E \) and \( \eta_{E_{s}'}^E \) are designed such that the collisional energy exchange between species satisfy \( \langle E_s C_{E_{ss}}^E(E_s) \rangle = -\langle E_s C_{E_{s}'}^E(E_s) \rangle \). So let us compute the following integral in velocity space,

\[
\langle E_s C_{E_{ss}}^E \rangle = \int E_s C_{E_{ss}}^E \, d^3v = -\eta_{E_{ss}}^E \frac{\Delta T_{ss}}{T_{\text{mean}}} \left\langle E_s \left( \frac{E_{ss}'}{T_{\text{mean}}} - \frac{3}{2} \right) \right\rangle_{v_\parallel,T_{\text{mean}}}
\]

which is equivalent to \( \langle E_s C_{E_{ss}}^E \rangle = -\eta_{E_{ss}}^E \frac{\Delta T_{ss}}{T_{\text{mean}}} I \), where

\[
I = \left\langle \left( \frac{m_s v_\parallel^2}{2} + m_s \left( \frac{V_{\parallel}^2 - V_{\text{mean}}^2}{2} \right) - m_s \left( \bar{v}_{\parallel} + V_{\text{mean}} \right) \frac{\Delta V_{E_{ss}'}^E}{2} + \mu_s B \left( 1 - \frac{2 \mu_s B}{T_{\text{mean}}} - 3 \right) \right) \right\rangle_{v_\parallel,T_{\text{mean}}}
\]

\[
= \left\langle \left( \frac{m_s v_\parallel^2}{2} - m_s \frac{\Delta V_{E_{ss}'}^E}{2} - m_s \frac{\Delta V_{E_{ss}'}^E}{8} + \mu_s B \left( 1 - \frac{2 \mu_s B}{T_{\text{mean}}} - 3 \right) \right) \right\rangle_{v_\parallel,T_{\text{mean}}}
\]

\[
= \left\langle \left( \frac{m_s v_\parallel^2}{2} + \mu_s B \left( \frac{m_s \Delta V_{E_{ss}'}^E}{2} - m_s \frac{\Delta V_{E_{ss}'}^E}{8} - \frac{3}{2} \right) + \frac{m_s v_\parallel^2 \mu_s B}{T_{\text{mean}}} + \frac{\mu_s B^2}{T_{\text{mean}}} - m_s \Delta V_{E_{ss}'}^E \right) \right\rangle_{v_\parallel,T_{\text{mean}}}
\]

Then, using the properties, \( \langle m_s v_\parallel^4 \rangle_{v_\parallel,T_{\text{mean}}} = 3T_{\text{mean}}^2 \) and \( \langle \mu_s B^2 \rangle_{v_\parallel,T_{\text{mean}}} = 2T_{\text{mean}}^2 \) and \( \langle m_s v_\parallel^2 \mu_s \rangle_{v_\parallel,T_{\text{mean}}} = T_{\text{mean}}^2 \), the collisional energy exchange between species occurs at the following rate

\[
\langle E_s C_{E_{ss}}^E \rangle = -\eta_{E_{ss}}^E \frac{3\Delta T_{ss}'}{8T_{\text{mean}}} \left( 1 - \frac{\Delta V_{E_{ss}'}^E}{2T_{\text{mean}}} \right) \quad (C.2)
\]

Let us use the property that the energy exchange term between two species is of the form (cf. [62], p.184)

\[
W_{ss'} = -\frac{4}{\sqrt{\pi}} \frac{\gamma_{ss'}}{m_{ss'}} \frac{v_{T,s} v_{T,s'} (T_s - T_{s'})}{(v_{T,s}^2 + v_{T,s'}^2)^{3/2}} \equiv \int d^3v \frac{1}{2} m_s v^2 C_{ss'} \quad \text{with} \quad \gamma_{ss'} \sim \frac{m_s v_{T,s} v_{T,s'}}{n_{ss'}} \nu_{ss'}
\]

where \( v_{T,s} = \max(v_{T,s}, v_{T,s'}) \) and the collision frequency \( \nu_{ss'} \) is deduced from the ion-ion collision frequency \( \nu_{ss} \) as

\[
\nu_{ss'} = \left( \frac{Z_s'}{Z_s} \right)^2 \frac{n_{ss'} v_{T,s}}{n_s v_{T,s'}} \nu_{ss} \quad \text{and} \quad \nu_{ss} = \left( \frac{m_s v_{T} v_{T}^2}{qR_0} \right)^{1/2} \nu_{ss}
\]

Therefore

\[
W_{ss'} = -\frac{4}{\sqrt{\pi}} \frac{v_{T,s} v_{T,s'}}{v_{T,s}} \frac{m_s}{n_s} \Delta T_{ss'} \left( 1 + \frac{v_{T,s}^2}{v_{T,s}^2} \right)^{-3/2} \quad (C.4)
\]
Finally, by analogy between (C.2) and eq.(C.4) (considering $\Delta V_{s's'}^2/8T_{\text{mean}} \ll 1$), this provides for the effective collision frequency $\eta_{s's'}^G$ the following expression,

$$\eta_{s's'}^G = \frac{8 \varepsilon^{3/2}}{3\sqrt{\pi}} n_{s'} \frac{m_s}{m_{s'}} \left( \frac{Z_{s'}}{Z_s} \right)^2 \left( 1 + \frac{\nu_{T'}}{\nu_T} \right)^{-3/2} \sqrt{\frac{T_{s',\text{coll}}/m_s}{qR_0}} \nu_{s's'}$$

which is equivalent to the one given by equation (39).

**Appendix C.2. Conservation properties of collisional parallel momentum transfer**

Let us consider the parallel momentum transfer defined by equations (40)-(42), i.e. $d\bar{F}_s/\text{d}t = C_{s's'}^{\nu_G^G}(E_s)$ and $d\bar{F}_{s'}/\text{d}t = C_{s's'}^{\nu_G^G}(E_{s'})$ with

$$C_{s's'}^{\nu_G^G}(E_s) \equiv -\eta_{s's'}^{\nu_G^G} \Delta V_{s's'}v_G \exp \left( -\frac{E_s}{T_{\text{mean}}} \right) \quad \text{and} \quad C_{s's'}^{\nu_G^G}(E_{s'}) \equiv -\eta_{s's'}^{\nu_G^G} \Delta V_{s's'}v_G \exp \left( -\frac{E_{s'}}{T_{\text{mean}}} \right)$$

Considering this approximation, the momentum exchanges between two species reads \(^{11}\)

$$\langle m_s v_G \| C_{s's'}^{\nu_G^G}(E_s) \rangle = \int \text{d}^3v \ m_s v_G \frac{\text{d}\bar{F}_s}{\text{d}t} = -\eta_{s's'}^{\nu_G^G} \Delta V_{s's'} \left( \frac{2\pi}{m_s} \right)^{3/2} T_{\text{mean}}^{5/2} \quad \text{(C.5)}$$

$$\langle m_s v_G \| C_{s's'}^{\nu_G^G}(E_{s'}) \rangle = \int \text{d}^3v \ m_s v_G \frac{\text{d}\bar{F}_{s'}}{\text{d}t} = -\eta_{s's'}^{\nu_G^G} \Delta V_{s's'} \left( \frac{2\pi}{m_s} \right)^{3/2} T_{\text{mean}}^{5/2} \quad \text{(C.6)}$$

Therefore, the action-reaction principle $\langle m_s v_G \| C_{s's'}^{\nu_G^G} \rangle = -\langle m_s v_G \| C_{s's'}^{\nu_G^G} \rangle$, leads to the first constraint $\eta_{s's'}^{\nu_G^G} = (m_{s'}/m_s)^{3/2}\eta_{s's'}^{\nu_G^G}$. The second constraint comes from the neoclassical friction relation,

$$- \int \text{d}^3v \ m_s v_G \frac{\text{d}\bar{F}_{s'}}{\text{d}t} = m_s n_{s's'} \Delta V_{s's'} = -m_s n_{s's'} \nu_{s's'} \Delta V_{s's'}$$

then using equation (C.5) leads to

$$\eta_{s's'}^{\nu_G^G} = m_s n_{s's'} \left( \frac{2\pi}{m_s} \right)^{3/2} T_{\text{mean}}^{-5/2} \nu_{s's'}$$

Finally, using the relation (C.3) for $\nu_{s's'}$ expression, $\eta_{s's'}^{\nu_G^G}$ can be expressed as

$$\eta_{s's'}^{\nu_G^G} = n_{s's'} m_s^2 \left( \frac{\varepsilon}{2\pi} \right)^{3/2} \frac{1}{qR_0} \frac{\sqrt{T_{s',\text{coll}}/m_s}}{T_{\text{mean}}} \left( \frac{Z_{s'}}{Z_s} \right)^2 \frac{\nu_T}{\nu_{T'}} \nu_{s's'}$$

which is equivalent to equation (42).

**Appendix D. Source terms**

Focussing on the source term, the gyrokinetic equation reads:

$$\frac{\text{d}\bar{F}_s}{\text{d}t} = S_E(r, \theta, v_G \|, \mu_s) S_y(r) \quad \text{(D.1)}$$

\(^{11}\langle m_s v_G \| C_{s's'}^{\nu_G^G} \rangle = -\eta_{s's'}^{\nu_G^G} \Delta V_{s's'} \left( \frac{2\pi T_{\text{mean}}}{m_s} \right)^{3/2} \langle m_s v_G^2 \rangle_{v_G \| \text{mean}} = -\eta_{s's'}^{\nu_G^G} \Delta V_{s's'} \left( \frac{2\pi T_{\text{mean}}}{m_s} \right)^{3/2} T_{\text{mean}}\)
The prescribed radial source profile $S_r(r)$ is defined as the sum of two hyperbolic tangents

$$S_r(r) = \frac{1}{2} \left[ \tanh \left( \frac{\rho - \rho_S - 3L_S}{L_S} \right) + \tanh \left( \frac{\rho_S - 3L_S - \rho}{L_S} \right) \right]$$

where $\rho_S$ and $L_S$ are input data and $\rho = (r - r_{\text{min}})/L_r$ where $L_r$ is the length of the radial box. The energy dependent part of the source is decomposed on the basis of orthogonal Hermite and Laguerre polynomials (cf. next section Appendix D.1 as a reminder):

$$S_E(r, \theta, v_{G\parallel}, \mu_s) = \sum_{\ell=0}^{+\infty} \sum_{h=0}^{+\infty} c_{\ell h} H_h(\bar{v}_{G\parallel}) L_\ell(\bar{\mu}_s)e^{-\bar{v}_{G\parallel}^2}$$

where the $c_{\ell h}$ coefficients depend on the space coordinates only. The following definitions have been introduced:

$$\bar{\mu}_s \equiv \frac{\mu_s B}{T_{s,\text{srce}}} ; \quad \bar{v}_{G\parallel} \equiv \frac{v_{G\parallel}}{\sqrt{2T_{s,\text{srce}}/m_s}}$$

with $T_{s,\text{srce}}$ the normalized source temperature.

**Appendix D.1. Hermite and Laguerre polynomials**

The Hermite and Laguerre polynomials form the set of orthogonal basis with respect to the following scalar products:

$$\text{Laguerre } L_\ell(x) : \quad \int_0^{+\infty} L_\ell(x)L\ell' e^{-x} \, dx = \delta_{\ell\ell'} |L_\ell|^2 \quad \text{(D.5)}$$

$$\text{Hermite } H_h(x) : \quad \int_{-\infty}^{+\infty} H_h(x)H_h' e^{-x^2} \, dx = \delta_{hh'} |H_h|^2 \quad \text{(D.6)}$$

The Laguerre polynomials are normalized: $|L_\ell|^2 = 1$. The norm of the Hermite polynomials is:

$$|H_h|^2 \equiv \int_{-\infty}^{+\infty} H_h^2 e^{-x^2} \, dx = \sqrt{\pi} 2^h h! \quad \text{(D.7)}$$

The five first Laguerre and Hermite polynomials are:

| $L_0(x)$ | 1 |
| $L_1(x)$ | $-x$ |
| $L_2(x)$ | $\frac{1}{2}(2 - 4x + x^2)$ |
| $L_3(x)$ | $\frac{1}{2}(6 - 18x + 9x^2 - x^3)$ |
| $L_4(x)$ | $\frac{1}{24}(24 - 96x + 72x^2 - 16x^3 + x^4)$ |

| $H_0(x)$ | 1 |
| $H_1(x)$ | $2x$ |
| $H_2(x)$ | $-2 + 4x^2$ |
| $H_3(x)$ | $-12x + 8x^3$ |
| $H_4(x)$ | $12 - 48x^2 + 16x^4$ |

**Appendix D.2. Corresponding sources for fluid moments**

Let us derive the corresponding source of matter $S_m$, of parallel momentum $S_{v_{G\parallel}}$, of energy $S_E$ and of vorticity $S_\Omega$. With the adopted definitions,

$$B_{\parallel s} = B(1 + J_{\parallel B}v_{G\parallel}) \quad \text{and} \quad J_{\parallel B} \equiv \frac{\sqrt{m_r}}{q_s} \sqrt{\frac{2T_{s,\text{srce}}}{B^2}} \mu_0 J_{\parallel}$$

