

A non-twisting flux tube for local gyrokinetic simulations

Justin Ball and Stephan Brunner

Ecole Polytechnique Fédérale de Lausanne (EPFL), Swiss Plasma Center (SPC),
CH-1015 Lausanne, Switzerland

E-mail: Justin.Ball@epfl.ch

Abstract.

Local gyrokinetic simulations use a field-aligned domain that twists due to the magnetic shear of the background magnetic equilibrium. However, if the magnetic shear is strong and/or the domain is long, the twist can become so extreme that it fails to properly resolve the turbulence. In this work, we derive and implement the “non-twisting flux tube,” a local simulation domain that remains rectangular at all parallel locations. Convergence and runtime tests indicate that it can calculate the heat flux more efficiently than the conventional flux tube. For one test case, it was 30 times less computationally expensive and we found no case for which it was more expensive. It is most advantageous when the magnetic shear is high and the domain includes at least two regions of turbulent drive (e.g. stellarator simulations, pedestal simulations, tokamak simulations with several poloidal turns). Additionally, it more accurately models the inboard midplane when the magnetic shear is large. Lastly, we show how the non-twisting flux tube can be generalized to allow further optimization and control of the simulation domain.

1. Introduction

In magnetic confinement fusion devices, plasma turbulence is usually an important, if not the dominant mechanism in determining the energy confinement time [1, 2]. Turbulence ejects significant amounts of energy from the plasma, which then must be replaced using heating systems and large amounts of external electricity. Thus, it is vital to understand turbulent transport in order to enable a fusion power plant to generate net electricity at a competitive price.

The most successful approach to understand and predict turbulent transport is gyrokinetics, a high-fidelity kinetic model of the plasma [3, 4, 5, 6]. Gyrokinetics has been rigorously derived from first principles using an asymptotic expansion of the Fokker-Planck and Maxwell’s equations. This expansion makes the problem much more computationally tractable because it removes one velocity dimension from the model and, more importantly, eliminates a fast timescale — the particle gyration around the magnetic field line. The crucial expansion parameter of gyrokinetics is $\rho_* \equiv \rho_i/a \ll 1$, the ratio of the ion gyroradius ρ_i to the plasma minor radius a . Indeed, present-day

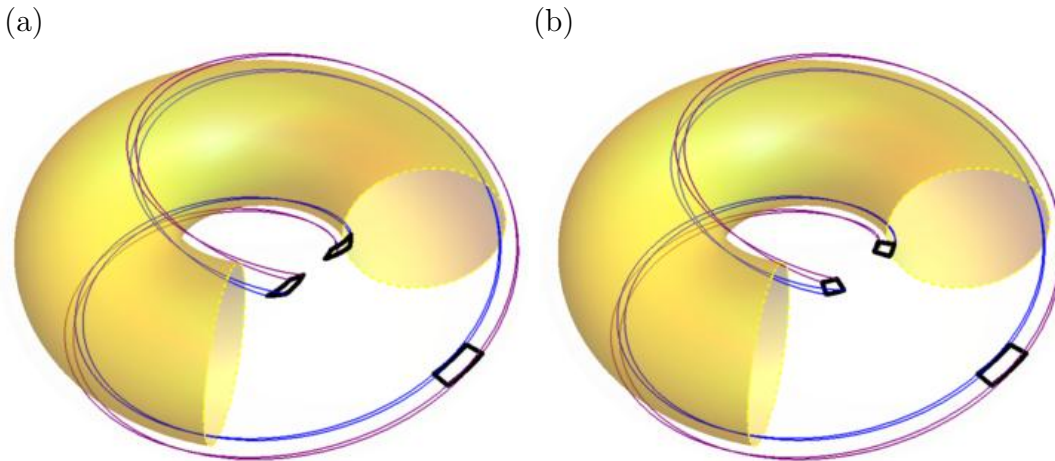


Figure 1. The boundaries of two local simulation domains (thin blue and purple lines): (a) a conventional flux tube and (b) a non-twisting flux tube. Both are one poloidal turn long. Note that at the outboard midplane, both flux tubes have a rectangular cross-section (thick black). However, away from the outboard midplane the cross-section of the conventional flux tube is twisted into a parallelogram, while the non-twisting flux tube remains rectangular. Also shown is the central flux surface of the flux tube (transparent yellow) with a toroidal wedge removed for visual clarity.

large fusion devices [7] can have $\rho_* \approx 1/300$ and ρ_* is anticipated to be even smaller in future high-performance devices [8]. Thus, we can be confident that our asymptotic expansion does not sacrifice much accuracy, making gyrokinetics one of the most reliable tools to simulate plasma turbulence in the core of fusion plasmas.

Despite its simplifications, gyrokinetic simulations remain near the limit of computational feasibility, with a single calculation typically requiring tens of thousands of CPU-hours and, hence, a supercomputer. To make simulations more accessible and practical, it is important to minimize the computational cost as much as possible. For grid-based codes, one straightforward way to do this is to minimize the number of grid points in the computational domain. This is a powerful motivation for so-called “local” simulations, which employ a reduced domain called a “flux tube” [9] (see figure 1(a)). The flux tube is designed to exploit the physical properties of the turbulent eddies in order to minimize the number of grid points required. This is accomplished by carefully choosing both the *locations* of grid points within the flux tube as well as the overall *shape* of the flux tube.

The *locations* of the grid points are chosen to be field-aligned, meaning that two of the three spatial coordinates are constant along the magnetic field lines. This helps because it matches the underlying structure of the turbulence. Due to the strong magnetic field, plasma turbulence in fusion devices is very anisotropic — individual turbulent eddies are very extended along the magnetic field lines, but are only a few gyroradii wide perpendicular to the field. Thus, the variation perpendicular to the magnetic field line has a small spatial scale, while the variation along the field line has a large spatial scale. By constructing a field-aligned grid, we take advantage of this and

can get by with a much larger distance between grid points in the parallel direction. If our grid was not field-aligned we would require a grid spacing comparable to the particle gyroradius in all three spatial directions, instead of just two.

Additionally, it is helpful if the *shape* of the flux tube mimics the turbulent structures it seeks to model. This allows us to shrink our domain as much as possible and best utilize its volume. Thus, the flux tube is very elongated along a central field line and narrow across it. Specifically, as shown in figure 1(a), it is typical for the domain to have a rectangular cross-section at the outboard midplane. Then, the four corners of the domain are fixed to four field lines, which determines how the cross-sectional shape changes along the parallel direction. This means that, due to magnetic shear, the domain twists into a parallelogram at other parallel locations.

With such a narrow domain in the minor radial direction (i.e. across flux surfaces), the statistical properties of the turbulence can be assumed to be identical on both radial boundaries. For simplicity, instead of using *statistical* periodicity, one can apply *exact* periodicity between the boundaries as long as the domain is wider than a few turbulent correlation lengths. The same can be done with the toroidal boundaries. Because of this narrow periodic domain, flux tube codes can employ a Fourier representation of the turbulence in the perpendicular plane. Such a representation is advantageous because it respects the boundary conditions by construction and, through use of the three-halves rule [10, 11], can cleanly avoid numerical aliasing issues that arise from quadratic nonlinear coupling. The parallel boundaries can also use periodicity, but it is more complicated — to ensure the parallel boundaries have statistically identical turbulence, the domain must be an integer number of poloidal turns long (i.e. the two ends must share the same poloidal location). This is because quantities like the strength and curvature of the magnetic field change the properties of turbulence and vary significantly with poloidal location. Additionally, implementing the parallel boundary condition while properly accounting for magnetic shear requires care [9, 12], as will be discussed. Regardless, the conventional flux tube is an elegant domain that appears to minimize the number of grid points, while still allowing for a physical treatment of turbulent eddies. However, the point of this paper is to demonstrate that further improvement is possible, thereby enabling more efficient numerical simulations. Specifically, we will formulate a simulation domain that does not twist due to the effect of magnetic shear, which we call the “non-twisting flux tube” and is shown in figure 1(b).

The idea of minimizing the twist of the simulation domain has been investigated before. Most recently, Watanabe *et al.* [13] devised the “flux tube train” — a domain formed by several conventional flux tubes coupled together into a chain. This enables a simulation domain to be several poloidal turns long without becoming extremely twisted. Normally, the twist would accrue along the entire length of the domain, but in the flux tube train it is reset between each individual flux tube (which can be just one poloidal turn long). However, the present paper is most closely related to the “shifted metric” procedure originally described in reference [14] and subsequently revisited in reference [11]. In the shifted metric procedure, one takes the discretized equations used

by a turbulence code and performs a spatial shift of the metric coefficients (also known as the geometric coefficients). Specifically, each grid point is shifted in the toroidal direction by an amount that can depend on both the radial location and the location along the field line. Using such discrete shifts, one can cancel the twisting caused by the magnetic shear at each grid point. This enables the cross-section of a flux tube to stay approximately rectangular, while still maintaining a field-aligned grid. For reasons that will be explained later, this was theorized to provide a more efficient computational domain, especially when the magnetic shear is large. Such a procedure was first applied to a simplified fluid model of turbulence in a domain with Dirichlet radial boundary conditions [14]. Later, it was implemented in a gyrokinetic code, again with Dirichlet radial boundary conditions [11]. Both works investigated the applicability of the shifted metric procedure to a flux tube with the usual periodic radial boundary condition, but concluded that the procedure was inapplicable or infeasible when using a Fourier representation in the perpendicular plane.

In this paper, we will present a novel derivation that uses a coordinate system transformation of the gyrokinetic equations *before* they are discretized. This gives a different perspective on the non-twisting flux tube and its boundary conditions, which will reveal a straightforward way to implement it while maintaining radial periodicity and a Fourier representation in the perpendicular plane. Aspects of its implementation are conceptually similar to the improved “wavevector-remap” scheme used to model $E \times B$ flow shear [15, 16] as magnetic shear twists the domain with parallel location analogously to how flow shear twists the domain with time. Flow shear will be omitted in this work for simplicity, but can be added to the non-twisting flux tube simply by including new terms analogously to the magnetic shear. The coordinate system discussed in this work may be useful for many purposes (e.g. global gyrokinetic simulations [11], fluid simulations [14]), but this paper will focus exclusively on its application to local gyrokinetic simulations. It has the potential to be most helpful for simulations with high global magnetic shear (e.g. near the tokamak separatrix), high local magnetic shear (e.g. stellarators), or many poloidal turns (e.g. turbulence with long parallel correlation lengths).

In section 2, we will review the analytic derivation of the conventional flux tube from reference [9] in order to provide appropriate context. Then, section 3 will present the novel derivation of the non-twisting flux tube, together with how it relates to both the conventional flux tube and the shifted metric approach. In section 4, we will benchmark the implementation of the non-twisting flux tube in the gyrokinetic code GENE [17, 18] and section 5 presents numerical results showing its computational performance relative to the conventional flux tube. Next, section 6 details how the non-twisting flux tube can be generalized to give even more flexibility in structuring the spatial grid. Lastly, section 7 offers some concluding thoughts.