(D.9)
The parallel current \( \mu_0 J_\parallel = \mu_0 b \cdot J \) is prescribed in GYSELA. Then, the integral over the velocity space reads as follows:

\[
\int d^3v \equiv \int_{-\infty}^{+\infty} dv_{G\parallel} \int_{0}^{+\infty} \frac{2\pi B^*_s}{m_s} d\mu_s = \frac{1}{\sqrt{\pi}} \left( \frac{2\pi T_{s,\text{srce}}}{m_s} \right)^{3/2} \int_{-\infty}^{+\infty} (1 + J_{\parallel} v_{G\parallel}) d\bar{v}_{G\parallel} \int_{0}^{+\infty} d\bar{\mu}_s
\]

Notice that \( (1 + J_{\parallel} v_{G\parallel}) = H_0(\bar{v}_{G\parallel}) + \frac{J_{\parallel} B}{2} H_1(\bar{v}_{G\parallel}) \) and \( L_0(\bar{\mu}_s) = 1 \). The fluid source of matter \( S_n \) is simply \( S_n \equiv \int d^3v \; S E S_r \). Using the decomposition of \( S_E \) on the basis of orthogonal polynomials (D.3), the source of matter becomes

\[ S_n = \left( \frac{2\pi T_{s,\text{srce}}}{m_s} \right)^{3/2} \left( c_{00} + J_{\parallel} (c_{00} + 4c_{20}) \right) S_r \]  

(D.10)

The fluid source of parallel momentum \( S_{v_{G\parallel}} \) reads as follows: \( S_{v_{G\parallel}} \equiv \int d^3v \; v_{G\parallel} S E S_r \). Following the same procedure than for the density source, one finally obtains:

\[ S_{v_{G\parallel}} = 2\pi^{3/2} \left( \frac{T_{s,\text{srce}}}{m_s} \right)^2 \left[ 2c_{20} + \frac{3}{2} c_{00} - c_{01} + \frac{5}{2} J_{\parallel} (c_{10} + 6J_{\parallel} c_{30} - J_{\parallel} c_{11}) \right] S_r \]  

(D.11)

The fluid source of energy \( S_E \) is defined as follows: \( S_E \equiv \int d^3v \; \left( m_s \frac{v_{G\parallel}^2}{2} + \mu_s B \right) S E S_r \). Notice that \( (m_s v_{G\parallel}^2/2 + \mu_s B) = T_{s,\text{srce}}(\bar{v}_{G\parallel}^2 + \bar{\mu}_s) \). Again, the energy source can also be recast in terms of the \( c_{h\ell} \) coefficients:

\[ S_E = \left( \frac{2\pi T_{s,\text{srce}}}{m_s} \right)^{3/2} \left( T_{s,\text{srce}} \right)^2 2c_{20} + \frac{3}{2} c_{00} - c_{01} + \frac{5}{2} J_{\parallel} (c_{10} + 6J_{\parallel} c_{30} - J_{\parallel} c_{11}) \right] S_r \]  

(D.12)

AppendixD.2.1. Source of vorticity

The fluid source of vorticity \( S_\Omega \) is simply: \( S_\Omega \equiv \int d^3v \; J_{0,\perp}(S_E S_r) \), where \( J_{0,\perp} \) is the gyro-average operator. We use the Padé approximation:

\[ J_{0,\perp} \approx 1 + \frac{m_s}{q_s^2} \frac{\mu_s}{2B} \nabla_\perp^2 = 1 + \alpha \; \bar{\mu}_s \; \nabla_\perp^2 \quad \text{with} \quad \alpha = \frac{m_s}{q_s^2} \frac{T_{s,\text{srce}}}{2B^2} \]

Again, the vorticity source can be recast in terms of the \( c_{h\ell} \) coefficients as

\[ S_\Omega = S_n + \alpha \left( \frac{2\pi T_{s,\text{srce}}}{m_s} \right)^{3/2} \left[ \nabla_\perp^2 \left( (c_{00} - c_{01}) S_r \right) + \frac{J_{\parallel} B}{2} \nabla_\perp^2 \left( (c_{10} - c_{11}) S_r \right) \right] \]  

(D.13)

AppendixD.3. Pure sources of momentum, energy and vorticity

The expressions of \( S_n \) (eq.(D.10)), \( S_{v_{G\parallel}} \) (eq.(D.11)), \( S_E \) (eq.(D.12)) and \( S_\Omega \) (eq.(D.13)) provide the constraints on the \( c_{h\ell} \) coefficients in order to impose independently zero source of density, of momentum, of energy or of vorticity. Let us consider three cases:

- Non vanishing source of energy, with no injection of particles nor of momentum.
- Non vanishing source of momentum, with no injection of particles of nor of energy.
- Non vanishing source of vorticity, with no injection of particles, of momentum nor of energy.
These three cases are considered hereafter. Imposing zero source of matter, the fluid sources of parallel momentum, of energy and of vorticity are proportional to:

\[ S_n = 0 \Rightarrow c_{00} + J_{\parallel B} c_{10} = 0 \]  \hspace{1cm} (D.14)

\[ S_{cG1} \propto \left( 2 - J_{\parallel B}^2 \right) c_{10} + 4 J_{\parallel B} c_{20} \]  \hspace{1cm} (D.15)

\[ S_c \propto 2 c_{20} - c_{00} + 6 J_{\parallel B} c_{30} - c_{01} - J_{\parallel B} c_{11} \]  \hspace{1cm} (D.16)

\[ S_{11} \propto -\nabla_\perp^2 (J_{\parallel B} c_{10} S_r) + J_{\parallel B} \nabla_\perp^2 (c_{10} S_r) - \nabla_\perp^2 (c_{01} S_r) - J_{\parallel B} \nabla_\perp^2 (c_{11} S_r) \]  \hspace{1cm} (D.17)

**Appendix D.3.1. Pure source of energy**

Killing the fluid sources of particles, momentum and vorticity, while keeping finite the fluid source of energy, imposes eq.(D.15) and eq (D.17) to vanish. Several solutions can be envisaged. Let’s choose \( c_{30} = 0 \) and \( (c_{01} + J_{\parallel B} c_{11}) = 0 \), with \( (2 c_{20} - c_{00}) \neq 0 \). Then, the fluid source of vorticity trivially vanishes for \( c_{10} = c_{11} \). To summarize, we propose the following set of coefficients for a pure source of energy,

\[ c_{11} = c_{10} = - \frac{4 J_{\parallel B}}{2 - J_{\parallel B}^2} c_{20} ; \quad c_{00} = c_{01} = \frac{4 J_{\parallel B}^2}{2 - J_{\parallel B}^2} c_{20} ; \quad c_{30} = 0 \text{ and } c_{20} \neq 0 \]  \hspace{1cm} (D.18)

In order to inject solely energy into the system, the source term that should appear in the right hand side of the gyrokinetic equation would then take the following form:

\[
S_E = S_r (c_{00} + c_{01} L_1 + c_{10} H_1 + c_{11} H_1 L_1 + c_{20} H_2) e^{-\vec{v}_G_{1s} \cdot \vec{\mu}_s} \\
= 4 c_{20} S_r \left[ \vec{v}_G_{1s} \frac{1}{2} + \frac{J_{\parallel B}}{2 - J_{\parallel B}^2} (2 - \vec{\mu}_s)(J_{\parallel B} - 2 \vec{v}_G_{1s}) \right] e^{-\vec{v}_G_{1s} \cdot \vec{\mu}_s} \text{ due to (D.18)}
\]

while the fluid source of energy \( S_E \) would have the following magnitude:

\[
S_E = \left( \frac{2 \pi T_{s,\text{arce}}}{m_s} \right)^{3/2} T_{s,\text{arce}} (2 c_{20} - c_{00}) S_r \text{ according to (D.12) and (D.18)} \\
= 2 \left( \frac{2 \pi T_{s,\text{arce}}}{m_s} \right)^{3/2} T_{s,\text{arce}} \left( 1 - \frac{2 J_{\parallel B}^2}{2 - J_{\parallel B}^2} \right) c_{20} S_r
\]

Let us introduce the normalized intensity \( S_E^c \equiv 2 c_{20} \left( \frac{2 \pi T_{s,\text{arce}}}{m_s} \right)^{3/2} T_{s,\text{arce}} \) then the previous equality becomes

\[
S_E = \left( 1 - \frac{2 J_{\parallel B}^2}{2 - J_{\parallel B}^2} \right) S_E^c S_r
\]  \hspace{1cm} (D.19)

Then, up to small terms proportional to \( J_{\parallel B}, S_r \) —which is normalized such that its volume integral is equal to unity— provides the radial shape of the energy source, while \( S_E^c \) gives its magnitude. Finally, \( S_E \) can be expressed as:

\[
S_E = \frac{S_E^c S_r}{\sqrt{2} \left( \frac{2 \pi T_{s,\text{arce}}}{m_s} \right)^{3/2} T_{s,\text{arce}}} \left[ \vec{v}_G_{1s} \frac{1}{2} + \frac{J_{\parallel B}}{2 - J_{\parallel B}^2} (2 - \vec{\mu}_s)(2 \vec{v}_G_{1s} - J_{\parallel B}) \right] e^{-\vec{v}_G_{1s} \cdot \vec{\mu}_s}
\]  \hspace{1cm} (D.20)

with \( J_{\parallel B} \) defined by (D.9).
Appendix D.3.2. Pure parallel momentum

Killing the fluid sources of particles, energy and vorticity, while keeping finite the fluid source of parallel momentum, imposes equations (D.16)-(D.17) to vanish. Again, several options could be considered. Following the same strategy as for the energy, namely \( c_{11} = c_{10} = c_{00} = c_{01} = -J_{\parallel B} c_{10} \), then the source of energy vanishes if \( c_{30} = 0 \) and \( 2 c_{20} = c_{00} \). Consistently, in order to inject only parallel momentum, the following source is proposed

\[
S_{\nu_{\parallel |}} = \frac{S_{0}^{\nu_{\parallel |}} S_{r}}{4\pi^{3/2} \left( \frac{T_{s,\text{src}}}{m_{s}} \right)^{2}} \left[ 2\nu_{G_{\parallel s}}(2 - \bar{\mu}_{s}) - J_{\parallel B} \left( 1 + 2\nu_{G_{\parallel s}}^{2} - \bar{\mu}_{s} \right) \right] e^{-\nu_{G_{\parallel s}}^{2} - \bar{\mu}_{s}}
\]

(D.21)

where the the normalized intensity \( S_{0}^{\nu_{\parallel |}} \) is defined as \( S_{0}^{\nu_{\parallel |}} \equiv 4\pi^{3/2} \left( \frac{T_{s,\text{src}}}{m_{s}} \right)^{2} c_{10} \). The corresponding fluid source of momentum is (according to (D.11))

\[
S_{\nu_{\parallel |}} = \left( 1 - \frac{3J_{\parallel B}^{2}}{2} \right) S_{0}^{\nu_{\parallel |}} S_{r}
\]

(D.22)

Then, at leading order in \( J_{\parallel B} \), \( S_{r} \) corresponds to the radial shape of the momentum source, and \( S_{0}^{\nu_{\parallel |}} \) to its magnitude.

Appendix D.3.3. Pure source of vorticity

So as to inject vorticity only, the simplest choice appears to be: \( c_{00} = c_{10} = c_{11} = 0 \) and \( c_{20} = 0 \). Then the source of vorticity is governed by the \( c_{01} \) coefficient only: \( S_{0} \propto -\nabla_{\perp}^{2} \left( c_{01} S_{r} \right) \), while that of momentum eq.(D.15) is set to zero. The source of energy eq.(D.16) vanishes provided that \( c_{30} = (1/6J_{\parallel B}) c_{01} \). Obviously, such a constraint is invalid for those simulations performed at vanishing parallel current. Alternatively, one decides to allow for some parallel momentum injection by taking \( c_{20} = c_{01}/2 \neq 0 \) and \( c_{30} = 0 \). Then, the source term to be considered is the following

\[
S_{0} = -\frac{S_{0}^{\Omega} S_{r}}{\alpha \left( \frac{T_{s,\text{src}}}{m_{s}} \right)^{3/2}} \left[ 2\nu_{G_{\parallel s}} - \bar{\mu}_{s} \right] e^{-\nu_{G_{\parallel s}}^{2} - \bar{\mu}_{s}} \quad \text{with} \quad \alpha = \frac{m_{s} T_{s,\text{src}}}{q_{s}^{2} 2B^{2}}
\]

(D.23)

where the normalized intensity \( S_{0}^{\Omega} \) is defined by \( S_{0}^{\Omega} \equiv -c_{01} \alpha \left( \frac{2T_{s,\text{src}}}{m_{s}} \right)^{3/2} \) and the resulting fluid source of vorticity is

\[
S_{\Omega} = S_{0}^{\Omega} \nabla_{\perp}^{2} \left( S_{r} \right)
\]

(D.24)

We recall that such a source does inject some momentum as well. However, its magnitude remains small, and equal to \( -\frac{1}{3} \sqrt{\frac{2m_{s}}{T_{s}}} J_{\parallel B} S_{0}^{\Omega} S_{r} \).