2. Analytic derivation of the conventional flux tube

A conventional flux tube [9] uses a field-aligned coordinate system given by (x, y, χ) . The flux surface label $x = r - r_0$ is the minor radial coordinate r relative to the center of the flux tube r_0 ,

$$y(x, \zeta, \chi) \equiv \pm C_y (\zeta - q(x)\chi) \quad (1)$$

is the binormal coordinate (which labels the different field lines within the flux surfaces), and χ is the straight-field line poloidal angle (see Appendix A). Here ζ is the toroidal angle and $q(x)$ is the safety factor. The normalization constant C_y will be left arbitrary, though it is often taken to be either $C_y = r_0/q_0$ or $C_y = (1/B_r)d\psi/dr$ (where $q_0 \equiv q(0)$, ψ is the poloidal flux, and B_r is a reference value of the magnetic field). Note that the sign in the definition of y depends on what coordinate system convention is used. For the default coordinate system used by the GENE code, the lower sign (e.g. the minus sign in equation (1)) should be taken in all formulas in this paper. For the opposite convention (e.g. that used by the GS2 gyrokinetic code [19, 20]), the upper sign should be taken in all formulas.

In local simulations, due to the narrow domain, the safety factor profile is linearized about the center of the flux tube $x = 0$, giving

$$q(x) = q_0 + \left. \frac{dq}{dx} \right|_{x=0} x. \quad (2)$$

This enables us to rewrite equation (1) as

$$y(x, \zeta, \chi) = \pm C_y \zeta \mp C_y q_0 \chi \mp \hat{s} \chi x. \quad (3)$$

by introducing the global magnetic shear, $\hat{s} \equiv C_y dq/dx|_{x=0}$. Note that this is a non-standard definition of \hat{s} if $C_y \neq r_0/q_0$. It is in this (x, y, χ) coordinate system that we apply the perpendicular periodic boundary conditions [9]. In the radial direction, the real space electrostatic potential $\bar{\phi}$ must follow

$$\bar{\phi}(x + L_x, y, \chi) = \bar{\phi}(x, y, \chi), \quad (4)$$

while the binormal boundary condition is

$$\bar{\phi}(x, y + L_y, \chi) = \bar{\phi}(x, y, \chi), \quad (5)$$

where L_x and L_y are the widths of the domain in the radial and binormal directions respectively. The parallel boundary condition is called the ‘‘twist-and-shift’’ condition [9], which is more complicated as it must account for the effect of magnetic shear. To appropriately maintain the twisting effect of magnetic shear across the boundary, poloidal periodicity should be applied at constant toroidal angle ζ , according to

$$\bar{\phi}(x, y(x, \zeta, \chi + 2\pi N_{\text{pol}}), \chi + 2\pi N_{\text{pol}}) = \bar{\phi}(x, y(x, \zeta, \chi), \chi), \quad (6)$$

where $N_{\text{pol}} \in \mathbb{Z}_+$ must be a positive integer to ensure that the parallel boundaries share the same poloidal location and have statistically identical turbulence. Using equation (3), we can see that $y(x, \zeta, \chi + 2\pi N_{\text{pol}}) = y(x, \zeta, \chi) \mp 2\pi N_{\text{pol}} C_y q_0 \mp 2\pi N_{\text{pol}} \hat{s}x$. Thus, equation (6) becomes

$$\bar{\phi}(x, y \mp 2\pi N_{\text{pol}} \hat{s}x, \chi + 2\pi N_{\text{pol}}) = \bar{\phi}(x, y, \chi), \quad (7)$$

where we have assumed that $\mp(2\pi N_{\text{pol}} C_y / L_y) q_0$ is very close to some integer $N_q \in \mathbb{Z}$ and applied binormal periodicity N_q times using equation (5). This assumption is acceptable because equation (1) shows that $C_y / L_y = 1 / L_\zeta$ at constant x and χ , where L_ζ is the width of the flux tube in toroidal angle such that $L_\zeta = 2\pi$ corresponds to the full toroidal domain. Thus, so long as the flux tube is very small compared to the full flux surface, $2\pi N_{\text{pol}} C_y / L_y$ will be a very big number and a negligibly small change in q_0 will ensure that our assumption holds.

Importantly, by evaluating the parallel boundary condition of equation (7) at $x \rightarrow x + L_x$, applying the radial boundary condition of equation (4) to both sides, and finally applying the parallel boundary condition to the left side we find

$$\bar{\phi}(x, y \mp 2\pi N_{\text{pol}} \hat{s}L_x, \chi) = \bar{\phi}(x, y, \chi). \quad (8)$$

This shows that the combination of parallel and radial boundary conditions will introduce artificial correlations between different y locations unless

$$L_x = \frac{N_{\text{asp}} L_y}{2\pi N_{\text{pol}} |\hat{s}|} \quad (9)$$

holds for a positive integer $N_{\text{asp}} \in \mathbb{Z}_+$, where the subscript ‘‘asp’’ indicates that it controls the domain aspect ratio. Fulfilling this constraint resolves the problem because the offending term in equation (8) can be eliminated by applying binormal periodicity N_{asp} times using equation (5). This constraint discretizes the aspect ratio of the domain. Note that it is a general consequence of the flux tube boundary conditions and must be satisfied regardless of whether one uses a real space or Fourier-space representation.

For the reasons outlined in the introduction, local gyrokinetic codes usually employ a Fourier representation in the perpendicular plane. Thus, we will use the Fourier-space electrostatic potential ϕ , defined by

$$\bar{\phi}(x, y, \chi) = \sum_{k_x, k_y} \phi(k_x, k_y, \chi) e^{ik_x x + ik_y y}. \quad (10)$$

According to Fourier’s theorem, we see that the binormal and radial boundary conditions of equations (4) and (5) require

$$k_x = \frac{2\pi}{L_x} m \quad (11)$$

$$k_y = \frac{2\pi}{L_y} n, \quad (12)$$

where $m \in \mathbb{Z}$ and $n \in \mathbb{Z}$ are integers. Note n is analogous to the toroidal mode number in the restricted domain of the flux tube. Thus, the perpendicular boundary conditions can be satisfied in a simulation simply by constructing the Fourier coordinate grids appropriately. Typically they are constructed according to

$$k_x \in \frac{2\pi}{L_x} m \quad \text{for all } m \in \left[-\frac{N_x - 1}{2}, \frac{N_x - 1}{2} \right] \quad (13)$$

$$k_y \in \frac{2\pi}{L_y} n \quad \text{for all } n \in [0, N_y - 1], \quad (14)$$

where N_x and N_y are the number of considered radial and binormal modes respectively and we have assumed N_x is odd for simplicity. Additionally, we have omitted the negative k_y values as they can be recovered using the reality condition $\phi(k_x, k_y, \chi) = \phi^*(-k_x, -k_y, \chi)$, where $*$ indicates the complex conjugate.

The parallel boundary condition is more complicated. After substituting the Fourier representation of equation (10), equation (7) becomes

$$\phi(k_x \pm 2\pi N_{\text{pol}} k_y \hat{s}, k_y, \chi + 2\pi N_{\text{pol}}) = \phi(k_x, k_y, \chi). \quad (15)$$

This is enforced by appropriately coupling different k_x modes across the ends of the flux tube. For these modes to always line up properly, the smallest possible jump in the radial wavenumber (i.e. $2\pi N_{\text{pol}} k_{y,\text{min}} |\hat{s}|$) must be equal to a multiple of the radial grid spacing (e.g. $N_{\text{asp}} k_{x,\text{min}}$). This condition turns out to be identical to the domain aspect ratio condition of equation (9) (since the minimum radial and binormal wavenumbers are $k_{x,\text{min}} = 2\pi/L_x$ and $k_{y,\text{min}} = 2\pi/L_y$ respectively). The only exceptions are the largest values of k_x , which are coupled to $\phi = 0$ because the corresponding modes are outside of the considered k_x grid.

In the (x, y, χ) coordinate system, the Fourier-space electrostatic gyrokinetic model (neglecting collisions and background flow) is given by the gyrokinetic equation [21]

$$\begin{aligned} \frac{\partial h_s}{\partial t} + v_{\parallel} \hat{b} \cdot \vec{\nabla} \chi \left. \frac{\partial h_s}{\partial \chi} \right|_{k_x, k_y} + i \vec{v}_{ds} \cdot (k_x \vec{\nabla} x + k_y \vec{\nabla} y) h_s + a_{s\parallel} \frac{\partial h_s}{\partial v_{\parallel}} \\ \mp \frac{1}{JB} \{h_s, \phi J_0(k_{\perp} \rho_s)\} = \frac{Z_s e F_{Ms}}{T_s} \frac{\partial \phi}{\partial t} J_0(k_{\perp} \rho_s) \mp i \frac{k_y}{JB} \phi J_0(k_{\perp} \rho_s) \frac{\partial F_{Ms}}{\partial x} \end{aligned} \quad (16)$$

and the quasineutrality equation

$$\phi = 2\pi \left(\sum_s \frac{Z_s^2 e^2 n_s}{T_s} \right)^{-1} \sum_s \frac{Z_s e B}{m_s} \int dv_{\parallel} \int d\mu J_0(k_{\perp} \rho_s) h_s, \quad (17)$$

where the perpendicular wavenumber is

$$k_{\perp} = \sqrt{k_x^2 |\vec{\nabla} x|^2 + 2k_x k_y \vec{\nabla} x \cdot \vec{\nabla} y + k_y^2 |\vec{\nabla} y|^2} \quad (18)$$

and the nonlinear term in Fourier-space is

$$\{h_s, \phi J_0(k_\perp \rho_s)\} = \sum_{k'_x, k'_y} (k'_x k''_y - k''_x k'_y) h'_s \phi'' J_0(k''_\perp \rho_s) \quad (19)$$

with $k''_x = k_x - k'_x$ and $k''_y = k_y - k'_y$. The unknowns are the Fourier-analyzed electrostatic potential ϕ and the Fourier-analyzed non-adiabatic portion of the distribution function $h_s \equiv \delta f + Z_s e \phi F_{Ms} / T_s$. The coordinates are the time t , the radial wavenumber k_x , the binormal wavenumber k_y , the straight-field line poloidal angle χ , the parallel velocity v_\parallel , and the magnetic moment $\mu \equiv m_s v_\perp^2 / (2B)$. The species subscript $s \in \{i, e\}$ indicates ions or electrons respectively. Here \hat{b} is the magnetic field unit vector, B is the strength of the magnetic field, \vec{v}_{ds} is the magnetic drift velocity, $a_{s\parallel}$ is the parallel acceleration from the magnetic mirror force, $J_0(\dots)$ is the 0th order Bessel function of the first kind, Z_s is the species charge number, e is the proton electric charge, F_{Ms} is the Maxwellian velocity distribution, δf is the turbulent portion of the Fourier-analyzed perturbed distribution function, T_s is the species temperature, n_s is the species density, m_s is the species mass, and $J^{-1} \equiv \mp (\vec{\nabla}x \times \vec{\nabla}y) \cdot \hat{b}$ is closely related to the coordinate system Jacobian and is defined such that it is positive for both sign conventions. Lastly, to reduce the computational cost, most codes compute the nonlinear term in real space according to

$$\begin{aligned} \{h_s, \phi J_0(k_\perp \rho_s)\} = & \frac{1}{N_x (2N_y - 1)} \sum_{x,y} \left[\left(\sum_{k'_x, k'_y} k'_x h'_s e^{ik'_x x + ik'_y y} \right) \left(\sum_{k''_x, k''_y} k''_y \phi'' J_0(k''_\perp \rho_s) e^{ik''_x x + ik''_y y} \right) \right. \\ & \left. - \left(\sum_{k'_x, k'_y} k'_y h'_s e^{ik'_x x + ik'_y y} \right) \left(\sum_{k''_x, k''_y} k''_x \phi'' J_0(k''_\perp \rho_s) e^{ik''_x x + ik''_y y} \right) \right] e^{-ik_x x - ik_y y} \end{aligned} \quad (20)$$

by using

$$\phi(k_x, k_y, \chi) = \frac{1}{N_x (2N_y - 1)} \sum_{x,y} \bar{\phi}(x, y, \chi) e^{-ik_x x - ik_y y}, \quad (21)$$

the inverse discrete Fourier transform of equation (10). Note the (double) primed quantities indicate they are evaluated at the (double) primed wavenumbers. In the calculation of the nonlinear term in real space, high wavenumber perturbations may be generated that then get fictitiously mapped to lower wavenumbers in the return Fourier transform. In the code, this is avoided by using the three-halves rule [10, 11]. First, before performing the Fourier transform to real space, the Fourier-space quantities are copied to an expanded wavenumber grid that is at least 3/2 times bigger in both perpendicular directions. These new high wavenumber modes are given zero amplitude. Then, the full calculation of the nonlinear term is performed according to equation (20) on this expanded grid. At the end, the final Fourier-space result is taken and the high wavenumber modes are discarded to return to the original wavenumber grid. Thus, in