Appendix D.3.4. Another possibility for the heating source

There exists a simplified version of the heating source possible in the code which is not exactly a pure source of heating. In this case the energy dependent part of the source is defined as:

\[
S_{\nu_{\parallel |}}(r, \theta, \nu_{G_{\parallel}}, \mu_{s}) = \frac{S_{0}^{\nu_{\parallel |}} S_{r}}{3\sqrt{2} \left( \frac{T_{s,\text{src}}}{m_{s}} \right)^{3/2}} \left( \frac{E}{T_{s,\text{src}}} - \frac{3}{2} \right) \exp \left( -\frac{E}{T_{s,\text{src}}} \right)
\]

(D.25)

where \( E \) the energy is equal to \( E = \frac{1}{2} m_{s} \nu_{G_{\parallel}} + \mu_{s} B \). Therefore,

\[
\int_{-\infty}^{\infty} d\nu_{G_{\parallel}} \int_{0}^{\infty} \frac{2\pi B^{*}}{m_{s}} d\mu_{s} E S_{\nu_{\parallel |}} = S_{0}^{\nu_{\parallel |}}
\]
Appendix D.4. A source of impurities

According to equation (D.13), it is clear that in this source formalism it is complicate to generate a source of matter without injecting a source of vorticity. So as first choice, we take \( c_{00} = c_{01} \) and \( c_{10} = c_{11} = 0 \), such that \( S_\Omega = S_n \). The injection of momentum and energy can be avoided by imposing in equations (D.11) and (D.12), \( c_{00} = 4c_{20} = 0 \) and \( 2c_{20} + (3/2)c_{00} - c_{01} + 6\bar{E}_C t_{30} = 0 \). Due to the previous assumptions, this leads to \( c_{30} = 0 \). Therefore, using the fact that \( c_{01} = c_{00} \) and \( c_{20} = -(1/4)c_{00} \), the matter source term take the form

\[
S_n = S_r c_{00} \left( \frac{5}{2} - \bar{\mu}_s - \bar{v}_{G||}^2 \right) e^{-\bar{v}_{G||}^2 - \bar{\mu}_s}
\]

Then, according to equation (D.10) and considering the normalized matter source \( S_n^0 = (2\pi T_{s,scrce}/m_s)^{3/2}c_{00} \),

\[
S_n = \frac{S_n^0 S_r}{(2\pi T_{s,scrce}/m_s)^{3/2}} \left( \frac{5}{2} - \bar{\mu}_s - \bar{v}_{G||}^2 \right) e^{-\bar{v}_{G||}^2 - \bar{\mu}_s} \tag{D.26}
\]

Let us denote by \( S_{n,s} \) (resp. \( S_{n,s'} \)) the source of matter associated to the majority species \( s \) (resp. to the impurity species \( s' \)). The injection of impurity must be compensated by the injection of the majority species such that

\[
Z_s \int S_{n,s} J_{s,s} d\nu_{G||} d\mu_s + Z_{s'} \int S_{n,s'} J_{s',s'} d\nu_{G||} d\mu_{s'} = 0
\]

Appendix D.5. Numerical treatment of the source terms

The source terms are taken into account by solving \( d/dt \bar{F}_s = S_E + S_{vG||} + S_n + S_n \) with \( S_E \) defined by eq.(D.20), \( S_{vG||} \) by eq.(D.21), \( S_n \) by eq.(D.23) and \( S_n \) by eq.(D.26). For the following, let us use the fact that each source is of the form \( S_E(r(t), \theta(t), v_{G||}(t), \mu_s)S_r(r(t)) \). For more readability let us consider one unique source knowing that the numerical method described below can be generalized to a sum of sources. Let us integrate in time between \( t \) and \( t + \Delta t \) the equation

\[
\frac{d\bar{F}_s}{dt}(r, \theta, \varphi, v_{G||}, \mu) = S(t) \quad \text{with} \quad S(t) = S_E(r(t), \theta(t), v_{G||}(t), \mu_s)S_r(r(t))
\]

then

\[
\bar{F}_s(t + \Delta t) - \bar{F}_s(t) = \int_t^{t+\Delta t} S(t') dt'
\]

\[
S_r(r(t')) = S_r(r(t + \Delta t)) + [r(t') - r(t + \Delta t)] \frac{\partial S_r}{\partial r}(r(t + \Delta t)) + O(\Delta t^2)
\]

\[
= S_r(r(t + \Delta t)) + (t' - t + \Delta t) \frac{dr}{dt}(t + \Delta t) \frac{\partial S_r}{\partial r}(r(t + \Delta t)) + O(\Delta t^2)
\]

Besides, let us define \( S_E(t') = S_E(r(t'), \theta(t'), v_{G||}(t'), \mu_s) \), then

\[
S_E(t') = S_E(r(t + \Delta t), \theta(t'), v_{G||}(t'), \mu_s) + (t' - t - \Delta t) \frac{dr}{dt}(t + \Delta t) \frac{\partial S_E}{\partial r}(r(t + \Delta t), \theta(t + \Delta t), v_{G||}(t + \Delta t), \mu_s) + O(\Delta t^2)
\]

\[
\approx S_E(r(t + \Delta t), \theta(t + \Delta t), v_{G||}(t + \Delta t), \mu_s) + (t' - t - \Delta t) \frac{dr}{dt}(t + \Delta t) \frac{\partial S_E}{\partial r}(r(t + \Delta t), \theta(t + \Delta t), v_{G||}(t + \Delta t), \mu_s) +
\]

\[
\frac{d\theta}{dt}(t + \Delta t) \frac{\partial S_E}{\partial \theta}(r(t + \Delta t), \theta(t + \Delta t), v_{G||}(t + \Delta t), \mu_s) +
\]

\[
\frac{dv_{G||}}{dt}(t + \Delta t) \frac{\partial S_E}{\partial v_{G||}}(r(t + \Delta t), \theta(t + \Delta t), v_{G||}(t + \Delta t), \mu_s) + O(\Delta t^2)
\]

55
Therefore, the distribution function $F_s$ at time $t + \Delta t$ is given at second order in time by

$$F_s(t + \Delta t) = F_s(t) + \Delta t S_E(t + \Delta t) S_r(t + \Delta t) - \frac{\Delta t^2}{2} \left[ S_E(t + \Delta t) \frac{d}{dt} t \Delta t \frac{\partial S_r}{\partial r} |_{t+\Delta t} + S_r(t + \Delta t) \left( \frac{d}{dt} t \Delta t \frac{\partial S_r}{\partial \theta} |_{t+\Delta t} + \frac{d}{dt} t \Delta t \frac{\partial S_r}{\partial v} |_{t+\Delta t} \right) \right]$$

In the code, the radial derivatives of the sources of: (i) energy $\partial_r S_E$, (ii) momentum $\partial_r S_{vG||}$, (iii) vorticity $\partial_r S_Q$ and (iv) density $\partial_r S_n$ are computed numerically as well as the poloidal derivatives. On the other hand, the derivatives in $v_{G||}$ direction are expressed analytically as

(i) for the pure source of energy:

$$\frac{\partial S_E}{\partial v_{G||}} = \frac{S_0^r S_r}{\sqrt{2 \pi \frac{T_{s,scce}}{m_s}}} \frac{1}{\sqrt{2 T_{s,scce}}} \exp \left( -\bar{v}_{G||s}^2 - \bar{\mu}_s \right) \times$$

$$\left[ -2K_h(2 - \bar{\mu}_s) - 2\bar{v}_{G||s} \left( \frac{3}{2} - K_h(2 - \bar{\mu}_s) \left( 2\bar{v}_{G||s} - J_{||B} \right) \right) \right]$$

(ii) for the pure source of momentum:

$$\frac{\partial S_{vG||}}{\partial v_{G||}} = \frac{S_0^r v_{G||} S_r}{4 \pi \frac{T_{s,scce}}{m_s}} \frac{1}{\sqrt{2 T_{s,scce}}} \exp \left( -\bar{v}_{G||s}^2 - \bar{\mu}_s \right) \times$$

$$\left[ 2(2 - \bar{\mu}_s) \left( 1 - 2\bar{v}_{G||s}^2 \right) + 2J_{||B}\bar{v}_{G||s} \left( -1 + 2\bar{v}_{G||s}^2 - \bar{\mu}_s \right) \right]$$

(iii) for the pure source of vorticity (according to eq.(D.23)):

$$\frac{\partial S_Q}{\partial v_{G||}} = -\frac{S_0^r Q}{\alpha \left( \frac{T_{s,scce}}{m_s} \right)^{3/2}} \frac{1}{\sqrt{2 T_{s,scce}}} \exp \left( -\bar{v}_{G||s}^2 - \bar{\mu}_s \right) \left[ 4\bar{v}_{G||s} - 2\bar{v}_{G||s} \left( 2\bar{v}_{G||s} - \bar{\mu}_s \right) \right]$$

(iv) for the source of impurity (according to eq.(D.26)):

$$\frac{\partial S_n}{\partial v_{G||}} = \frac{S_0^r S_r}{4 \pi \frac{T_{s,scce}}{m_s}} \frac{1}{\sqrt{2 T_{s,scce}}} \exp \left( -\bar{v}_{G||s}^2 - \bar{\mu}_s \right) \left[ -2\bar{v}_{G||s} \left( \frac{7}{2} - \bar{\mu}_s - \bar{v}_{G||s}^2 \right) \right]$$

**AppendixE. A Crank-Nicolson scheme for diffusion terms**

In this appendix the semi-implicit Crank-Nicolson scheme, which is implemented in GYSELA to take into account a diffusion term of the form

$$\partial_t F_s = \frac{1}{r} \partial_r \left( r \chi(r) \partial_r F_s \right) \tag{E.1}$$

is detailed. Classically, let us write: \( f^n_r \equiv F_s(r, \theta, \varphi, v_{G||}, t_n) \), where subscript \( j \in \{ 0, \ldots, N \} \) refers to the discretised index in radial direction and superscript \( n \) refers to the time. Let us also define \( D \equiv r \chi(r) \) with \( \chi(r) = \chi_0 H_{\text{buff}}(r) \). The diffusion amplitude \( \chi_0 \) is modulated by a radial profile \( H_{\text{buff}} \) which is equal to 1 in the buffer region and equal to 0 elsewhere. For each value of \( \mu_s \), each term in eq.(E.1) reads for all
\[ f_i^{n+1} = f_i^n + \frac{\Delta t}{2 \Delta x^2 r_i} \left\{ \mathcal{D}_{i-\frac{1}{2}} \left[ \left( f_i^{n+1} - f_{i-1}^{n+1} \right) + \left( f_i^{n+1} - f_i^n \right) \right] - \mathcal{D}_{i+\frac{1}{2}} \left[ \left( f_i^{n+1} - f_{i+1}^{n+1} \right) + \left( f_i^n - f_i^{n-1} \right) \right] \right\} \] (E.2)

Let us introduce \( \alpha_i = \frac{\Delta t}{2 \Delta x^2 r_i} \), then

\[-\alpha_i \mathcal{D}_{i-\frac{1}{2}} f_i^{n+1} + \left[ 1 + \alpha_i \left( \mathcal{D}_{i-\frac{1}{2}} + \mathcal{D}_{i+\frac{1}{2}} \right) \right] f_i^{n+1} - \alpha_i \mathcal{D}_{i+\frac{1}{2}} f_i^{n+1} = R_i^n \]

with

\[ R_i^n = \alpha_i \mathcal{D}_{i-\frac{1}{2}} f_i^{n+1} + \left[ 1 - \alpha_i \left( \mathcal{D}_{i-\frac{1}{2}} + \mathcal{D}_{i+\frac{1}{2}} \right) \right] f_i^n + \alpha_i \mathcal{D}_{i+\frac{1}{2}} f_i^n \]

Therefore, for all \( i = 0, \ldots, N \),

\[ A_i f_i^{n+1} + B_i f_i^n + C_i f_{i+1}^{n+1} = A_i f_i^n + (2 - B_i) f_i^n - C_i f_{i+1}^n \] (E.3)

where the coefficients \( A_i, B_i \) and \( C_i \) are defined as:

\[ A_i = -\alpha_i \mathcal{D}_{i-\frac{1}{2}} ; \quad B_i = 1 + \alpha_i \left( \mathcal{D}_{i-\frac{1}{2}} + \mathcal{D}_{i+\frac{1}{2}} \right) ; \quad C_i = -\alpha_i \mathcal{D}_{i+\frac{1}{2}} \]

Using the fact that \( r_{i+1/2} = 0.5 \left( r_i + r_{i+1} \right) \), then \( \mathcal{D}_{i+\frac{1}{2}} = \frac{1}{2} \left( \mathcal{D} r_i + \mathcal{D} r_{i+1} \right) \). Concerning the boundary conditions, let us consider that the second derivative of the radial profile \( \chi \) is equal to 0 at the boundaries (i.e. \( \frac{\partial^2 \chi}{\partial r^2} \big|_{r_{\text{min}}} = \frac{\partial^2 \chi}{\partial r^2} \big|_{r_{\text{max}}} = 0 \) ), then

\[ \mathcal{D}_{-\frac{1}{2}} = \frac{1}{2} \left( r_0 - \frac{\Delta r}{2} \right) (3 \chi_0 - \chi_1) \] \[ \text{and} \quad \mathcal{D}_{N+\frac{1}{2}} = \frac{1}{2} \left( r_N + \frac{\Delta r}{2} \right) (3 \chi_N - \chi_{N-1}) \]

Besides, a non-homogeneous Dirichlet boundary condition is applied to the outer radial boundary such that \( \hat{F}_s (r = r_{\text{max}}, \theta, \varphi, \nu_G, \mu_s) = \hat{F}_{s, \text{eq}} (r = r_{\text{max}}, \theta, \nu_G, \mu_s) \). Concerning the inner radial boundary, there are two possibilities: (i) non-homogeneous Dirichlet boundary condition \( \hat{F}_s (r = r_{\text{min}}, \theta, \varphi, \nu_G, \mu_s) = \hat{F}_{s, \text{eq}} (r = r_{\text{min}}, \theta, \nu_G, \mu_s) \) in the case of thermal bath or (ii) Neumann boundary conditions (i.e. \( f_{-1} = f_1 \)) in the case of flux-driven simulation. Let us consider the boolean \( \kappa \) which is equal to 1 if Neumann boundary conditions are applied (i.e. in the case of flux driven boundary conditions) and is equal to 0 otherwise. The system can be rewritten in the compact tridiagonal form:

\[
\begin{pmatrix}
(1 - \kappa) + \kappa B_0 & \kappa (A_0 + C_0) \\
A_1 & B_1 & \ddots & \end{pmatrix}
\begin{pmatrix}
\vdots \\
A_{N-1} & B_{N-1} & C_{N-2} & \end{pmatrix}
\begin{pmatrix}
f_0^{n+1} \\
f_1^{n+1} \\
\vdots \\
f_N^{n+1} \\
\end{pmatrix} =
\begin{pmatrix}
R_0^n \\
\vdots \\
R_N^n \\
\end{pmatrix}
\] (E.4)

with

\[
\begin{pmatrix}
R_0^n \\
\vdots \\
R_N^n \\
\end{pmatrix} =
\begin{pmatrix}
(1 - \kappa) + \kappa (2 - B_0) & -\kappa (A_0 + C_0) \\
-\kappa A_1 & (2 - B_1) & \ddots & \end{pmatrix}
\begin{pmatrix}
f_0^n \\
\vdots \\
f_N^n \\
\end{pmatrix}
\]

The tridiagonal system (E.4) is solved by using a modified Thomas algorithm.
Appendix F. GYSELA normalized equations

This appendix contains the expression of all the normalized quantities and normalized equations implemented in the code GYSELA.

Appendix F.1. GYSELA normalization

Let consider \( m_0 = A_0 m_p \) (kilograms) a reference ionic mass, \( q_0 = Z_0 e \) (coulombs) a reference ionic charge, \( B_0 \) (tesla) a reference magnetic induction and \( T_0 \) (electron volts) a reference temperature. \( A_0 \) and \( Z_0 \) are the (dimensionless) mass number and charge state of the principal ionic species and \( e \) the modulus of the electron charge. These quantities are used to define the reference ion cyclotron frequency \( \Omega_{co} \), the reference thermal speed \( v_{T0} \) and the reference Larmor-radius \( \rho_0 \) as:

\[
\Omega_{co} = \frac{Z_0 \ e \ B_0}{m_0} \quad ; \quad v_{T0} = \sqrt{\frac{T_0}{m_0}} \quad ; \quad \rho_0 = \frac{v_{T0}}{\Omega_{co}} = \frac{\sqrt{T_0 m_0}}{Z_0 \ e \ B_0} \tag{F.1}
\]

Actually, the seven normalizing quantities \( (A_0, Z_0, n_0, T_0, B_0, a, R_0) \) are not completely free. Three dimensionless parameters are further specified in the input data, which provide relationships between these seven parameters. These are the aspect ratio \( A = R_0/a \), rhostar \( \rho_{s0} = \sqrt{A_0 m_p T_0/Z_0 e B_0} \) and mustar \( \nu_{s0} = e^{-3/2}(q R_0/v_{T0})_\text{coll} \sim Z_0^2 n_0 R_0/T_0^2 \) (where \( e \) and \( q \) are taken at \( r_{\text{peak}} = r_{\text{min}} + 0.5(r_{\text{max}} - r_{\text{min}}) \)). It follows that, given \( A, \rho_{s0} \) and \( \nu_{s0} \), it remains only four free normalizing quantities. Hereafter, \( \hat{X} \) refers to the dimensionless expression of the \( X \) quantity. The normalizations used in the code are summarized in the table F.12.

$$
\begin{align*}
\begin{array}{ll}
m_s &= m_0 \hat{A}_s \\
q_s &= Z_0 e \hat{Z}_s \\
l &= \rho_0 \hat{l} \\
t &= t/\Omega_{co} \\
n_s &= n_0 \hat{n}_s \\
T_s &= T_0 \hat{T}_s \\
B &= B_0 \hat{B}
\end{array}
\Rightarrow
\begin{array}{ll}
v_{G\parallel} &= v_{T0} \hat{v}_{G\parallel} \\
v_{E\times B_s} &= (v_{T0}/\sqrt{A_s}) \hat{v}_{E\times B_s} \\
v_{D_s} &= v_{T0} \hat{v}_{D_s} \\
\mu_s &= (T_0/B_0) \hat{\mu}_s \\
U &= (T_0/\{Z_0 e\}) \hat{U} \\
E &= (v_{T0} B_0) \hat{E} \\
F_s &= (n_0/v_{T0}) \hat{F}_s \\
\end{array}
\end{align*}
\]

and

$$
\begin{align*}
\chi &= \rho_0^2 \Omega_{co} \hat{\chi} \\
\nu &= \Omega_{co} \hat{\nu} \\
D_{\parallel s} &= v_{T0} \Omega_{co} \hat{D}_{\parallel s} \\
V_{\parallel s} &= v_{T0} \Omega_{co} \hat{V}_{\parallel s} \\
S_{\parallel s} &= n_0 \Omega_{co} T_0 \hat{S}_{\parallel s} \\
S_{0\parallel s} &= n_0 \Omega_{co} v_{T0} \hat{S}_{0\parallel s} \\
\end{align*}
\]

Table F.12: Links between physical and normalized quantities.

Notice that the parallel velocity is normalized to \( v_{T0} \) while all drift velocities are normalized to \( v_{T0} \).

The normalized system of equations is made of the 5D gyrokinetic equation (including source terms and collisions) self-consistently coupled to the 3D quasi-neutrality. The gyrokinetic equation involves Poisson brackets \([\cdot,\cdot]\). Introducing the unit vector \( \mathbf{b} = B/\|B\| \) along the magnetic field, its covariant components \( b_k \), and the jacobian \( J_x \) of the configuration space, these brackets read as follows:

\[
[F,G] = \mathbf{b} \cdot (\nabla F \times \nabla G) = J_x^{-1} \epsilon^{ijk} \partial_i F \partial_j G b_k \tag{F.2}
\]

Appendix F.2. Normalized Boltzmann equation

The evolution of the guiding-center distribution function \( \hat{F}_s \) is governed by the following normalized Boltzmann equation:

\[
\frac{\partial \hat{F}_s}{\partial t} + \frac{1}{B_{\parallel s}} \nabla \cdot \left( B_{\parallel s} \frac{\partial \hat{F}_s}{\partial \hat{x}_s} \right) + \frac{1}{B_{\parallel s}^2} \frac{\partial}{\partial v_{G\parallel s}} \left( B_{\parallel s}^2 \frac{\partial \hat{F}_s}{\partial v_{G\parallel s}} \right) = \hat{\mathcal{D}}_s(\hat{F}_s) + \hat{\mathcal{K}}(\hat{F}_s) + \hat{\mathcal{C}}(\hat{F}_s) + \hat{\mathcal{S}} \tag{F.3}
\]
where $\hat{D}_s$ and $\hat{K}$ are respectively a diffusion term and a Krook operator applied on a radial buffer region (see section Appendix F.4), while $\hat{C}$ refers to a collision operator (see section Appendix F.5) and $\hat{S}$ corresponds to source terms (see Appendix A in [106] for detailed expressions). The evolution of the gyro-center coordinates $(x_G, v_{G||s}, \mu_s)$ of species $s$ is given by:

$$\frac{d\hat{v}_{G||s}}{dt} = -\frac{1}{\sqrt{A_s}} \hat{\nu}_{G||s} \cdot \nabla \phi - \hat{\phi} \cdot \nabla x^i + \hat{\phi}_s \cdot \nabla x^i$$

$$\frac{d\hat{v}_{G||s}}{dt} = -\frac{1}{\sqrt{A_s}} \hat{\mu}_s \hat{b}_s \cdot \nabla \hat{B} - \hat{Z}_s \hat{b}_s \cdot \nabla \hat{\phi} + K_{\nabla B} \hat{v}_{G||s} \nabla \frac{1}{B} \cdot \nabla \hat{B}$$

The $i$-th covariant coordinates of the normalized drift velocities are given by:

$$\hat{v}_{E \times B_s} \cdot \nabla x^i = \hat{v}_{E \times B_s}^i = \frac{1}{B_{||s}} \left[ \hat{\phi}, \hat{\phi}^i \right]$$

$$\hat{v}_{D_s} \cdot \nabla x^i = \hat{v}_{D_s}^i = K_{\nabla B} \left( \frac{\hat{\nu}_{G||s}^2 + \hat{\mu}_s \hat{B}}{Z_s B_{||s}} \right) \left[ \hat{B}, \hat{\phi} \right]$$

while $\hat{b}_s$ and $\hat{B}_{||s}$ are defined as:

$$\hat{b}_s = \frac{1}{B_{||s}} \left( \hat{B} + \frac{\sqrt{A_s}}{Z_s} \hat{b}_{G||s} \hat{J} \right)$$

$$\hat{B}_{||s} = \hat{B} + \frac{\sqrt{A_s}}{Z_s} \hat{r} \cdot \hat{b} \cdot \hat{J}$$

The normalized magnetic field $\hat{B}$ is defined as:

$$\hat{B} = \frac{R_0}{R(r, \theta)} \left[ \zeta(r) e_\theta + e_\varphi \right] \text{ with } \zeta(r) = \frac{\hat{r}}{q R_0}$$

while the normalized current reads

$$\hat{J} = \hat{J}_T \hat{R} \nabla \varphi \text{ with } \hat{J}_T = \frac{R_0}{R} \frac{\zeta}{\hat{r}} \left( 1 + \frac{\hat{r}}{R} \frac{d\zeta}{dF} - \frac{\hat{r}}{R} \cos \theta \right)$$

The constant $K_{\nabla B}$ has been added for tests in the code. $K_{\nabla B}$ is equal to 1 if the curvature of the magnetic field is taken into account and 0 otherwise. In equations (F.4) and (F.5) the parallel projection $\nabla_{||F} F$ is computed as $\hat{b}_s \cdot \nabla F = \hat{b}_s \cdot \hat{\phi} F$. In equations (F.6) and (F.7), the explicit expressions of the Poisson brackets $[\hat{\phi}, x^i]$ and $[B, x^i]$ are given by:

$$[\hat{\phi}, x^i] = \frac{1}{J_{xB}} \left\{ B_2 \partial_x \hat{\phi} - B_3 \partial_z \hat{\phi} \right\}$$

$$[\hat{\phi}, x^2] = \frac{1}{J_{xB}} \left\{ -B_1 \partial_x \hat{\phi} + B_3 \partial_z \hat{\phi} \right\}$$

$$[\hat{\phi}, x^3] = \frac{1}{J_{xB}} \left\{ B_1 \partial_x \hat{\phi} - B_2 \partial_z \hat{\phi} \right\}$$

$$[B, x^1] = -\frac{1}{J_{xBB}} \times B_3 \partial_x B$$

$$[B, x^2] = \frac{1}{J_{xBB}} \times B_2 \partial_x B$$

$$[B, x^3] = \frac{1}{J_{xBB}} \left( B_1 \partial_x B - B_2 \partial_z B \right)$$

Appendix F.3. Normalized quasi-neutrality equation

The normalized quasi-neutrality equation is $^{12}$:

$$-\frac{1}{\hat{n}_e} \sum_s \hat{A}_s \nabla_{\perp} \cdot \left( \frac{\hat{n}_G}{B} \nabla_{\perp} \hat{\phi} \right) + \frac{1}{Z_f T_e} \left[ \hat{\phi} - \hat{\lambda}(\hat{\phi})_{FS} \right] = \frac{1}{\hat{n}_e} \sum_s \hat{Z}_s \left[ \hat{n}_G_s - \hat{n}_G_{s, eq} \right]$$