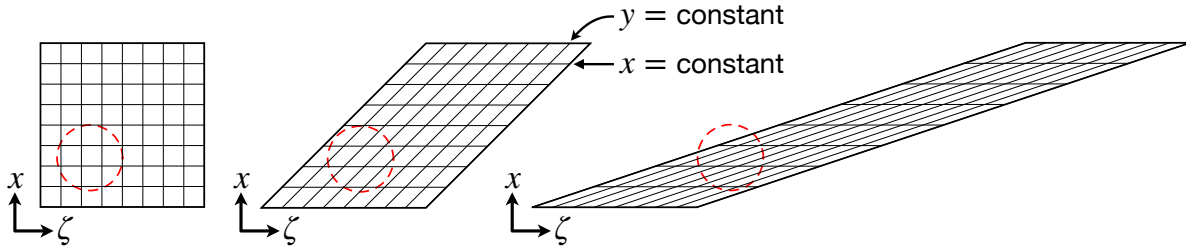


Figure 2. The (x, y) coordinate system grid at three χ locations along a conventional flux tube. We see from the example ion gyroradius (red dashed circle) that, as the domain becomes increasingly sheared, ions will more effectively average over any of the turbulent perturbations allowed in the domain.

GENE the summations in equation (21) are really taken over an expanded wavenumber grid, $x \in 2mL_x / (3(N_x + 1))$ for all $m \in [0, 3(N_x + 1)/2 - 1]$, and $y \in nL_y / (3N_y)$ for all $n \in [0, 3N_y - 1]$.

Now let us consider the implications of this (x, y, χ) coordinate system (or equivalently (k_x, k_y, χ)). A regularly-spaced rectangular grid in these coordinates produces a simulation domain that becomes sheared into a parallelogram as it extends along the field lines due to the effect of magnetic shear. This has important consequences for what type of turbulence can be modeled in the domain. For certain conditions, the twist of the coordinate system is appropriate. If the three-dimensional shape of a turbulent eddy is primarily an extrusion along field lines of its two-dimensional cross-section at the outboard midplane, the coordinate system will model it efficiently. However, as an eddy extends along the field lines, magnetic shear twists it and decreases the spatial scales of the structure. If the magnetic shear is large or the eddy is very extended along the magnetic field, the spatial scales of the eddy will become very small as is reflected in the extreme tilt of the parallelogram (see figure 2). Such fine structures are strongly damped due to finite gyroradius effects, so it would be surprising if they were important. Yet, at these parallel locations the conventional flux tube is optimized to model them.

To see this mathematically, one can look at the Bessel functions in the gyrokinetic model, which represent the finite gyroradius effects. In the conventional flux tube, the radial grid is centered around $k_x = 0$ at all parallel locations, for which we see that

$$J_0(k_{\perp}\rho_s) = J_0\left(|k_y|\rho_s\left|\vec{\nabla}y\right|\right) \quad (22)$$

using equation (18). When one calculates the geometric coefficient

$$\vec{\nabla}y = \pm C_y \vec{\nabla}\zeta \mp C_y q_0 \vec{\nabla}\chi \mp \hat{s}\chi \vec{\nabla}x \quad (23)$$

at $x = 0$, we see that the third term has a secular dependence on the parallel location χ . Thus, as we move along the field line, $J_0(k_{\perp}\rho_s) \rightarrow 0$ for the central radial mode in the grid, indicating that it is damped. This means that, for large values of \hat{s} or long flux-tube domains, the radial wavenumber grid is centered around a mode that is strongly damped at most values of χ .

To summarize, the perpendicular grid of a conventional flux tube is optimized to model perturbations that are field line following. Away from $\chi = 0$, this is a distinct goal from efficiently modeling the perturbations that are least damped by finite gyroradius effects. The question then becomes empirical — does turbulence prioritize following field lines from the outboard midplane or does it prioritize minimizing the amount of finite gyroradius damping? The short answer is both, to some degree, and the balance depends on the specifics of a particular simulation. We will delay a more in depth answer until section 5 and instead define a grid that is optimized for turbulence that minimizes finite gyroradius damping — the non-twisting flux tube.

3. Analytic derivation of the non-twisting flux tube

In order to produce a flux tube that does not twist (e.g. figure 1(b)), we want to perform a coordinate system transformation that removes the effect of the magnetic shear — both global shear and local shear. This can be accomplished by defining a new “non-twisting” binormal coordinate Y such that $\vec{\nabla}x \cdot \vec{\nabla}Y = 0$ at all parallel locations. In other words, we want (x, Y) to be an orthogonal coordinate system. The effect of global shear is clearly contained in the third term of equation (3), but the local correction to the shear is more subtle. It manifests through the definition of the straight-field line poloidal angle χ (see Appendix A) in the second term of equation (3) because $\vec{\nabla}x \cdot \vec{\nabla}\chi \neq 0$. To remove all magnetic shear, we make the coordinate transformation [14, 22]

$$Y(x, y, \chi) \equiv y - \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2}x. \quad (24)$$

Using equation (23) and $\vec{\nabla}x \cdot \vec{\nabla}\zeta = 0$, this geometric factor can be calculated to be

$$\frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2} = \mp \hat{s}\chi \mp C_y q_0 \frac{\vec{\nabla}x \cdot \vec{\nabla}\chi}{|\vec{\nabla}x|^2}, \quad (25)$$

where the first term contains the global magnetic shear and the second term contains the local correction to the magnetic shear through the definition of χ (see equations (A.4) and (A.9)). Thus, we can rewrite equation (24) as

$$Y(x, \zeta, \chi) = \pm C_y \zeta \mp C_y q_0 \left(\chi - \frac{\vec{\nabla}x \cdot \vec{\nabla}\chi}{|\vec{\nabla}x|^2}x \right) \quad (26)$$

using equation (3). Importantly, we can directly substitute this expression to show that $\vec{\nabla}x \cdot \vec{\nabla}Y = 0$, remembering that the geometric coefficients are evaluated at $x = 0$ in the *local* gyrokinetic model.

To accomplish this coordinate system transformation, we will first transform the boundary conditions. The binormal boundary condition is applied at constant x and χ ,

so it is written as

$$\bar{\Phi}(x, Y(x, y + L_y, \chi), \chi) = \bar{\Phi}(x, Y(x, y, \chi), \chi). \quad (27)$$

From equation (24) we see that $Y(x, y + L_y, \chi) = Y(x, y, \chi) + L_y$, demonstrating that the transformed binormal boundary condition is simply

$$\bar{\Phi}(x, Y + L_y, \chi) = \bar{\Phi}(x, Y, \chi). \quad (28)$$

Note that we have introduced a new symbol for the electrostatic potential $\bar{\Phi}$. This is because it has a different functional dependence on its arguments according to

$$\bar{\Phi}(x, Y, \chi) = \bar{\phi}\left(x, Y + \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2}x, \chi\right) = \bar{\phi}(x, y(x, Y, \chi), \chi). \quad (29)$$

We also define a new distribution function \bar{H}_s from \bar{h}_s in an analogous fashion. More broadly, throughout the paper we will match the style of these symbols with the associated coordinate (i.e. lowercase $\bar{\phi}$ and ϕ are functions of lowercase y and k_x respectively, capital $\bar{\Phi}$ and Φ are functions of capital Y and K_x , script φ will be a function of script \mathcal{K}_x , etc.).

The parallel boundary condition must still be taken at constant toroidal angle ζ and x , according to

$$\bar{\Phi}(x, Y(x, \zeta, \chi + 2\pi N_{\text{pol}}), \chi + 2\pi N_{\text{pol}}) = \bar{\Phi}(x, Y(x, \zeta, \chi), \chi). \quad (30)$$

Equation (26) shows that $Y(x, \zeta, \chi + 2\pi N_{\text{pol}}) = Y(x, \zeta, \chi) \mp 2\pi N_{\text{pol}} C_y q_0$ because the geometric coefficients $\vec{\nabla}x \cdot \vec{\nabla}\chi$ and $|\vec{\nabla}x|^2$ are both 2π -periodic in χ . This implies that

$$\bar{\Phi}(x, Y, \chi + 2\pi N_{\text{pol}}) = \bar{\Phi}(x, Y, \chi), \quad (31)$$

again assuming that $(2\pi N_{\text{pol}} C_y / L_y) q_0$ is very close to an integer and applying binormal periodicity as we did for the conventional flux tube. This parallel boundary condition is simpler than equation (7), which is intuitive as a non-twisting flux tube has an identical rectangular cross-section at both parallel boundaries making it straightforward to copy turbulent structures across it.

Instead the complications from magnetic shear appear in the radial boundary condition. It is applied at constant y and χ , according to

$$\bar{\Phi}(x + L_x, Y(x + L_x, y, \chi), \chi) = \bar{\Phi}(x, Y(x, y, \chi), \chi). \quad (32)$$

Using equation (24) we see that $Y(x + L_x, y, \chi) = Y(x, y, \chi) - (\vec{\nabla}x \cdot \vec{\nabla}y)L_x / |\vec{\nabla}x|^2$, so the radial boundary condition in the new coordinates is

$$\bar{\Phi}\left(x + L_x, Y - \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2}L_x, \chi\right) = \bar{\Phi}(x, Y, \chi). \quad (33)$$

We see that a new term, $(\vec{\nabla}x \cdot \vec{\nabla}y)L_x/|\vec{\nabla}x|^2$, appears, which is needed to properly pair up field lines across the radial boundary. In the conventional flux tube, as one moved along the domain in χ , the domain cross-section twisted with the field lines, so the field lines always stayed at the same position relative to the boundaries. In the new coordinates, all the field lines not at $x = 0$ will shift relative to the binormal boundaries as you move in χ . At some parallel location(s) they may exit through the binormal boundary of the domain and immediately re-enter through the opposite binormal boundary. Note that this means the “non-twisting flux tube” is technically not a true “flux tube” according to some definitions. However, we think the terminology is appropriate as the domain still changes area to retain a constant magnetic flux through its cross-section. Since the field lines move relative to the domain boundaries, the new term in equation (33) is needed to properly maintain field line identity across the radial boundary, accounting for the local magnetic shear. If radial periodicity identifies two field lines on opposite radial boundaries to be the same at a particular parallel location, the two field lines must remain matched up at all other parallel locations. Otherwise a particle could move to a different field line by moving purely in the parallel direction, which is unphysical.

Field line identity must also be preserved across the parallel boundary condition, which can be checked by combining the parallel and radial boundary conditions of equations (31) and (33). Specifically, we evaluate the radial boundary condition at $\chi \rightarrow \chi + 2\pi N_{\text{pol}}$, apply the parallel boundary condition to both sides, and then apply the radial boundary condition to the left side. In doing this, it is important to note that equation (25) shows

$$\left. \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2} \right|_{\chi+2\pi N_{\text{pol}}} - \left. \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2} \right|_{\chi} = \mp 2\pi N_{\text{pol}} \hat{s}, \quad (34)$$

since $\vec{\nabla}x \cdot \vec{\nabla}\chi$ and $|\vec{\nabla}x|^2$ are both 2π -periodic in χ . Equation (34) states that after an integer number of poloidal turns the local correction to the shear integrates to zero and only the twisting from global shear remains. Using this fact, we find the combination of parallel and radial boundary conditions in the new coordinate system gives

$$\bar{\Phi}(x, Y \pm 2\pi N_{\text{pol}} \hat{s} L_x, \chi) = \bar{\Phi}(x, Y, \chi). \quad (35)$$

This implies that the domain will have fictitious correlations between different Y positions unless equation (28) can be used to eliminate the offending term. Enforcing this discretizes the aspect ratio of the domain according to equation (9) just like the conventional flux tube, as is expected.

We emphasize that the boundary conditions of the non-twisting flux tube (i.e. equations (28), (31), and (33)) are consistent with those of the conventional flux tube (i.e. equations (4), (5), and (7)). This can be checked by taking a boundary condition from the non-twisting flux tube, applying equation (29) to the left and right sides, and then substituting equation (24) to replace Y with y .

As with the conventional flux tube, we would like to use a Fourier representation in the perpendicular plane. Because the binormal boundary condition in real space (i.e. equation (28)) has the same form as in the conventional flux tube (i.e. equation (5)), it will be satisfied using the standard Fourier representation

$$\bar{\Phi}(x, Y, \chi) = \sum_{k_y} \hat{\Phi}(x, k_y, \chi) e^{ik_y Y} \quad (36)$$

with the wavenumbers discretized according to equation (12). However, the radial boundary condition of equation (33) is more complicated than in the conventional flux tube. To find the allowed radial wavenumbers, we first Fourier analyze using equation (36) to get the radial boundary condition into the form

$$\hat{\Phi}(x + L_x, k_y, \chi) e^{-ik_y \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2} L_x} = \hat{\Phi}(x, k_y, \chi). \quad (37)$$

Then, using Floquet's theorem we make the substitution

$$\hat{\Phi}(x, k_y, \chi) = P(x, k_y, \chi) e^{ik_y \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2} x} \quad (38)$$

and see that the radial boundary condition becomes simple periodicity, $P(x + L_x, k_y, \chi) = P(x, k_y, \chi)$. This is satisfied by writing $P(x, k_y, \chi)$ as a standard Fourier series in x with wavenumber discretized according to equation (11). Thus, by substituting this result into equation (38) and then equation (36), we find the final form of

$$\bar{\Phi}(x, Y, \chi) = \sum_{K_x, k_y} \Phi(K_x, k_y, \chi) e^{iK_x x + ik_y Y}, \quad (39)$$

where the allowed radial wavenumbers are given by

$$K_x = \frac{2\pi}{L_x} m + k_y \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2} \quad (40)$$

for any integer $m \in \mathbb{Z}$. Comparing this with equation (11), we see that the new radial wavenumber is defined by

$$K_x \equiv k_x + k_y \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2}, \quad (41)$$

which means that $K_x x + k_y Y = k_x x + k_y y$. Equation (40) indicates that the allowed radial wavenumbers now depend on the poloidal angle and the binormal wavenumber, which might seem strange. However, this is *not* because something physical has changed about the situation. We will see that it is purely a semantic consequence of our coordinate system transform.

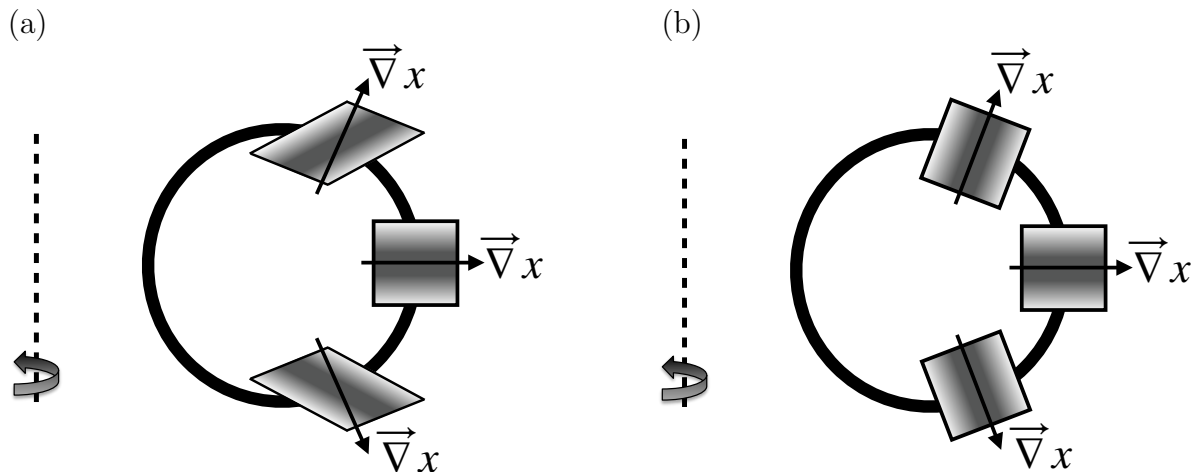


Figure 3. Three perpendicular cross-sections at different poloidal angles for (a) a cartoon conventional flux tube with a $(k_x = 0, k_y = 2\pi/L_y)$ perturbation and (b) a cartoon non-twisting flux tube with a $(K_x = 0, k_y = 2\pi/L_y)$ perturbation. Note that the three perturbations depicted in (a) are linearly coupled, unlike those in (b). Also shown is the central $x = 0$ flux surface of the flux tube (thick black), which is circular, and the axis of toroidal symmetry (dotted black).

Equation (41) defines our coordinate system transformation in Fourier-space just as equation (24) defined our coordinate system transform in real space, so it is important to understand the physical meaning of K_x relative to k_x . We will call K_x the *local* radial wavenumber and k_x the *ballooning* radial wavenumber. We see from equation (41) that the two wavenumbers are identical wherever $\vec{\nabla}x \cdot \vec{\nabla}y = 0$, such as at $\chi = 0$. The ballooning radial wavenumber k_x is defined by performing the Fourier analysis while holding y and χ constant, so it is the radial wavenumber of the turbulence *along lines of constant y and χ* . However, lines of constant y and χ are *not* perpendicular to the flux surfaces (i.e. lines of constant x) because the (x, y) coordinate system is not orthogonal (see equation (25) or figure 2). Thus, as shown in figure 3(a), $(k_x = 0, k_y \neq 0)$ does not everywhere refer to a Fourier mode that is constant in the $\vec{\nabla}x$ direction, even for circular flux surfaces. Lines of constant y and χ only run in the the $\vec{\nabla}x$ direction at the parallel location where the flux tube is rectangular, typically at the outboard midplane. Thus, away from the midplane the Fourier mode with zero variation in the $\vec{\nabla}x$ direction actually has $k_x \neq 0$. Instead $k_x = 0$ refers to the Fourier mode that, if you traced it along field lines back to the outboard midplane, would have zero variation in the $\vec{\nabla}x$ direction. Hence, its name — the ballooning radial wavenumber. In contrast, the local radial wavenumber K_x is defined by performing the Fourier analysis while holding Y constant. Since we defined Y such that $\vec{\nabla}x \cdot \vec{\nabla}Y = 0$, lines of constant Y run in the $\vec{\nabla}x$ direction at all parallel locations. Thus, the $K_x = 0$ Fourier mode never varies in the $\vec{\nabla}x$ direction, as is shown in figure 3(b). Note that the $k_x = 0$ Fourier modes at each parallel location are linearly coupled, while the $K_x = 0$ Fourier modes are not.

The final Fourier-space boundary condition is in the parallel direction. By using

equation (39) to perform a Fourier analysis of equation (31), we find that it is

$$\Phi(K_x, k_y, \chi + 2\pi N_{\text{pol}}) = \Phi(K_x, k_y, \chi). \quad (42)$$

Despite this simple form, it is not appropriate to use a Fourier representation in χ because the K_x grid varies with χ as do the geometric coefficients in the gyrokinetic equations. Like before, we can combine the parallel and radial boundary conditions of equations (40) and (42) to derive a constraint on the aspect ratio of the domain. Using equations (12) and (34), we find that the usual condition of equation (9) still holds and ensures that the same values of K_x exist at both χ and $\chi + 2\pi N_{\text{pol}}$. Before we move on, we note that equation (42) is consistent with the parallel boundary condition in the conventional flux tube (i.e. equation (15)). To prove this we must first relate the functional forms of Φ and ϕ by substituting equations (10) and (39) into equation (29) to see

$$\Phi(K_x, k_y, \chi) = \phi\left(K_x - k_y \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2}, k_y, \chi\right) = \phi(k_x(K_x, k_y, \chi), k_y, \chi). \quad (43)$$

Then, we apply this to the left and right sides of equation (42) and substitute equation (41) to arrive at equation (15).

Now that we know the boundary conditions and allowed modes, we will construct the coordinate system grids. We could determine our K_x grid by taking equation (40) and choosing $m \in [-(N_x - 1)/2, (N_x - 1)/2]$ at all parallel locations. However, this will accomplish nothing new. We would have created the exact same grid for the exact same physical situation as with the conventional flux tube, just with different labels on the grid points. In other words, the new K_x grid would still be centered around $k_x = 0$. Instead, we would like to lay down a grid centered around $K_x = 0$. However, in general it is not possible to do this exactly. In order to respect the radial boundary condition, we must adhere to equation (40), which prohibits a grid point at exactly $K_x = 0$ at most parallel locations. Thus, we must be content to center the local radial wavenumber grid around $K_x \approx 0$, instead of exactly 0. This can be accomplished by first finding the allowed mode number that comes closest to $K_x = 0$, given by

$$m_0(k_y, \chi) = -\text{NINT}\left[\frac{L_x}{2\pi} k_y \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2}\right], \quad (44)$$

where $\text{NINT}[\dots]$ is a function that takes a real number and returns the nearest integer. Using equation (44), we can construct the K_x grid around this mode number according to

$$K_x \in \frac{2\pi}{L_x} (m + m_0(k_y, \chi)) + k_y \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2} \quad \text{for all } m \in \left[-\frac{N_x - 1}{2}, \frac{N_x - 1}{2}\right]. \quad (45)$$

Thus, we see that the K_x grid changes for different binormal wavenumbers and at different parallel locations, but always stays as close as possible to $K_x = 0$. The k_y grid remains the same as in the conventional flux tube, specified according to equation (14).