$^{12}$In the code, to avoid the expensive gyroaverage operation for each value of $v_{G||s}$—which occur for the computation of the RHS of the quasi-neutrality equation (F.11)—we use the definition (F.8) of $B_{||s}$ and therefore the fact that the term $\hat{n}_G_s - \hat{n}_G_{s, eq}$
where the normalized electron density \( \hat{n}_{ce} \) is defined as \( \hat{n}_{ce} = \sum_s Z_s \hat{Z}_s \hat{n}_{so} \). Notice that, in the polarization term (first term of eq. (F.11)), \( \hat{B} \) has been replaced by \( \hat{B}_0 = 1 \). The integral \( \langle \phi \rangle_{\Omega s} = \int \hat{\phi} J_s d\theta d\phi / \int J_s d\theta d\phi \) represents the flux surface average of \( \phi \) (with \( \hat{J}_s = 1/(\hat{B} \cdot \hat{\nabla} \theta) \) being the normalized jacobian space). The parameter \( \lambda \) has been added for tests. It can be chosen equal to 1 or 0. The normalized guiding-center density \( \hat{n}_{G_s} \) of species \( s \) is given by:

\[
\hat{n}_{G_s} = \int d\hat{\mu}_s \int \hat{J}_s d\hat{\nu}_{G||s} \hat{J}_0 \cdot \hat{F}_s \tag{F.12}
\]

with the normalized jacobian in velocity equal to \( \hat{J}_v = 2\pi \hat{B}^*_s \). The correction term \( \hat{n}_{G_s, eq} \) in the right hand side is defined as follows:

\[
\hat{n}_{G_s, eq} = \int d\hat{\mu}_s \int \hat{J}_s d\hat{\nu}_{G||s} \hat{J}_0 \cdot \hat{F}_{s, eq} \tag{F.13}
\]

where

\[
\hat{F}_{s, eq} = c_s \frac{\hat{n}_{so}}{(2\pi T_s)^{3/2}} \exp \left[ -\left( \frac{\hat{V}_{G||s}^2}{2} + \hat{\mu}_s \hat{B} \right) \right] \tag{F.14}
\]

represents the equilibrium part of the distribution function. The concentration \( c_s \) is such that \( \sum_s c_s Z_s \hat{Z}_s = 1 \). The normalized gyroaverage operator \( \hat{J}_0 \) approximated by Padé corresponds to \( \hat{J}_0 \approx 1 + \frac{1}{2} \frac{\hat{A}_s}{\hat{B} \hat{\nabla}^2} \).

where, as in the quasi-neutrality equation, \( \hat{B} \) is replaced by \( \hat{B}_0 = 1 \) in the code.

Appendix F.4. Normalized diffusion terms and Krook operator in buffer regions

A radial diffusion and an artificial damping are applied in buffer regions. These buffer regions are defined at each boundaries of the radial domain \( \hat{r} \in [\hat{r}_{\min}, \hat{r}_{\max}] \) as a sum of hyperbolic tangents:

\[
\hat{H}_{\text{buff}}(\hat{r}) = 1 + \frac{1}{2} \left[ \tanh \left( \frac{\hat{r} - \hat{r}_{\max} + B_L \hat{L}_r}{B_S \hat{L}_r} \right) - \tanh \left( \frac{\hat{r} - \hat{r}_{\min} - B_L \hat{L}_r}{B_S \hat{L}_r} \right) \right] \tag{F.15}
\]

where \( \hat{L}_r \) is the normalized length of the radial domain. \( B_L \) and \( B_S \) are respectively the location and the stiffness of the buffer regions. The function \( \hat{H}_{\text{buff}} \) plays the role of a mask which is equal to 1 in the buffer regions and equal to 0 elsewhere.

A radial diffusion \(^{13}\) is applied on this buffer region by solving the following equation

\[
\frac{\partial \hat{F}_s}{\partial \hat{u}} = \hat{D}_r(\hat{F}_s) \quad \text{with} \quad \hat{D}_r(\hat{F}_s) = \frac{1}{\hat{r} \hat{B}^*_s} \frac{\partial}{\partial \hat{r}} \left( \hat{r} \hat{\chi}(\hat{r}) \frac{\partial}{\partial \hat{r}} \left( \hat{B}^*_s \hat{F}_s \right) \right) \tag{F.16}
\]

where \( \hat{\chi}(\hat{r}) = \hat{\chi}_0 \hat{H}_{\text{buff}}(\hat{r}) \), where the normalized magnitude \( \hat{\chi}_0 \) of the diffusion coefficient in the buffer region is \( \hat{\chi}_0 = \chi_0/\rho_0^2 \Omega_{ce} \).

\(^{13}\)In the code, there exists the possibility to apply the diffusion term to \( (\hat{F}_s - \hat{F}_{s, eq}) \) instead of \( \hat{F}_s \).
A Krook operator is applied by solving $\partial_t \hat{F}_s = \hat{K}_s(\hat{F}_s)$ with $\hat{K}_s(\hat{F}_s) = -\hat{v}(\hat{r})(\hat{F}_s - \hat{F}_{s,eq})$ where $\hat{v}(\hat{r}) = \hat{v}_0 H_{\text{buf}}(\hat{r})$ such that $\hat{v}(\hat{r}) = \nu(\hat{r})/\Omega_0$.

**Appendix F.5. Normalized collisionality**

A collision operator $C_s$ is present on the RHS of the normalized Vlasov equation (F.3), such that the normalized Boltzmann equation reads:

$$\frac{d\hat{F}_s}{dt} = \hat{\mathcal{C}}_s(\hat{F}_s) \quad \text{with} \quad \hat{\mathcal{C}}_s(\hat{F}_s) = \frac{1}{B^*_{||s}} \frac{\partial}{\partial \hat{v}_{G||s}} \left[ \hat{B}^*_{||s} \left( \hat{D}_{||s}^* \frac{\partial \hat{F}_s}{\partial \hat{v}_{G||s}} - \hat{V}_{||s} \hat{F}_s \right) \right]$$  \hspace{1cm} (F.17)

where the diffusion term $\hat{D}_{||s}$ and the drag term $\hat{V}_{||s}$ in the parallel direction are defined as:

$$\hat{D}_{||s} = \hat{A}_s(\hat{r}) \left( \frac{\Phi(\hat{v}) - G(\hat{v})}{2\hat{v}} \right) ; \quad \hat{V}_{||s} = -\left( \frac{\hat{v}_{G||s} - \hat{V}_{s,\text{coll}}}{\hat{v}_{T_s,\text{coll}}^2} \right) \hat{D}_{||s}$$  \hspace{1cm} (F.18)

where $\hat{v} = \sqrt{\frac{1}{2} \hat{v}_{G||s}^2 + \hat{\mu}_s \hat{B}_{\text{axis}}} / \hat{T}_{s,\text{coll}}$ while the Chandrasekhar function $G$ is defined as follows:

$$G(\hat{v}) = \frac{\Phi(\hat{v}) - \hat{v} \Phi'(\hat{v})}{2\hat{v}^2} \quad ; \quad \Phi(\hat{v}) = \frac{2}{\sqrt{\pi}} \int_0^{\hat{v}} e^{-x^2} \, dx \quad ; \quad \Phi'(\hat{v}) = \frac{2}{\sqrt{\pi}} e^{-\hat{v}^2}$$  \hspace{1cm} (F.19)

The radial profile $\hat{A}_s$ is given the form

$$\hat{A}_s(\hat{r}) = \frac{1}{\sqrt{\hat{A}_s}} \left( \frac{3}{2} \frac{\hat{v}_{T_s,\text{coll}}^3}{q(\hat{r})R_0} \right) \hat{\nu}_{s,s}$$  \hspace{1cm} (F.20)

The normalized collisionality $\hat{\nu}_{s,s}$ is an input data. The normalized collisional frequency is then obtained as follows:

$$\hat{\nu}_{s,s} = \frac{1}{\sqrt{\hat{A}_s}} \left( \frac{3}{2} \frac{\hat{v}_{T_s,\text{coll}}^3}{q(\hat{r})R_0} \right) \hat{\nu}_{s,s}$$  \hspace{1cm} (F.21)

Considering that the species $s$ is the major ion species, the collisionality $\hat{\nu}_{s,s'}$ of each minority ion species $s'$ is determined by

$$\hat{\nu}_{s,s'} = \left( \frac{\hat{n}_{s'}}{\hat{n}_s} \right) \left( \frac{\hat{Z}_{s'}}{\hat{Z}_s} \right)^4 \left( \frac{\hat{T}_{s,\text{coll}}}{\hat{T}_{s',\text{coll}}} \right)^2 \hat{\nu}_{s,s}$$  \hspace{1cm} (F.22)

In practice, $\hat{T}_{s,\text{coll}} / \hat{T}_{s',\text{coll}}$ is approximated by $\hat{T}_s / \hat{T}_{s'}$. The ratio $\hat{n}_{s'}/\hat{n}_s$ and $\hat{T}_{s,\text{coll}} / \hat{T}_{s',\text{coll}}$ are computed at the radial point $\hat{r}_p$ which usually corresponds to the middle of the simulation radial domain. The normalized mean velocity $\hat{V}_{||s,\text{coll}}$ and the normalized mean temperature $\hat{T}_{s,\text{coll}}$ can be calculated as follows:

$$\hat{V}_{||s,\text{coll}} = \hat{P}^{-1} \left( \hat{I}_4 \times \hat{I}_1 - \hat{I}_2 \times \hat{I}_3 \right) \quad ; \quad \hat{T}_{s,\text{coll}} = \hat{P}^{-1} \left( \hat{I}_0 \times \hat{I}_2 - \hat{I}_1^2 \right)$$  \hspace{1cm} (F.23)

with the normalized mean pressure defined by $\hat{P} = \hat{I}_0 \times \hat{I}_4 - \hat{I}_1 \times \hat{I}_3$ and where the five integrals $\hat{I}_0$, $\hat{I}_1$, $\hat{I}_2$, $\hat{I}_3$ and $\hat{I}_4$ are:

$$\hat{I}_0 = \langle \hat{D}_{||s} \rangle \quad ; \quad \hat{I}_1 = \langle \hat{D}_{||s} \hat{v}_{G||s} \rangle \quad ; \quad \hat{I}_2 = \langle \hat{D}_{||s} \hat{v}_{G||s}^2 \rangle$$  \hspace{1cm} (F.24)

$$\hat{I}_3 = \langle \frac{1}{\hat{B}^*_{||s}} \partial_{v_{G||s}} \left( \hat{B}^*_{||s} \hat{D}_{||s} \right) \rangle \quad \text{and} \quad \hat{I}_4 = \langle \frac{1}{\hat{B}^*_{||s}} \partial_{v_{G||s}} \left( \hat{B}^*_{||s} \hat{v}_{G||s} \hat{D}_{||s} \right) \rangle$$  \hspace{1cm} (F.25)
Appendix F.6. Normalized collisional transfers between two species

Normalized collisional energy transfer. Energy exchange between two species can be expressed by the following reduced collision operator:

\[
\begin{align*}
\frac{d\tilde{F}_s}{dt} &= -\frac{\tilde{\eta}_s r}{(2\pi \tilde{T}_{\text{mean}})^{3/2} \tilde{T}_{\text{mean}}} \Delta \tilde{T}_{ss'} \left( \frac{\tilde{E}_{ss'}}{\tilde{T}_{\text{mean}}} - \frac{3}{2} \right) \exp \left( -\frac{\tilde{\eta}_s r}{\tilde{T}_{\text{mean}}} \right) \\
\frac{d\tilde{F}_s'}{dt} &= -\frac{\tilde{\eta}_s r}{(2\pi \tilde{T}_{\text{mean}})^{3/2} \tilde{T}_{\text{mean}}} \Delta \tilde{T}_{s's} \left( \frac{\tilde{E}_{s's}}{\tilde{T}_{\text{mean}}} - \frac{3}{2} \right) \exp \left( -\frac{\tilde{\eta}_s r}{\tilde{T}_{\text{mean}}} \right)
\end{align*}
\]  

with

\[
\tilde{T}_{\text{mean}} = \frac{\tilde{T}_s + \tilde{T}_{s'}}{2} ; \quad \Delta \tilde{T}_{ss'} = \tilde{T}_s - \tilde{T}_{s'} = -\Delta \tilde{T}_{s's} ; \quad \tilde{V}_{\text{mean}} = \frac{\tilde{V}_{\|s} + \tilde{V}_{\|s'}}{2}
\]

\[
\tilde{\eta}_s r = \frac{8 \tilde{E}_s}{3\sqrt{\pi}} \frac{\tilde{Z}_s}{\tilde{Z}_s'} \left( \frac{\tilde{Z}_s'}{\tilde{Z}_s} \right)^2 \left( \frac{\tilde{\nu}_{s'} \sqrt{(\tilde{T}_{s\text{coll}})_{\text{FS}}}}{qR_0} \right) \left( 1 + \frac{\tilde{A}_s \tilde{T}_{s'}^2}{\tilde{A}_{s'} \tilde{T}_s^2} \right)^{-3/2}
\]

Let us notice that in practice, the flux surface average of \( \tilde{T}_{s\text{coll}} \) is used (compare to the general expression given by (42)), to consider only the radial dependency of \( \tilde{\eta}_s r \). The normalized fluid moments \( \tilde{V}_{\|s} \) and \( \tilde{T}_s \) are computed as \( 3\tilde{T}_s(r)/2 = \langle \int \tilde{E}_s \tilde{F}_s d^2\tilde{V} \rangle_{\text{FS}}/\tilde{N}_s(r) \) and \( \tilde{V}_{\|s}(r) = \langle \int \tilde{v}_{G\|s} \tilde{F}_s d^2\tilde{V} \rangle_{\text{FS}}/\tilde{N}_s(r) \) where \( \tilde{E}_s \) is defined as \( \tilde{E}_s = \frac{1}{2} \left( \tilde{v}_{G\|s} - \tilde{V}_{\|s} \right)^2 + \tilde{\mu}_s \tilde{B} \) while \( \tilde{N}_s(r) = \langle \int \tilde{F}_s d^2\tilde{V} \rangle_{\text{FS}} \).