We can now see the impact of this new K_x grid on how a gyrokinetic code models turbulence. In our previous discussion concerning figure 2, we saw that the conventional flux tube prioritized turbulence that follows field lines rather than turbulence with minimal finite gyroradius damping. This could be seen by noting the secular dependence on $\hat{s}\chi$ in the argument of the Bessel function for $k_x = 0$, the central radial mode in the simulation domain (see equations (22) and (23)). Now we can repeat the exercise for $K_x = 0$, the central radial mode in the new coordinate system. By substituting the definitions of Y and K_x (i.e. equations (24) and (41)) into equation (18), we find that the perpendicular wavenumber becomes

$$K_{\perp} = \sqrt{K_x^2 |\vec{\nabla}x|^2 + k_y^2 |\vec{\nabla}Y|^2}, \quad (46)$$

which is similar to equation (18) except the cross term vanishes because $\vec{\nabla}x \cdot \vec{\nabla}Y = 0$. Thus, when we calculate the Bessel function for the central radial mode in the grid, $K_x \approx 0$, we find

$$J_0(K_{\perp}\rho_s) = J_0\left(|k_y|\rho_s \left|\vec{\nabla}Y\right|\right). \quad (47)$$

Using equation (26), this new geometric coefficient can be found to be

$$\vec{\nabla}Y = \pm C_y \vec{\nabla}\zeta \mp C_y q_0 \vec{\nabla}\chi \pm C_y q_0 \frac{\vec{\nabla}x \cdot \vec{\nabla}\chi}{|\vec{\nabla}x|^2} \vec{\nabla}x, \quad (48)$$

which has no secular dependence with χ . Thus, while the flux tube cross-section may expand and contract periodically due to the effect of flux surface shaping and toroidicity, it will not twist at all. This means the domain will include the modes that are least damped by finite gyroradius effects, regardless of the strength of the magnetic shear or the length of the simulation domain.

Lastly, the final step in changing our coordinate system is to transform the gyrokinetic model. Substituting equations (24), (41), and (43), we find that the gyrokinetic equation (i.e. equation (16)) becomes

$$\begin{aligned} \frac{\partial H_s}{\partial t} + v_{\parallel} \hat{b} \cdot \vec{\nabla}\chi \left. \frac{\partial H_s}{\partial \chi} \right|_{K_x - k_y \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2}, k_y} + i \vec{v}_{ds} \cdot \left(K_x \vec{\nabla}x + k_y \vec{\nabla}Y \right) H_s + a_{s\parallel} \frac{\partial H_s}{\partial v_{\parallel}} \\ \mp \frac{1}{JB} \{H_s, \Phi J_0(K_{\perp}\rho_s)\} = \frac{Z_s e F_{Ms}}{T_s} \frac{\partial \Phi}{\partial t} J_0(K_{\perp}\rho_s) \mp i \frac{k_y}{JB} \Phi J_0(K_{\perp}\rho_s) \frac{\partial F_{Ms}}{\partial x}. \end{aligned} \quad (49)$$

where H_s is defined analogously to the definition of Φ in equation (43). The nonlinear term is either

$$\{H_s, \Phi J_0(K_{\perp}\rho_s)\} = \sum_{K'_x, k'_y} (K'_x k''_y - K''_x k'_y) H'_s \Phi'' J_0(K''_{\perp}\rho_s) \quad (50)$$

using $K_x'' = K_x - K_x'$ and $k_y'' = k_y - k_y'$ or

$$\begin{aligned} \{H_s, \Phi J_0(K_\perp \rho_s)\} &= \frac{1}{N_x(2N_y - 1)} \sum_{x,Y} \left[\left(\sum_{K_x', k_y'} K_x' H_s' e^{iK_x' x + ik_y' Y} \right) \left(\sum_{K_x'', k_y''} k_y'' \Phi'' J_0(K_\perp'' \rho_s) e^{iK_x'' x + ik_y'' Y} \right) \right. \\ &\quad \left. - \left(\sum_{K_x', k_y'} k_y' H_s' e^{iK_x' x + ik_y' Y} \right) \left(\sum_{K_x'', k_y''} K_x'' \Phi'' J_0(K_\perp'' \rho_s) e^{iK_x'' x + ik_y'' Y} \right) \right] e^{-iK_x x - ik_y Y} \end{aligned} \quad (51)$$

using the new Fourier transform of

$$\Phi(K_x, k_y, \chi) = \frac{1}{N_x(2N_y - 1)} \sum_{x,Y} \bar{\Phi}(x, Y, \chi) e^{-iK_x x - ik_y Y}. \quad (52)$$

Importantly, in order to retain a field-aligned grid, we see that the parallel derivative in equation (49) must still be taken at constant $k_x = K_x - k_y \vec{\nabla} x \cdot \vec{\nabla} y / |\vec{\nabla} x|^2$. If we applied the chain rule in order to take the parallel derivative at constant K_x , we would no longer be able to use a coarse grid in χ . Quasineutrality remains the same as equation (17), except H_s and Φ are functions of K_x instead of k_x and the perpendicular wavenumber is given by equation (46). Thus, the complete non-twisting flux tube in Fourier-space is defined by equations (14), (17) (using Φ , H_s , and K_\perp), (42), (45), (46), (49), and (51). The new geometric coefficients and equilibrium quantities that appear in these equations are simple to relate to those already calculated for the conventional flux tube. Using equation (24), we find

$$\vec{\nabla} x \cdot \vec{\nabla} Y = 0 \quad (53)$$

$$|\vec{\nabla} Y|^2 = |\vec{\nabla} y|^2 - \frac{(\vec{\nabla} x \cdot \vec{\nabla} y)^2}{|\vec{\nabla} x|^2} \quad (54)$$

$$\vec{\nabla} Y \cdot \vec{\nabla} \chi = \vec{\nabla} y \cdot \vec{\nabla} \chi - \frac{\vec{\nabla} x \cdot \vec{\nabla} y}{|\vec{\nabla} x|^2} \vec{\nabla} x \cdot \vec{\nabla} \chi \quad (55)$$

$$J^{-1} \equiv \mp (\vec{\nabla} x \times \vec{\nabla} Y) \cdot \hat{b} = \mp (\vec{\nabla} x \times \vec{\nabla} y) \cdot \hat{b} \quad (56)$$

$$\vec{B} = \frac{1}{C_y} \frac{d\psi}{dx} \vec{\nabla} x \times \vec{\nabla} y = \frac{1}{C_y} \frac{d\psi}{dx} \vec{\nabla} x \times \vec{\nabla} Y \quad (57)$$

$$\vec{\nabla} B = \left(\frac{\partial B}{\partial x} + \frac{\vec{\nabla} x \cdot \vec{\nabla} y}{|\vec{\nabla} x|^2} \frac{\partial B}{\partial y} \right) \vec{\nabla} x + \frac{\partial B}{\partial y} \vec{\nabla} Y + \frac{\partial B}{\partial \chi} \vec{\nabla} \chi. \quad (58)$$

Before we move on, we have two comments that are important to provide perspective on the non-twisting flux tube. First, we believe that the crucial novel insight of this paper is specifying the radial wavenumber grid according to equations

Parameter	Value	Parameter	Value
Minor radius of flux tube, x_0/a	0.54	Major radius, R_0/a	3.0
Safety factor, q_0	1.4	Magnetic shear, \hat{s}	0.8
Temperature gradient, a/L_{Ts}	2.3	Density gradient, a/L_n	0.733
Ion-e ⁻ mass ratio, m_i/m_e	4590	Ion-e ⁻ temperature ratio, T_i/T_e	1.0
Effective ion charge, Z_{eff}	1.0	4 th order y hyperdiffusion [23], ϵ_y	0.0
4 th order χ hyperdiffusion [23], ϵ_χ	0.2	4 th order $v_{ }$ hyperdiffusion [23], $\epsilon_{v_{ }}$	0.2

Table 1. The nominal Cyclone Base Case (CBC) parameters [24] used in this work. Unless otherwise noted the geometry is specified using the s- α model [25] with $\alpha = 0$ and the simulations are electrostatic and collisionless.

(44) and (45). It reveals that not every parallel location permits Fourier modes with a radial wavenumber of zero. Prior works [11, 14] appear to have presupposed the existence of a local radial wavenumber of $K_x = 0$ at all parallel locations. This created a number of problems, which made a practical implementation appear impossible. Notably, if a mode exists at $K_x = 0$, this can violate the radial boundary condition, necessitate a prohibitively small K_x grid spacing, and/or require arbitrary non-uniform K_x grid spacing when using general geometry. All of these difficulties are resolved by constructing the grid according to equation (45), which prohibits the $K_x = 0$ mode at most parallel locations.

Second, the analytic derivation presented in this section actually accomplishes remarkably little. We have simply rewritten the Fourier-analyzed gyrokinetic model in terms of a new radial wavenumber K_x , which absorbs some of the information that was previously contained in the geometric coefficient $\vec{\nabla}y$. As noted above, if we constructed a K_x grid using equation (40) and $m \in [-(N_x - 1)/2, (N_x - 1)/2]$ at all parallel locations, nothing consequential would change about the representation. Our change of coordinate system would be purely semantics and running a simulation would produce the exact same results as the conventional coordinate system. This again highlights the importance of equations (44) and (45). Performing the change of coordinates to K_x is not what creates a non-twisting flux tube, rather the essential step is *constructing a near-rectangular grid* in the new (K_x, k_y) coordinates. Alternatively, one could have created a non-twisting flux tube without changing coordinates by constructing a non-rectangular grid in (k_x, k_y) . Specifically one would use a grid centered around $k_x \approx -k_y \vec{\nabla}x \cdot \vec{\nabla}y / |\vec{\nabla}x|^2$, which can be found by expressing the grid of equation (45) in terms of k_x by substituting equation (41). In fact, this is essentially what is done in the shifted metric approach. Nevertheless, we believe that using a coordinate system transformation, as presented in this paper, is more intuitive and demonstrates the constraints on the system (e.g. boundary conditions) well.

Coordinate	Grid range		Number of grid points	
	Adiabatic	Kinetic	Adiabatic	Kinetic
x/ρ_i	$[-100, 100)$	$[-92, 92)$	N_x	N_x
y/ρ_i	$[-63, 63)$	$[-63, 63)$	64	64
χ	$[-N_{\text{pol}}\pi, N_{\text{pol}}\pi)$	$[-N_{\text{pol}}\pi, N_{\text{pol}}\pi)$	$20N_{\text{pol}}$	$48N_{\text{pol}}$
$v_{\parallel}/v_{th,s}$	$[-3, 3]$	$[-3, 3]$	32	64
$\sqrt{\mu/(T_s/B_r)}$	$(0, 2.75)$	$(0, 2.75)$	10	10
$t/(a/c_S)$	$[1000N_{\text{pol}}, 6000]$	$[30, 200]$	time step < CFL limit [26]	

Table 2. The nominal GENE coordinate grids used in this work. Note that all grids are equally spaced, $v_{th,s} = \sqrt{2T_s/m_s}$ is the thermal velocity of species s , and $c_S = \sqrt{T_e/m_i}$ is the sound speed.

4. Code benchmarking

In this section, we will benchmark our implementation of the non-twisting flux tube in GENE. The practical details involved in implementing the results of section 3 are fairly technical and are therefore relegated to Appendix B. However, this section on benchmarking is important because it presents a concrete illustration of the relationship between the grids in the non-twisting and conventional flux tubes. Unless otherwise noted, we will use standard Cyclone Base Case (CBC) parameters [24] given in table 1 using the resolutions given in table 2. Our first test is to compare linear growth rates against the conventional flux tube when the two coordinate systems are identical. To see when this occurs, we must return to the discussion surrounding equation (44). Remember, if $m_0(k_y, \chi) = 0$ for all binormal modes and all parallel locations, our modified radial grid becomes functionally identical to that of the conventional flux-tube. In other words, the non-twisting and conventional flux tubes model the exact same physical situation using the exact same grid, just with different labels on the grid points.

To simulate a single linear mode, one sets $N_{\text{asp}} = 1$ and $k_y = 2\pi/L_y$. Thus, we see from equation (44) that $m_0(k_y, \chi)$ becomes

$$m_0\left(\frac{2\pi}{L_y}, \chi\right) = -\text{NINT} \left[\mp \frac{1}{2} \text{sign}(\hat{s}) \frac{\chi}{\pi N_{\text{pol}}} \mp \frac{C_y q_0}{2\pi N_{\text{pol}} |\hat{s}|} \frac{\vec{\nabla}x \cdot \vec{\nabla}\chi}{|\vec{\nabla}x|^2} \right], \quad (59)$$

where we have used equations (9) and (25). In general this expression is *not* necessarily zero because of the local correction to the magnetic shear, which is represented by the second term. For tokamaks with extreme shaping or a tight aspect ratio, the second term can be made arbitrarily large, meaning the expression will round to a non-zero value at some parallel locations. However, for typical tokamak parameters the local correction to the shear is often small. In fact, for circular flux surfaces in the large aspect ratio limit, this local correction vanishes entirely. This can be accomplished in

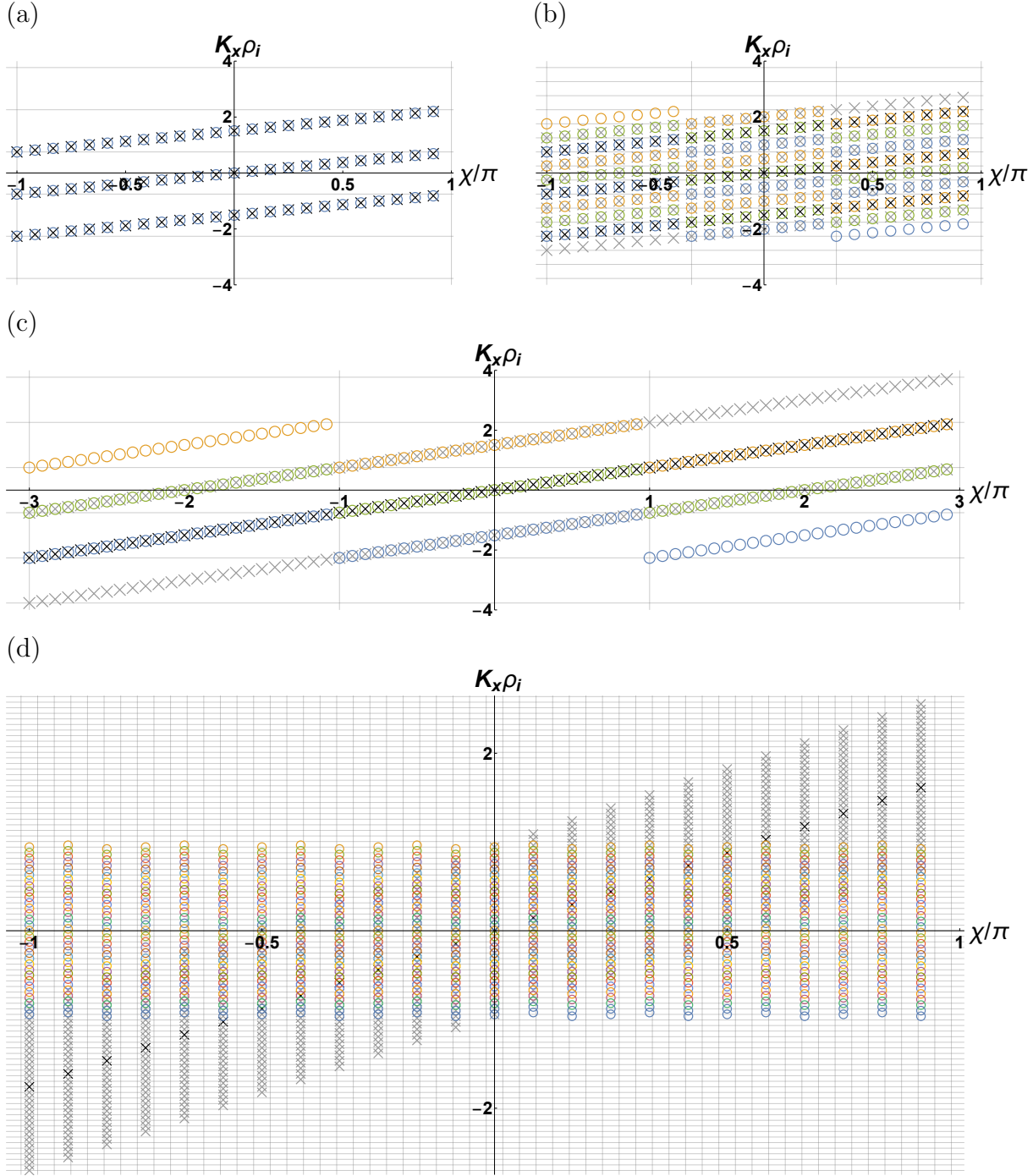


Figure 4. The parallel-radial spatial grid expressed in (χ, K_x) from the non-twisting (colored circles) and conventional (black and gray crosses) flux tubes for (a) $N_{\text{asp}} = 1$, $N_{\text{pol}} = 1$, $N_x = 3$; (b) $N_{\text{asp}} = 3$, $N_{\text{pol}} = 1$, $N_x = 9$; (c) $N_{\text{asp}} = 3$, $N_{\text{pol}} = 3$, $N_x = 3$; and (d) $N_{\text{asp}} = 4$, $N_{\text{pol}} = 1$, $N_x = 31$. The first three plots are grids for linear simulations and use a single binormal mode $k_y \rho_i = 0.3$, while the last plot is a nonlinear grid with $k_y \rho_i = 0.7$ and a minimum wavenumber of $k_{y,\text{min}} \rho_i = 0.05$. All use $N_\chi = 24N_{\text{pol}}$ parallel grid points and have only global magnetic shear. The vertical grid lines and changes in color indicate where linear modes couple to a new radial index in the non-twisting flux tube (i.e. the argument to the NINT function is half-integer), the horizontal grid lines correspond to the connections across the parallel boundary, and the black crosses indicate the linear mode with zero ballooning angle.

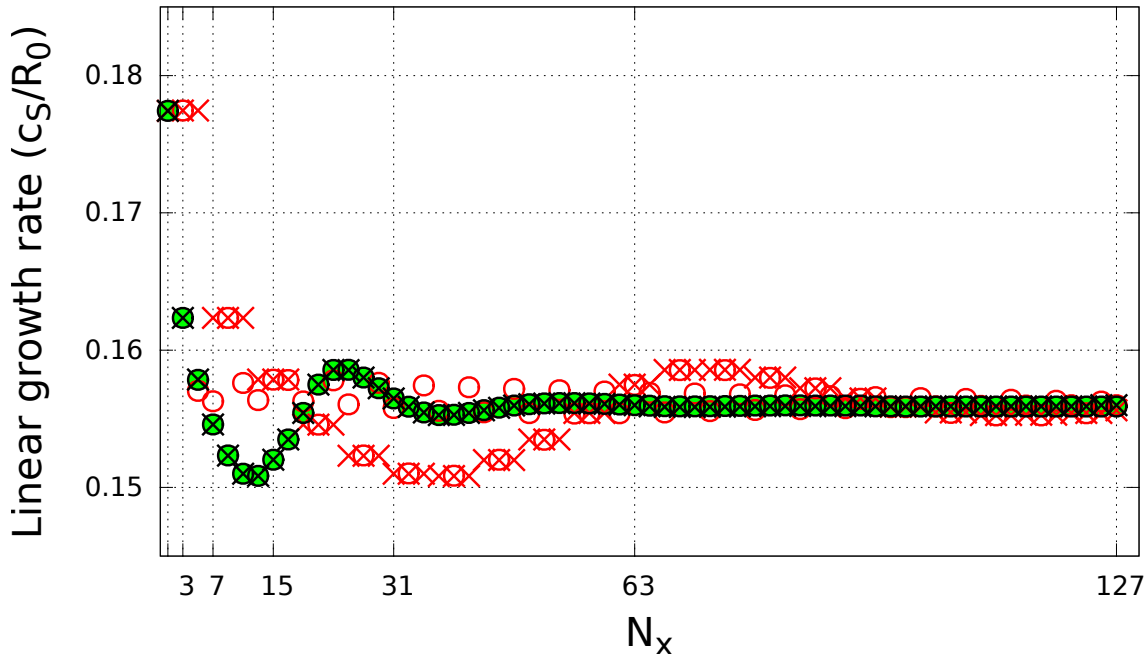


Figure 5. A resolution study of the linear growth rate of the most unstable mode using the non-twisting (empty or filled circles) or conventional (crosses) flux tube with $N_{\text{asp}} = 1$ and $N_{\text{pol}} = 1$ (black), $N_{\text{asp}} = 3$ and $N_{\text{pol}} = 1$ (red), or $N_{\text{asp}} = 3$ and $N_{\text{pol}} = 3$ (filled green circles). These correspond to the grids from figure 4(a)-(c).

GENE by specifying the magnetic geometry using the s - α representation [25] and setting $\alpha = 0$. Thus, with only global shear equation (59) becomes

$$m_0 \left(\frac{2\pi}{L_y}, \chi \right) = -\text{NINT} \left[\mp \frac{1}{2} \text{sign}(\hat{s}) \frac{\chi}{\pi N_{\text{pol}}} \right]. \quad (60)$$

Given the range of $\chi \in [-\pi N_{\text{pol}}, \pi N_{\text{pol}})$, this expression will necessarily round to zero at all parallel locations as long as we are careful to properly round at $\chi = -\pi N_{\text{pol}}$ (see Appendix B for more details). This ensures that the two grids are identical, as is shown in figure 4(a). Therefore, we have derived a simple linear test case for which the non-twisting and conventional flux tubes should produce identical results — using a single k_y mode with $N_{\text{asp}} = 1$ in the s - α geometry model with $\alpha = 0$. The linear growth rates for this test are shown in figure 5 by the black markers. We see perfect agreement as expected, regardless of the number of radial grid points N_x (which determines how far the linear mode is allowed to extend in ballooning space). This test gives confidence that the modifications to the geometric coefficients was done properly (e.g. replacing $|\vec{\nabla}y|^2$ with $|\vec{\nabla}Y|^2$).

In the next test case, we change N_{asp} from 1 to 3. While N_x determines how far the linear modes extend in ballooning space, the value of N_{asp} determines the number of independent linear modes in the system. This can be seen in figure 4(b). Note that *both* the non-twisting and conventional flux tubes now contain three linear modes, whereas

they only contained one linear mode in figure 4(a). We also see that when $N_{\text{asp}} \neq 1$ the non-twisting and conventional grids can differ. In particular, we see that linear modes in the non-twisting flux tube now can fall off the computational grid at internal parallel locations. This differs from the conventional flux tube, in which linear modes can only terminate at the parallel boundaries. However, for the CBC parameters used for this study, we know that the linear mode with zero ballooning angle (i.e. the one passing through the origin) will have the largest growth rate. Looking closely at the example shown in figure 4(b), which uses $N_x = 9$, we see that the linear mode with zero ballooning angle has the exact same grid points in the two flux tubes, even though the grids differ for the other two linear modes. This coincidence only occurs at specific values of N_x (i.e. N_x is an integer multiple of N_{asp}), which makes it a good test case. Studying the growth rates for this test case (shown in red in figure 5), we see that at $N_x = 9$ the two flux tubes produce the same growth rate, as is expected from figure 4(b). Moreover, the two flux tubes agree at every third value of N_x , while they do not for other values of N_x . This pattern exactly reflects when the grids of the zero ballooning angle linear mode are identical and when they differ. Additionally, we see that the growth rate at every third value of N_x agrees with the growth rate from the $N_{\text{asp}} = 1$ case (shown in black) at $N_x/3$. This is because they also have identical grids for the zero ballooning angle linear mode, which is the case between figures 4(a) and 4(b). The fact that the growth rates exactly inherit all these underlying patterns from the coordinate systems grids confirms that the mode coupling in the parallel derivative is implemented correctly.