Normalized collisional momentum transfer. Momentum exchange between two species can be expressed by the following reduced collision operator

\[
\begin{align*}
\frac{d\tilde{F}_s}{dt} &= -\tilde{\eta}_s^{\text{col}} \tilde{A}_s^{-1/2} \Delta \tilde{V}_{ss'} \tilde{v}_{G\|s} \exp \left( -\frac{\tilde{\eta}_s r}{\tilde{T}_{\text{mean}}} \right) \\
\frac{d\tilde{F}_s'}{dt} &= -\tilde{\eta}_s^{\text{col}} \tilde{A}_{s'}^{-1/2} \Delta \tilde{V}_{s's} \tilde{v}_{G\|s'} \exp \left( -\frac{\tilde{\eta}_s r}{\tilde{T}_{\text{mean}}} \right)
\end{align*}
\]  

with

\[
\Delta \tilde{V}_{ss'} = \tilde{V}_{\|s} \tilde{A}_s^{1/2} - \tilde{V}_{\|s'} \tilde{A}_{s'}^{-1/2}
\]

The averages \( \tilde{T}_{\text{mean}} \) and \( \tilde{V}_{\text{mean}} \) have already been introduced. The energy \( \tilde{E}_s \) is equal to \( \frac{1}{2} \tilde{v}_{G\|s}^2 + \tilde{\mu}_s \tilde{B} \). The radial profile \( \tilde{\eta}_s^{\text{col}} \) reads:

\[
\tilde{\eta}_s^{\text{col}}(r) = \left( \frac{\pi}{2} \right)^{3/2} \frac{1}{qR_0 \tilde{\nu}_s} \frac{\tilde{v}_{G\|s} \sqrt{(\tilde{T}_{s\text{coll}})_{\text{FS}}}}{\tilde{T}_{\text{mean}}^{3/2}} \left( \frac{\tilde{Z}_s'}{\tilde{Z}_s} \right)^2 \tilde{\nu}_s
\]  

where the velocity \( \tilde{v}_{G\|s} \) is calculated as \( \tilde{v}_{G\|s} = \max(\tilde{A}_s^{-1/2} \tilde{v}_{T_s}, \tilde{A}_{s'}^{-1/2} \tilde{v}_{T_{s'}}) \) where \( \tilde{v}_{T_s} = \sqrt{\tilde{T}_s} \) and \( \tilde{v}_{T_{s'}} = \sqrt{\tilde{T}_{s'}} \) and the collision frequencies \( \tilde{\nu}_s, \tilde{\nu}_{s'} \) are linked by the relation (F.22).
Appendix F.7. Normalized source terms

Focusing on the source term, the gyrokinetic equation is:

\[
\frac{d\hat{\psi}}{dt} = \hat{S}_\text{heat}(\hat{r}, \theta, \hat{\psi}_G, \hat{\mu}_s) + \hat{S}_\text{moment}(\hat{r}, \theta, \hat{\psi}_G, \hat{\mu}_s) + \hat{S}_\text{vorticity}(\hat{r}, \theta, \hat{\psi}_G, \hat{\mu}_s)
\]  

(F.33)

where the heat source can be defined as

\[
\hat{S}_\text{heat} = \frac{\hat{S}_\text{heat}}{3\sqrt{2}\left(\pi T_{s,\text{src}}\right)^{3/2}} \left(\hat{E}_s - \frac{3}{2}\right) e^{-\hat{E}_s/T_{s,\text{src}}} \hat{S}_r^\xi
\]  

(F.34)

with \(\hat{E}_s = \frac{1}{2} \hat{v}_G + \hat{\mu}_s \hat{B}\) or as a pure source of heating

\[
\hat{S}_\text{heat} = \left\{ \begin{array}{l}
\left[ \hat{v}_G^2 - \frac{1}{2} \hat{J}_B - \frac{2}{3} \hat{\mu}_s \left( 2 \hat{\mu}_s - \hat{J}_B \right) \right] \frac{2\hat{S}_0^{\xi}}{\left( 2\pi T_{s,\text{src}} \right)^{3/2}} \hat{S}_r^\xi \\hat{S}_r^\xi \quad \text{e}^{-\hat{E}_s/T_{s,\text{src}}} \hat{S}_r^\xi
\end{array} \right.
\]  

(F.35)

with \(\hat{\mu}_s = \frac{\hat{\mu}_s \hat{B}}{T_{s,\text{src}}} \), \(\hat{v}_G^2 = \frac{\sqrt{2} T_{s,\text{src}}}{B^2} \) and \(\hat{J}_B = \frac{\hat{A}_s \hat{T}_{s,\text{src}}}{Z_s \hat{\rho}_{s,\text{src}}} \). The source \(\hat{S}_\text{moment} \) is a pure source of momentum expressed as

\[
\hat{S}_\text{moment} = \left\{ \begin{array}{l}
\left[ 2\hat{\psi}_G^2 \left( 2 \hat{\mu}_s - \hat{J}_B \right) + \hat{\psi}_G^2 \left( 1 + 2 \hat{\psi}_G \hat{\mu}_s - \hat{\psi}_G \hat{\mu}_s \right) \right] \frac{\hat{S}_0^{\psi_G}}{4\pi^{3/2} T_{s,\text{src}}^{2/3}} \hat{S}_r^\psi \hat{S}_r^\psi \quad \text{e}^{-\hat{E}_s/T_{s,\text{src}}} \hat{S}_r^\psi
\end{array} \right.
\]  

(F.36)

and the \(\hat{S}_\text{vorticity} \) is a pure source of vorticity is defined as

\[
\hat{S}_\text{vorticity} = -\left\{ \begin{array}{l}
\left[ \hat{v}_G^2 \left( \hat{\mu}_s - \hat{J}_B \right) \right] \frac{\hat{S}_0^{\xi}}{\left( 2\pi T_{s,\text{src}} \right)^{3/2}} \hat{S}_r^\xi \quad \text{e}^{-\hat{E}_s/T_{s,\text{src}}} \hat{S}_r^\xi
\end{array} \right.
\]  

(F.37)

The radial components of the sources (\(i.e\), \(\hat{S}_r^\xi, \hat{S}_r^\psi, \hat{S}_r^\xi \)) are defined as:

\[
\hat{S}_r^\xi(\hat{r}) = -\frac{1}{2} \left[ \tanh \left( \frac{\hat{\hat{\rho}} - \left( \hat{\hat{\rho}} + 3\hat{L}_{\xi} \right) \hat{L}_{\xi}}{L_{\xi}} \right) + \tanh \left( \frac{\hat{\hat{\hat{\rho}}} - \left( \hat{\hat{\hat{\rho}}} - 3\hat{L}_{\xi} \right) \hat{L}_{\xi}}{L_{\xi}} \right) \right]
\]  

(F.38)

where \(\hat{\hat{\rho}} = (\hat{r} - \hat{r}_{\text{min}})/\hat{L}_r \), \(\hat{\hat{\rho}} \) and \(\hat{L}_\xi \) are input data corresponding to the radial position and the normalized width of the different sources. These radial parts are normalized such that

\[
\int_{\hat{r}_{\text{min}}}^{\hat{r}_{\text{max}}} \hat{r} d\hat{r} \hat{S}_r^\xi(\hat{r}) = 1
\]

Appendix F.8. Normalized source of impurity

Let us consider the species \(s \) as the major species and the species \(s' \) as an impurity species. Then, it is possible to add a source of impurities, by adding to the right hand side of equation (F.33) a source of matter. The source \(\hat{S}_n(\hat{r}, \theta, \hat{\psi}_G, \hat{\mu}_s) \) of impurity \(s' \) is defined by

\[
\hat{S}_n = \frac{\hat{S}_0^{\psi_G} \hat{S}_r}{\left( 2\pi T_{s,\text{src}} \right)^{3/2}} \left( \frac{5}{2} - \hat{\mu}_{s'} - \hat{\psi}_{G}^2 \right) \exp \left( -\hat{E}_{G}^{2} \hat{\mu}_{s'} \right)
\]  

(F.39)
where \( \tilde{v}_{G,s'} = \sqrt{2T_{s',\text{srce}}} \) and \( \mu_{s'} = \mu_{s'} B / T_{s',\text{srce}} \). \( S_0^{n,s} \) corresponds to the normalized intensity of the source. To avoid any injection of charges, the injection of impurities must be compensated by the injection of source of matter for the majority species such that

\[
Z_s \int S_{n,s} J_{v,s} \, dv_G \, d\mu_s + Z_{s'} \int S_{n,s'} J_{v,s'} \, dv_G \, d\mu_{s'} = 0
\]

with \( J_{v,s} \) (resp. \( J_{v,s'} \)) the jacobian in velocity space for species \( s \) (resp. species \( s' \)). The normalized radial profiles of the sources are assumed equal to \( \tilde{S}_r \). Therefore,

\[
\tilde{S}_{n,s} = \frac{\tilde{S}_{0,s} \tilde{S}_{0,s'}}{(2\pi \tilde{T}_{s,\text{srce}})^{3/2}} (5/2 - \tilde{\mu}_s - \tilde{v}_{G,s}^2) e^{-\tilde{v}_{G,s}^2 - \tilde{\mu}_s} \quad \text{with} \quad \tilde{S}_0^{n,s} = - \frac{Z_{s'}}{Z_s} \left( \frac{\tilde{T}_{s',\text{srce}}}{\tilde{T}_{s,\text{srce}}} \right)^{3/2} \tilde{S}_0^{n,s'}
\]

(F.40)

In the code, the temperature profile of the density sources are taken equal, i.e. \( \tilde{T}_{s,\text{srce}} = \tilde{T}_{s',\text{srce}} \).

**Appendix G. Invariance**

Let us consider a reference simulation without source terms. Let us also define a second simulation similar to the reference one but where the mass and charge have been respectively multiplied by a factor \( \alpha_A \) and \( \alpha_Z \). Then, it is shown in this appendix that 7 control parameters \( (\alpha_Z, \alpha_T, \alpha_{\text{diff}}, \alpha_{\text{coll}}, \alpha_{\text{charge}}, \alpha_{\text{Krook}}) \) that respectively rescale the space, the time, the electron temperature, the electrostatic potential, the collision operator, the diffusion term and the Krook operator) are sufficient to ensure that both simulations –reference and scaled– are equivalent. The constraints on the different control parameters are summarized in table G.13 and the proof follows.

<table>
<thead>
<tr>
<th>Scaling</th>
<th>Charge Mass Length Time</th>
<th>( T_e )</th>
<th>Elec. potential</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collision</td>
<td>( \alpha_Z )</td>
<td>( \alpha_A )</td>
<td>( \sqrt{\alpha_A / \alpha_Z} )</td>
</tr>
<tr>
<td>Krook</td>
<td>( \alpha_Z / \alpha_A )</td>
<td>( \alpha_A^{-1} )</td>
<td></td>
</tr>
</tbody>
</table>

Table G.13: Scaling which must be applied to the different characteristic quantities to obtain an equivalence between any reference simulation and a scaled simulation.