Next, we set $N_{\text{pol}} = 3$ to extend the length of the flux tube to three poloidal turns. By also choosing $N_{\text{asp}} = 3$, we model three identical linear modes that are each the same as our first $N_{\text{pol}} = 1$ test case at all values of N_x . The only thing that has changed is that some of the mode coupling that was occurring at the ends of the flux tube has been moved into the interior. Accordingly, we see that the growth rates in figure 5 (green and black points) agree well at all values of N_x . This verifies that the non-twisting flux tube has no distinction between mode coupling at the parallel boundary condition and mode coupling within the domain. We note that all of the test cases, in both the non-twisting and conventional flux tubes, converge to the same value at sufficiently high radial resolution. This is important because all are modeling the same physical situation just with different numerical grids.

Lastly, we will perform a nonlinear benchmark. Unless the magnetic shear is zero at all parallel locations, the grids used for nonlinear simulations will always be different. Figure 4(d) shows that the non-twisting flux tube has a grid centered around $K_x = 0$, while the conventional grid is centered around $k_x = 0$. This is made possible by the fact that linear modes in the non-twisting flux tube can fall off of the computational grid at any parallel location and do not necessarily span an integer number of poloidal turns. On the other hand, figure 4(d) shows that linear modes in the conventional flux tube only ever terminate at the parallel boundary and are, thus, always an integer number of poloidal turns long. Regardless, as with the linear studies, both flux tubes

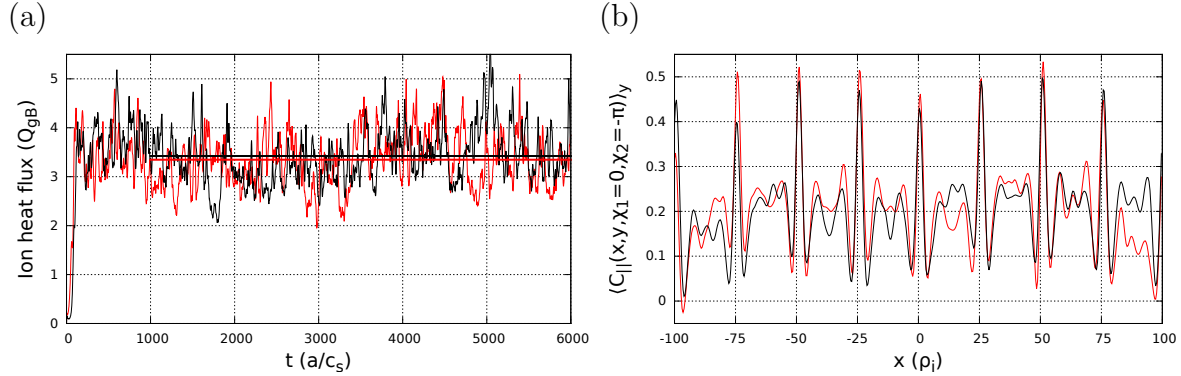


Figure 6. A nonlinear benchmark at the high radial resolution of $N_x = 512$ with $N_{\text{pol}} = 1$: (a) the time traces of the turbulent ion heat flux (thin lines) together with their time average over the nonlinearly saturated state (thick lines) and (b) the radial variation of the parallel correlation function for the non-twisting (black) and conventional (red) flux tubes. In (b) the locations of the pseudo-integer magnetic surfaces are indicated by the vertical grid lines.

should produce identical results at sufficiently high radial resolution. Accordingly, we performed a nonlinear simulation with the very high radial resolution of $N_x = 512$ using the CBC parameters given in table 1, the resolutions given in table 2, $N_{\text{pol}} = 1$, and adiabatic electrons. Figure 6(a) shows that the time-averaged heat flux agrees very well between the two flux tubes and the time traces look qualitatively similar. Note the heat flux is normalized to the gyroBohm value of $Q_{gB} \equiv (\rho_i/a)^2 n_e T_e c_S$.

As an additional check, we calculate the two-point parallel correlation function

$$C_{||}(x, y, \chi_1, \chi_2) \equiv \frac{\langle \bar{\phi}_{NZ}(x, y, \chi_1, t) \bar{\phi}_{NZ}(x, y, \chi_2, t) \rangle_t}{\sqrt{\langle \bar{\phi}_{NZ}^2(x, y, \chi_1, t) \rangle_t \langle \bar{\phi}_{NZ}^2(x, y, \chi_2, t) \rangle_t}}, \quad (61)$$

where the subscript NZ signifies the non-zonal portion of the quantity and $\langle \dots \rangle_u$ indicates an average over any coordinate u . The quantity $C_{||}$ indicates the degree of correlation between two points χ_1 and χ_2 on *the same field line*, which is accomplished by holding x and y constant (as opposed to x and Y). To transform to the field-aligned coordinate y from the grid used by the non-twisting flux tube, one must use the inverse Fourier transform

$$\bar{\phi}(x, y, \chi) = \sum_{K_x, k_y} \Phi(K_x, k_y, \chi) e^{-ik_y(\vec{\nabla}_x \cdot \vec{\nabla}_y / |\vec{\nabla}_x|^2)x} e^{iK_x x + ik_y y}, \quad (62)$$

which can be derived by substituting equations (24) and (29) into equation (39). Figure 6(b) shows the y -averaged correlation between the inboard and outboard midplanes $\langle C_{||}(x, y, \chi_1 = 0, \chi_2 = -\pi) \rangle_y$ for the two flux tubes. We see that, despite using different grids, the two parallel correlation functions are very similar. As should be the case, we see that both flux tubes display spikes in the parallel correlation function at “pseudo-integer” surfaces [12]. On these surfaces, due to the parallel boundary condition, the

magnetic field lines close on themselves after just one poloidal turn, which is artificial unless the flux tube covers the full flux surface. Thus, if turbulent eddies span a full poloidal turn, they will “bite their own tails.” This self-interaction alters the statistical properties of the turbulence, resulting in the radially localized spikes shown in figure 6(b). Interestingly, the extreme radial resolution reveals that each spike is actually shifted slightly outwards (i.e. to the right in figure 6(b)) from its pseudo-integer surface by a few ion gyroradii. This appears to be a consequence of the magnetic drifts, which prevent the single particle trajectories from precisely following magnetic field lines. Thus, the turbulence may not be *exactly* field-aligned, but can have a slight bias one way or the other. A field line that just misses closing on itself can actually compensate for the drifts and enable a larger degree of correlation. While this slight shift is a minor effect, its presence in both flux tube models gives even more confidence that our implementation of the non-twisting flux tube is correct.

5. Nonlinear performance results

In this section, we will investigate the performance of the non-twisting flux tube relative to the conventional flux tube using five test cases:

- (1) CBC with adiabatic electrons,
- (2) CBC with a long computational domain,
- (3) CBC-like with high magnetic shear,
- (4) a shaped DEMO equilibrium with kinetic electrons, and
- (5) CBC-like with high magnetic shear and a long computational domain.

We will compare the runtime of the code (i.e. the total computational time needed to reach a given physical time) to see how it is affected by the modifications for the non-twisting flux tube. Additionally, we will study the convergence with radial resolution while holding the radial domain size L_x constant. In other words, we will hold the minimum value of the radial wavenumber constant and increase the maximum value by adding more Fourier modes. This will reveal if centering a grid around $K_x \approx 0$ enables convergence with fewer radial modes.

Test case 1

For the first test case, we performed standard CBC simulations with the parameters and resolutions given in tables 1 and 2, the s- α geometry, and adiabatic electrons. These simulations have a modest value of $\hat{s} = 0.8$ and use domains that are just one poloidal turn long. The time averaged turbulent heat fluxes from both flux tubes are shown in figure 7, as is the speed-up (i.e. the total wall clock simulation time of the conventional flux tube divided by that of the non-twisting flux tube). We see that both the heat flux and the computational cost of the two flux tubes are similar at all radial resolutions. Surprisingly, the biggest difference is that the non-twisting flux tube is slightly *faster* at

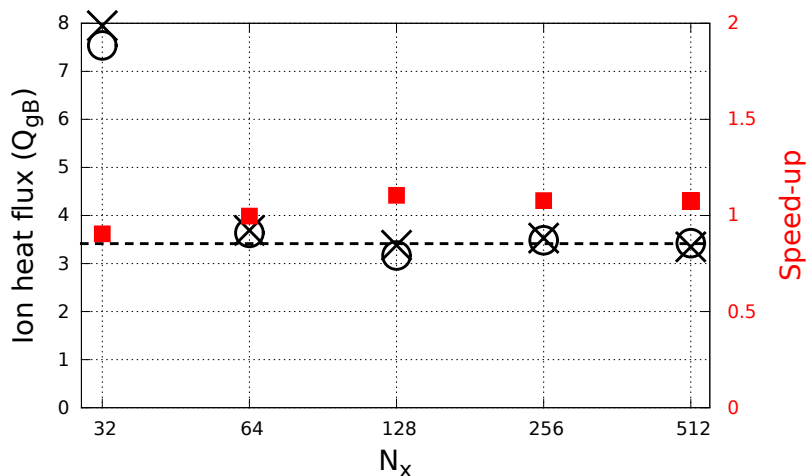


Figure 7. A nonlinear resolution study for the non-twisting (black circles) and conventional (black crosses) flux tube using adiabatic electrons, $N_{\text{pol}} = 1$, CBC parameters with $\hat{s} = 0.8$, and the parameters and resolutions of tables 1 and 2. The ion heat flux (black) and computational speed-up (red squares) data use the left and right axes respectively and the horizontal dashed line indicates the fully converged heat flux. Each simulation was run on $10N_x$ CPUs (except 2560 CPUs were used for $N_x = 512$).

high resolution. Looking into the details, while the CPU time required to compute one time step in the non-twisting flux tube is longer by a factor of ~ 1.2 , the physical time step itself is longer by a factor of ~ 1.3 . We believe that the slowdown in calculating a time step is because the parallel derivative at fixed k_x and k_y is no longer taken across continuous sections of memory (see Appendix B). However, longer time steps are possible because the local radial wavenumbers of the non-twisting flux tube grid stay closer to $K_x = 0$ than in the conventional flux tube, especially at the inboard midplane. This means that, at the same value of N_x , switching to the non-twisting flux tube simulation will eliminate the finest spatial scales, which allows for a longer time step according to the CFL condition [26]. Since GENE has a routine for calculating the maximum allowable timestep, it automatically takes advantage of this. In these simulations (as well as all subsequent simulations), no difference was observed in the temporal convergence properties of the turbulent state in the two flux tubes (i.e. all simulations converged to quasi-steady state at roughly the same simulation time).