For the proof let us first consider the following normalized reference Vlasov equation (deduced from eqs. (F.3) to (F.8)) for a species of charge \( \tilde{Z}_s \) and mass \( \tilde{A}_s \)

\[
\frac{\partial \tilde{F}_s}{\partial t} + \left( \frac{1}{\sqrt{\tilde{A}_s}} \tilde{v}_{G||} \tilde{b}^* + \left( \frac{\tilde{v}_{G||}^2 + \tilde{\mu}_s \tilde{B}}{\tilde{Z}_s \tilde{B}_{||}^* \tilde{B}} \right) \left( \tilde{b} \times \tilde{\nabla} \tilde{B} \right) + \frac{1}{\tilde{B}_{||}^*} \left( \tilde{b} \times \tilde{\nabla} \tilde{\phi} \right) \right) \cdot \tilde{\nabla} \tilde{F}_s
\]

\[
+ \left( - \frac{1}{\sqrt{\tilde{A}_s}} \tilde{\mu}_s \tilde{b}^* \cdot \tilde{\nabla} \tilde{B} - \frac{\tilde{Z}_s}{\sqrt{\tilde{A}_s}} \tilde{b}^* \cdot \tilde{\nabla} \tilde{\phi} + \tilde{v}_{G||} \tilde{E}_s \cdot \tilde{\nabla} \ln \tilde{B} \right) \frac{\partial \tilde{F}_s}{\partial \tilde{v}_{G||}}
\]

\[
= \frac{1}{\tilde{B}_{||}^*} \frac{\partial}{\partial \tilde{v}_{G||}} \left[ \tilde{B}_{||}^* \tilde{D}_{||s} \left( \frac{\partial \tilde{F}_s}{\partial \tilde{v}_{G||}} + \frac{(\tilde{v}_{G||} - \tilde{V}_{s,\text{coll}})}{\tilde{v}_{G||}^2} \tilde{F}_s \right) \right] + \frac{1}{\tilde{B}_{||}^*} \left[ 1 \frac{\partial}{\partial \tilde{r}} \left( \tilde{r} \tilde{\chi} \tilde{B}_{||}^* \frac{\partial \tilde{F}_s}{\partial \tilde{r}} \right) \right]
\]

\[
\equiv \tilde{b}^* = \frac{1}{\tilde{B}_{||}^*} \left( \tilde{B} + \sqrt{\tilde{A}_s \tilde{Z}_s \tilde{B}} \tilde{\nabla} \times \tilde{B} \right)
\]

64
Therefore, the rescaled equation reads:

\[ \tilde{D}_s(\tilde{r}, \tilde{v}) = \hat{A}_s(\tilde{r}) \left( \frac{\Phi(\tilde{v}) - G(\tilde{v})}{2\tilde{v}} \right) \]  

with  

\[ \hat{A}_s(\tilde{r}) = \frac{1}{\sqrt{A_s}} \left( 3\sqrt{\pi} \frac{v^3_{T,\text{coll}}}{2 q(\tilde{r}) B_0} \right) \tilde{v}_{s} \]

The Chandrasekhar function \( G \) and the function \( \Phi \) are defined by equation (F.19). Let us apply a scaling with the control parameters defined below, this leads to the following new Vlasov equation

\[ \frac{1}{\alpha_t} \frac{\partial \hat{F}_s}{\partial t} + \frac{1}{\sqrt{\alpha_A} A_s} \frac{\tilde{v}_{G||}}{\alpha_s} \hat{b}_s^* \left( \frac{v^2_{G||} + \tilde{\mu}_s \hat{B}}{\alpha_s \bar{Z}_s B^*_s \hat{B}} \right) \left( \mathbf{b} \times \frac{1}{\alpha_x} \nabla \bar{B} \right) + \frac{1}{\hat{B}_s^*} \left( \mathbf{b} \times \frac{1}{\alpha_x} \nabla (\alpha \hat{\phi}) \right) \} \cdot \nabla \hat{F}_s = \frac{1}{\alpha_x} \frac{\partial}{\partial \phi} \left( \frac{\hat{F}_s}{\alpha_x} \right) \]

Concerning the left hand side, this equation is equivalent to the reference one, if and only if:

\[ \frac{1}{\alpha_t} = \frac{1}{\sqrt{\alpha_A} \alpha_s} = \frac{1}{\alpha_s \alpha_x} = \frac{1}{\alpha_x} \alpha_s \alpha_x \]

\[ \alpha_A = \alpha_s \alpha_x \]

(G.1)

(G.2)

Therefore, according to the last equality of equation (G.2), \( \alpha_x = \sqrt{\alpha_A}/\alpha_Z \). Besides, due to the last equality of equation (G.1), \( \alpha_s = \alpha_x^{-1} \). And finally, due to the first equality of equation (G.1), \( \alpha_t = \sqrt{\alpha_A}/\alpha_x \), so \( \alpha_t = \alpha_A/\alpha_Z \). Taking into account the collision operator \( 1/\alpha_t = \alpha_{\text{coll}}/(\sqrt{\alpha_A} \alpha_x) \) which gives \( \alpha_{\text{coll}} = 1 \). In addition, the equivalence condition for the Krook operator leads to \( 1/\alpha_t = \alpha_{\text{Krook}} \) thus \( \alpha_{\text{Krook}} = \alpha_Z/\alpha_A \) while for the diffusion term the condition reads \( 1/\alpha_t = \alpha_{\text{diff}}/\alpha_x \). i.e \( \alpha_{\text{diff}} = \alpha_Z^2 \).

Concerning the equivalence for the quasi-neutrality let us consider normalized equation

\[- \sum_s \hat{A}_s \hat{\mathbf{v}}_s \left( \frac{\phi - \langle \hat{\phi} \rangle_{FS}}{B} \right) + \frac{1}{Z_\alpha} \sum_s \bar{Z}_s \hat{n}_{s, \text{eq}} \left( \frac{\phi - \langle \hat{\phi} \rangle_{FS}}{B} \right) = \sum_s \hat{Z}_s \int \hat{J}_0 \left( \hat{F}_s - \hat{F}_{s, \text{eq}} \right) \]  

d\( \hat{v} \)

Therefore, the rescaled equation reads:

\[- \frac{\alpha A \phi}{\alpha_x^2} \sum_s \hat{A}_s \hat{\mathbf{v}}_s \left( \frac{\hat{n}_{s, \text{eq}}}{B} \right) + \frac{\alpha Z \phi}{\alpha T_e} \sum_s Z_s \hat{n}_{s, \text{eq}} \left( \frac{\phi - \langle \hat{\phi} \rangle_{FS}}{B} \right) = \sum_s \alpha Z \hat{Z}_s \int \hat{J}_0 \left( \hat{F}_s - \hat{F}_{s, \text{eq}} \right) \]  

d\( \hat{v} \)

The equivalence between reference and rescaled QN equations is ensured if and only if:

\[ \alpha A \phi = \alpha Z \phi \Rightarrow \alpha_t = \alpha_T^2 \frac{\alpha Z}{\alpha A} \]  

and  

\[ \alpha_Z = \alpha A \phi \frac{\alpha Z}{\alpha T_e} \Rightarrow \alpha_\phi = \alpha_T \]

Therefore \( \alpha_T = 1/\alpha_Z \) which confirms the previous equality \( \alpha_\phi = \alpha_Z^{-1} \).

Finally, regarding the gyroaverage operator in the case of a Padé approximation, i.e:  

\[ \hat{J}_0 \sim 1 + \frac{1}{2} \frac{\alpha T_e}{\alpha_x^2} \hat{\mathbf{v}}_s^2 \].
its invariance is ensured provided the equality $\alpha_A = \alpha_Z^2 \alpha_x^2$ is fulfilled. This equality is already satisfied by the previous constraints.

Appendix H. Rosenbluth-Hinton test

In this appendix, we report on simulations of the so-called Rosenbluth-Hinton (R-H) test [102]. This consists in studying the collisionless relaxation of an initial $E \times B$ poloidal flow, including the transient GAM oscillation (Geodesic Acoustic Mode), towards a non vanishing residual value. Such simulations are interesting validation tests for gyrokinetic codes because the damping, the frequency and the residual value of the GAM can be compared to theoretical predictions [59, 130]. In practice, a zonal perturbation in ion density is initialized, with a radial profile of the form $\sin(\pi r/a)$. This perturbation generates a zonal radial electric field which evolves in time as predicted by the R-H theory. The parameters used for the following simulations are the same as those used by Biancalani [8], namely: (i) an analytical circular equilibrium with large aspect ratio ($\epsilon = a/R = 0.1$); (ii) flat density and temperature profiles with $\tau = T_i/T_e = 1$ and (iii) flat $q$-profiles varying from $q = 1.5$ to $3.5$. All simulations are performed for a mesh $(N_r,N_\theta,N_\phi,N_{v_\|},N_\mu) = (256,256,16,128,32)$. The results obtained with GYSELA are displayed on figs.H.16,H.17, where the GAM frequency, the damping rate and the residual value are plotted as a function of the safety factor $q$. They are comparable to those published by the ORB5 team (see figures 2 and 3 in [8]). Consistently with the theory, the FOW (Finite Orbit Width) effects are already significant at moderate values of $q$ (typically for $q \leq 2$, cf. fig.H.16a). Regarding the residual values, two ratio are considered: either the ratio of the finite over the initial zonal electric potential, or the ratio of the radial electric fields. Both ratios would be equal if the radial profiles of the electric potential would not evolve in time, as assumed by the theory. As evident on fig.H.17, this is actually not the case in these global simulations.

<table>
<thead>
<tr>
<th>$\rho_s$</th>
<th>$A_s$</th>
<th>$Z_s$</th>
<th>$\epsilon_0$</th>
<th>$r_{int}/a$</th>
<th>$r_{ext}/a$</th>
<th>Torus</th>
<th>$n_b v_{th0}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/160</td>
<td>1</td>
<td>1</td>
<td>0.1</td>
<td>0.01</td>
<td>1</td>
<td>1</td>
<td>7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\mu_{max}$</th>
<th>$r_{peak}/a$</th>
<th>$\kappa_{n_{s0}}$</th>
<th>$\Delta_{n_{s0}}$</th>
<th>$\kappa_{T_{s0}}$</th>
<th>$\Delta_{T_{s0}}$</th>
<th>$T_i/T_e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>0.5</td>
<td>1.0e$^{-7}$</td>
<td>0.2</td>
<td>1.0e$^{-7}$</td>
<td>0.1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table H.14: Common parameters for GAM test. The velocity phase space is defined by $-nb_{v_{th0}} v_{T_{s0}} \leq V_{Gi} \leq nb_{v_{th0}} v_{T_{s0}}$ and $0 \leq \mu \leq \mu_{max} T_0/B_0$. Torus indicates the fraction of the torus simulated. The radial density profile is defined by its gradient $\frac{d \log n_{s0}(r)}{dr} = -\kappa_{n_{s0}} \cosh^{-2} \left[ (r - r_{peak}/a)/\Delta r_{n_{s0}} \right]$. The same analytical expression is used for the temperature with $\kappa_{T_{s0}}$ and $\Delta r_{T_{s0}}$. 

66
Figure H.16: Damping rate (left) and frequency (right) of the electric field for simulation parameters given in table H.14. Comparison with explicit analytical values given by Sugama and Watanabe [59] and Zonca [130] where FOW effects are taken into account or not.

Figure H.17: Residual values divided by the initial value of the electric potential (green stars) or electric field (magenta triangles) for the 5 different values of constant $q$ profile :1.5, 2., 2.5, 3. and 3.5. Comparison with analytical theory given by Rosenbluth-Hinton [102] (red line) and Xiao and Catto [125] (blue line).

Figure H.18: Radial profile of the $(0,0)$ mode of $\phi$ at the initial time (blue line) and at the time $t = 30000/\Omega_{C0}$ (red dotted line).

67
Appendix I. Input data files

This appendix contains the input data files of all simulations used in this paper.

15/01/2015 A150_FD_allwcoll2_semitore 1

FILE

15/01/2015 A150_FD_allwcoll2_semitore 2

FILE

Figure I.19: Input data file used in the code for the first simulation for invariance tests (see section 4.2).


Figure I.20: Example of input data file used for Rosenbluth-Hinton test (see section 4.3.1).
Figure I.21: Example of input data file used for the Cyclone Base Case benchmark with GENE code (see section 4.3.2).
Appendix J. Detailed computations for local conservation laws

In this appendix, we present several detailed calculations useful for the derivation of the gyrokinetic conservation laws in section 5.

Appendix J.1. Useful integrals

Let us consider two arbitrary fields $X$, $Y$ and $Z$. The aim of this paragraph is to compute the general integral

$$I \equiv \int J_x \, d\chi \, d\theta \, d\varphi \, X \nabla \cdot (Y \nabla_{\perp} Z) \quad (J.1)$$

where $J_x$ is the jacobian is space, i.e $J_x = \sqrt{g}$ with $g$ representing the determinant of the metric tensor. In order to perform this integration it is useful to write the operator $\nabla \cdot (Y \nabla_{\perp} Z)$ by using Einstein notations.

Given that $\nabla \varphi \cdot \nabla \theta = \nabla \varphi \cdot \nabla \chi = 0$, we can write this operator as

$$\nabla \cdot (Y \nabla_{\perp} Z) = 1 \sqrt{g} \partial_i \left( \sqrt{g} Y g^{\chi j} \partial_j Z \right) + 1 \sqrt{g} \partial_\varphi \left( \sqrt{g} Y g^{\varphi j} \partial_j Z \right)$$

Then, the previous integral (J.1) can be written as

$$I = \int J_x \, d\chi \, d\theta \, d\varphi \, X \left\{ \frac{1}{J_x} \partial_i \left( Y J_x g^{\chi j} \partial_j Z \right) \right\}$$

Then, using an integration by parts on the coordinate $i$

$$I = - \int d\chi \, d\theta \, d\varphi \left( \partial_i X \right) Y J_x g^{\chi j} \partial_j Z + \left[ \int d\theta \, d\varphi \, X Y J_x g^{\chi j} \partial_j Z \right]_{\partial L_x} + \left[ \int d\theta \, d\varphi \, X Y J_x g^{\varphi j} \partial_j Z \right]_{\partial L_x} 2\pi \quad (J.2)$$

where $i$ and $j$ are still $\vartheta$ or $\chi$. Because of the periodicity in $\theta$, the surface term resulting from the integration by parts on $i = \theta$ (i.e. last term of previous equation) is equal to 0. Therefore,

$$I = - \int J_x \, d\chi \, d\theta \, d\varphi \left( \partial_i X \right) Y J_x g^{\chi j} \partial_j Z$$

Using the fact that $g^{ij} = \nabla x^i \cdot \nabla x^j$ and $\nabla X = (\partial_i X) \nabla x^i$ for all field $X$, then

$$I = - \int J_x \, d\chi \, d\theta \, d\varphi \left[ Y (\partial_i X) \nabla x^i \cdot \nabla x^j \partial_j Z \right] - \int J_x \, d\theta \, d\varphi \, X Y \nabla \chi \cdot \nabla x^i \partial_j Z$$

therefore,

$$I = - \int J_x \, d\chi \, d\theta \, d\varphi \left[ Y \nabla_{\perp} X \cdot \nabla_{\perp} Z \right] + \int J_x \, d\theta \, d\varphi \, X Y \nabla \chi \cdot \nabla_{\perp} Z \quad (J.3)$$

\[\text{14} \] The direction $\perp$ corresponds to an approximation of the poloidal plane because it is actually perpendicular to $\nabla \varphi$ rather than $B$. 

71
Then, using again an integration by parts on the coordinate \(j\) for equation (1.2) and the fact that the surface term \(\int d\chi d\phi (\partial_t X) Y J_\chi g^{\chi j} Z \mid_{L_\chi} = 0\) due to periodicity in \(\theta\),

\[
\mathcal{I} = \int d\chi d\theta d\phi \partial_j (Y J_\chi g^{\chi j} \partial_t X) Z - \left[ \int d\theta d\phi (\partial_t X) Y J_\chi g^{\chi j} Z \right]_{\partial L_\chi} + \left[ \int d\theta d\phi X Y J_\chi g^{\chi j} (\partial_j Z) \right]_{\partial L_\chi}
\]

Finally, previous equation gives

\[
\int J_\chi d\chi d\theta d\phi X \nabla_\chi \cdot (Y \nabla_\chi Z) = \int J_\chi d\chi d\theta d\phi \nabla_\chi \cdot (Y \nabla_\chi X) Z - \left[ \int J_\chi d\theta d\phi X Y J_\chi g^{\chi j} (\partial_j Z) \right]_{\partial L_\chi}
\]

Appendix J.2. Expression of \(\sum_s q_s \int d\tau^s \bar{U} \partial_t \bar{F}_s\) for local energy conservation

The opposite of the total energy source \(-W = \sum_s q_s \int d\tau^s \bar{U} \partial_t \bar{F}_s\) can be divided into two parts

\[
\sum_s q_s \int d\tau^s \bar{U} \partial_t \bar{F}_s = \sum_s q_s \int d\tau^s U (J_\theta \cdot \partial_t \bar{F}_s) + \sum_s q_s \int d\tau^s \{ (J_\theta \cdot U) \partial_t \bar{F}_s - U (J_\theta \cdot \partial_t \bar{F}_s) \}
\]  

(3.5)

Then, using the quasi-neutrality equation (12), the first term in eq.(3.5) can be expressed as

\[
\sum_s q_s \int d\tau^s U (J_\theta \cdot \partial_t \bar{F}_s) = -\frac{1}{2} \sum_s q_s \int J_\theta d\theta d\phi U \nabla_\phi \cdot \left( \frac{n_{s,eq} \Omega_s}{\beta s} \nabla_\phi \partial_t U \right)
\]

Besides, using the general following equality (J.3) the previous equation is equivalent to

\[
\sum_s q_s \int d\tau^s U (J_\theta \cdot \partial_t \bar{F}_s) = \partial_t \left[ \frac{1}{2} \sum_s q_s \int d\tau^s U (J_\theta \cdot \bar{F}_s) \right] \equiv \partial_t E_p
\]

Let us define the potential energy as \(E_p \equiv \frac{1}{2} \sum_{\text{species}} q_s \int d\tau^s U (J_\theta \cdot \bar{F}_s)\) this term can be identified to the time derivative of potential energy,

\[
\sum_s q_s \int d\tau^s U (J_\theta \cdot \partial_t \bar{F}_s) = \partial_t \left[ \frac{1}{2} \sum_{\text{species}} q_s \int d\tau^s U (J_\theta \cdot \bar{F}_s) \right] \equiv \partial_t E_p
\]

Let us consider the low wavenumber approximation of the gyroaverage operator used in the code, \(i.e\) the Padé approximation \(J_\theta \simeq 1 + \frac{1}{2} \nabla \cdot \left( \frac{m_v \nabla U}{\mu B} \nabla \perp \right)\) (see section 2.5.1 for more details) then for each species

\[
\mathcal{I}_E = q_s \int d\tau^s \{ (J_\theta \cdot U) \partial_t \bar{F}_s - U (J_\theta \cdot \partial_t \bar{F}_s) \} = \frac{m_s}{2q_s} \int d\tau^s \left\{ \partial_t \bar{F}_s \nabla \cdot \left( \frac{\mu_s}{B} \nabla \perp U \right) - U \nabla \cdot \left( \frac{\mu_s}{B} \nabla \perp \partial_t \bar{F}_s \right) \right\}
\]

and using the result (J.3) of Appendix J

\[
\mathcal{I}_E = \frac{m_s}{2q_s} \partial_\chi \left[ \int d\tau \left\{ \partial_t \bar{F}_s \nabla \cdot \left( \frac{\mu_s}{B} \nabla \perp U \right) - U \nabla \cdot \left( \frac{\mu_s}{B} \nabla \perp \partial_t \bar{F}_s \right) \right\} \right] \\
= \frac{m_s}{2q_s} \partial_\chi \left[ -\int J_\chi d\chi d\theta d\phi d^3v \frac{\mu_s}{B} \nabla \perp (\partial_t \bar{F}_s) \cdot \nabla \perp U + \int J_\chi d\theta d\phi d^3v \partial_t \bar{F}_s \frac{\mu_s}{B} \nabla \perp (\partial_t \bar{F}_s) \right] \\
+ \frac{m_s}{2q_s} \partial_\chi \left[ \int J_\chi d\chi d\theta d\phi d^3v \frac{\mu_s}{B} \nabla \perp U \cdot \nabla \perp (\partial_t \bar{F}_s) - \int J_\chi d\theta d\phi d^3v U \frac{\mu_s}{B} \nabla \perp (\partial_t \bar{F}_s) \right]
\]

72
The first and third terms cancel each other out. The remaining terms
\[
\mathcal{I}_E = \frac{m_s}{2q_s} \partial_\chi \left[ \int d\tau^* \partial_\ell \bar{F}_s \frac{\mu}{B} \nabla \chi \cdot \nabla_{\perp} U - \int d\tau^* \frac{\mu}{B} \nabla \chi \cdot \nabla_{\perp} (\partial_\ell \bar{F}_s) \right]
\]
So finally, using the gyrocenter perpendicular stress \(P_{s,\perp} = \int d^3 v \bar{F}_s \mu_s B,\)
\[
\sum_s q_s \int d\tau^* \partial_\ell \bar{F}_s = \partial_\ell F_p + \frac{m_s}{2q_s} \partial_\chi \left[ \int d\tau^* \partial_\ell \bar{F}_s \frac{\mu}{B} \nabla \chi \cdot \nabla_{\perp} U - \int d\tau^* \frac{\mu}{B} \nabla \chi \cdot \nabla_{\perp} (\partial_\ell \bar{F}_s) \right]
\]

**Appendix J.3. Effect of the electric potential on the toroidal canonical momentum**

The objective of this section is to prove that
\[
dt P_\varphi = -q_s \partial_\varphi \bar{U} \quad \text{with} \quad P_\varphi = -q_s \chi + \frac{m_s I}{B} v_G \parallel
\]
Due to the fact that \(P_\varphi\) is an invariant of the equilibrium motion, this result is equivalent to proving that
\[
[q_s \bar{U}, P_\varphi]_{GC} = q_s \partial_\varphi \bar{U} \tag{J.6}
\]
where \([\cdot, \cdot]_{GC}\) indicate the Poisson brackets in the gyro-center coordinates. In the gyrokinetic framework, we recall the expression of the Poisson brackets for two given fields \(X\) and \(Y\)
\[
B^*_{[\parallel]} [X, Y]_{GC} = \frac{B^*_{\parallel}}{m_s} \cdot (\nabla X \partial_{v_{\parallel}} Y - \partial_{v_{\parallel}} X \nabla Y) - \frac{b}{q_s} \cdot (\nabla X \times \nabla Y)
\]
where \(b = B/B\) is the unit vector parallel to the magnetic field. The quantities \(B^*_{\parallel}\) and \(B^*_{[\parallel]}\) are respectively defined by eq.(3) and \(B^*_{[\parallel]} = B^*_{\parallel} \cdot B/B, i.e. B^*_{[\parallel]} \equiv B + \frac{m_s v_G}{q_s} \nabla \times b\) and \(B^*_{[\parallel]} \equiv B + \frac{m_s v_G}{q_s} b \cdot (\nabla \times b).\)
\[
-q_s B^*_{[\parallel]} [\bar{U}, P_\varphi]_{GC} = -q_s \nabla \cdot Y \left( \frac{I}{B} B - b \times \nabla \chi + \frac{m_s v_G}{q_s} \left( \frac{I}{B} (\nabla \times b) - \nabla \left( \frac{I}{B} \right) \times b \right) \right) \tag{J.7}
\]
In the chosen coordinate system, we recall that the magnetic field \(B\) can be written as \(B = I(\chi) \nabla \varphi + \nabla \times \nabla \chi\) with \(I\) a flux function, then
\[
B \times \nabla \chi = IB - I^2 \nabla \varphi = IB - B^2 R^2 \nabla \varphi \tag{J.8}
\]
\[
\nabla \times B = -(\partial_\chi I)B - R^2 \partial_\chi P \nabla \varphi \tag{J.9}
\]
Besides,
\[
\nabla \left( \frac{I}{B} \right) \times b = -I^2 B^2 (\nabla B \times b) + \frac{I}{B} (\nabla I \times b) = -I \frac{I}{B^2} (\nabla B \times b) + \frac{1}{B} (\partial_\chi I) \nabla \chi \times b
\]
\[
= -I \frac{I}{B^2} (b \times \nabla B) - \frac{1}{B} (\partial_\chi I) b \times \nabla \chi \tag{J.10}
\]
Then according to (J.8) and (J.10), equation (J.7) becomes
\[
-q_s B^*_{[\parallel]} [\bar{U}, P_\varphi]_{GC} = -q_s \nabla \cdot \left[ R^2 B \nabla \varphi + \frac{m_s v_G}{q_s} \left( \frac{I}{B} (\nabla \times b) - \frac{I}{B^2} (b \times \nabla B) + \frac{1}{B} (\partial_\chi I) b \times \nabla \chi \right) \right]
\]
Using the fact that \(\nabla \times b = b \times \nabla B + \frac{I}{B} \nabla \times B\), the previous equation reads
\[
-q_s B^*_{[\parallel]} [\bar{U}, P_\varphi]_{GC} = -q_s \nabla \cdot \left[ R^2 B \nabla \varphi + \frac{m_s v_G}{q_s} \left( \frac{I}{B^2} \nabla \times B + \frac{1}{B} (\partial_\chi I) b \times \nabla \chi \right) \right] \tag{J.11}
\]
Therefore,
\[
\frac{1}{B^2} \nabla \times B + \frac{1}{B}(\partial_s I) \mathbf{b} \times \nabla \chi = -R^2 \nabla \varphi \left( \frac{1}{B^2} \partial_s P + \partial_s I \right)
\]
So, using the fact that the volume element in guiding-center velocity space \( B^*_s \) can be expressed as
\[
B^*_s = B - \frac{m_s}{q_s} v_G \left( \partial_s I + \frac{1}{B^2} \partial_s P \right) \tag{J.12}
\]
equation (J.11) becomes
\[
-q_s B^*_s \left[ \bar{U}, P_s \right]_G = -q_s \nabla \cdot \left( R^2 B^*_s \nabla \varphi \right) = -q_s \partial_s \bar{U} \nabla \varphi \cdot \nabla \varphi R^2 B^*_s = -q_s B^*_s \partial_s \bar{U}
\]
which is equivalent to equation (J.6).


Fusion 41.