Test case 2

The second test case is identical to the first, except the domain is lengthened from one to three poloidal turns by setting $N_{\text{pol}} = 3$. Though simulation domains with more than one poloidal turn have fallen out of fashion, they were originally recommended to ensure properly converged results [9]. Without testing for convergence in N_{pol} (or alternatively L_y), parallel self-interaction may be artificially large and affect the accuracy of the results [12, 27]. Simulations with $N_{\text{pol}} > 1$ are also important for stellarators as

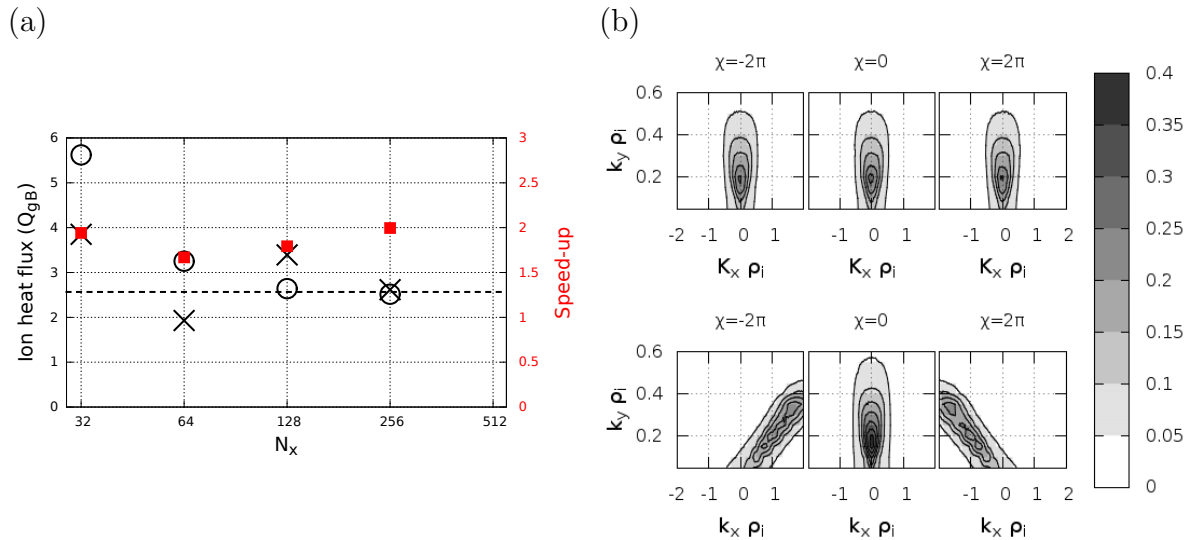


Figure 8. (a) A resolution study for the non-twisting (black circles) and conventional (black crosses) flux tube using adiabatic electrons, $N_{\text{pol}} = 3$, CBC parameters with $\hat{s} = 0.8$, and the parameters and resolutions of tables 1 and 2. (b) The two-dimensional wavenumber spectra of the electrostatic potential (i.e. $\langle |\Phi| \rangle_t$ or $\langle |\phi| \rangle_t$) at the three outboard midplanes from the $N_x = 128$ simulations. Note the full range in radial wavenumber used by the simulations is shown, while the binormal range is restricted for visual clarity. Each simulation was run on $20N_x$ CPUs.

each poloidal turn is physically different due to the three-dimensional geometry [22, 28]. While this test uses a tokamak geometry, it is still expected to be a good indicator for the performance in a stellarator, so long as the magnetic shear is comparable. The important distinguishing feature of this test is that there is now more than one region of turbulent drive. Specifically, turbulence is driven at all three outboard midplanes in the simulation and should be statistically identical because the physical conditions are identical due to axisymmetry.

Figure 8(a) shows the results of the resolution study. First, we note that, for the same value of N_x , the non-twisting calculation is almost twice as fast. Breaking this down, we find that, while the CPU time required to compute one time step in the non-twisting flux tube is longer by a factor of ~ 1.7 , the physical time step itself is longer by a factor of ~ 3.3 . As in the first test, these two effects counteract one another, but the net effect is a computational speed-up for the non-twisting flux tube. Of course, the magnitude of the speed-up is significantly larger for this test case, which is intuitive as the twist from one end of the computational domain to the other is more extreme.

Looking at the convergence in N_x , we see fairly different behavior. The non-twisting flux tube data converges monotonically to the high resolution limit, while the conventional flux tube data jumps around. To make sense of this, figure 8(b) shows the turbulent spectrum of the electrostatic potential at all three outboard midplanes for both $N_x = 128$ simulations. For the non-twisting flux tube, we see the expected results — since all three outboard midplanes are physically identical, they each have statistically

identical turbulence. However, for the conventional flux tube, the turbulence at the central outboard midplane is different from that at the other two outboard midplanes. It has a maximum value of the electrostatic potential that is almost double the maximum value at the other two midplanes. This is because, despite the fact that all three locations are physically identical, they are modeled using different perpendicular grids. The conventional flux tube has a rectangular cross-section at $\chi = 0$, but is twisted into parallelograms at $\chi = -2\pi$ and 2π . This means that the turbulent activity is tilted diagonally across the perpendicular wavenumber grid and actually extends off the grid at $\chi = -2\pi$ and 2π , even at the high radial resolution of $N_x = 128$. As a consequence, the conventional flux tube requires a radial resolution that is more than double that of the non-twisting flux tube. Comparing the runtime of the non-twisting simulation at $N_x = 128$ with that of the conventional flux tube simulation at $N_x = 256$, we find that the overall computational cost of a properly resolved $N_{\text{pol}} = 3$ CBC simulation is reduced by a factor of 7 by employing the non-twisting flux tube.

Test case 3

The third test case uses $N_{\text{pol}} = 1$ and is identical to the first, except the strength of the global magnetic shear is quintupled to $\hat{s} = 4.0$. Due to the larger magnetic shear, the background temperature gradient had to be increased to $a/L_{Ti} = 6.5$ to ensure instability. To more efficiently model these new conditions, the radial and binormal domain widths were approximately halved to $L_x = 80\rho_i$ and $L_y = 63\rho_i$ along with a corresponding decrease in the number of binormal grid points to $N_y = 32$. Such a reduced grid still resolves the turbulence well, which is not surprising as stronger magnetic shear tends to reduce the correlation lengths. Additionally, the number of parallel grid points was increased to $N_\chi = 40N_{\text{pol}}$ to better resolve shorter parallel correlation lengths, the parallel hyperdiffusion was lowered to $\epsilon_\chi = 0.02$ to match the linear growth rate, and the time average over the nonlinearly saturated state was performed from $t = 500$ to $2000a/c_s$. The resulting resolution study is shown in figure 9(a). We see that the non-twisting flux tube benefits from a speed-up of around 2 (although it is fairly inconsistent).

Figure 9(a) also shows that the convergence with N_x is fairly similar between the two flux tubes. This may be surprising as it seems like the non-twisting flux tube should be beneficial when the magnetic shear is high because the conventional flux tube cross-section becomes strongly twisted at the inboard midplane. However, figure 9(c) shows that the turbulence is very localized around the outboard midplane. Thus, the *flux surface averaged* fluxes are insensitive to resolution problems around the *inboard* midplane because the averages are dominated by behavior at the outboard midplane. Looking carefully at figure 9(c) we see that, as N_x is lowered, an unphysical grid-scale oscillation develops in the parallel direction away from the outboard midplane in the conventional flux tube. Similar behavior was also observed in reference [11]. Such oscillations do not occur in the non-twisting flux tube, presumably because it includes

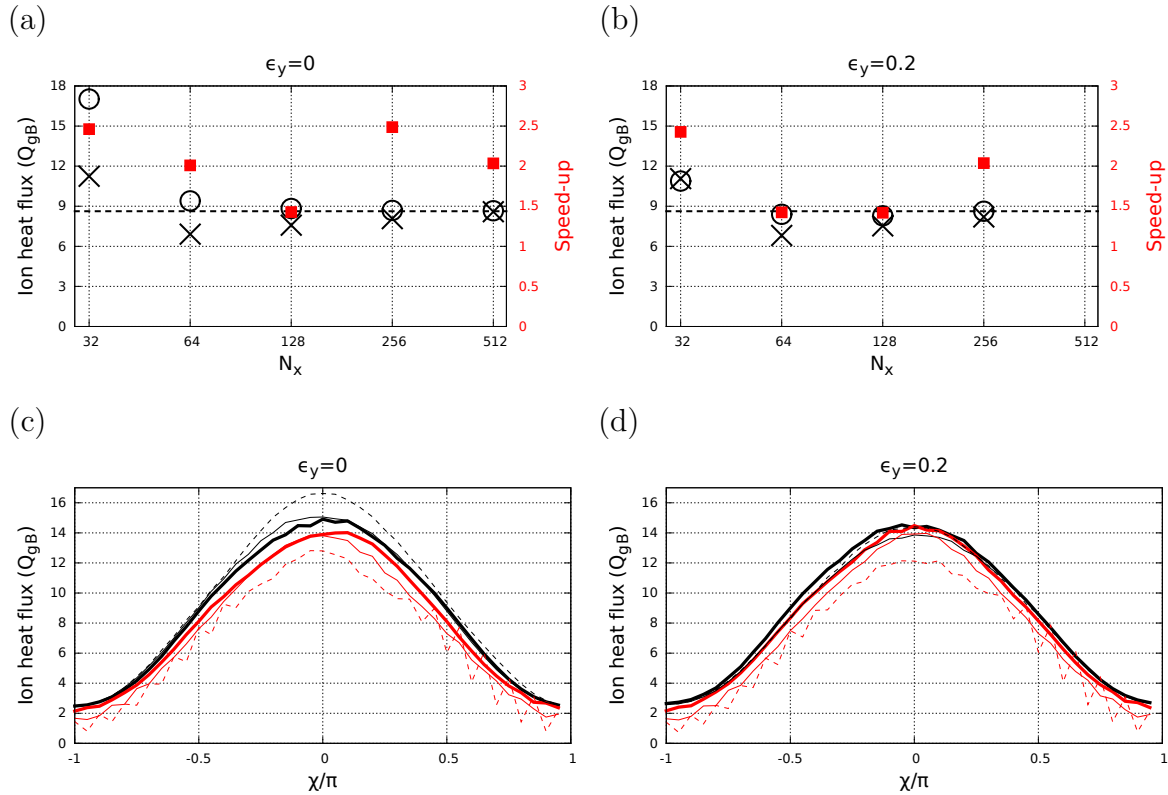


Figure 9. The top row shows a resolution study for the non-twisting (black circles) and conventional (black crosses) flux tube (a) without hyperdiffusion in y or (b) with $\epsilon_y = 0.2$. It uses adiabatic electrons, $N_{\text{pol}} = 1$, CBC parameters with $\hat{s} = 4.0$, and the parameters and resolutions of tables 1 and 2 (with the modifications detailed in the text). The bottom row shows the corresponding poloidal distribution of the heat flux (c) without or (d) with hyperdiffusion in y for the non-twisting (black) and conventional (red) flux tube with radial resolutions of $N_x = 256$ (thick solid), $N_x = 128$ (thin solid), and $N_x = 64$ (thin dashed). Each simulation was run on $1.25N_x$ CPUs.

modes near $K_x = 0$ at all poloidal locations and the turbulent activity does not fall off the grid. To summarize, while the non-twisting flux tube better resolves the turbulence away from $\chi = 0$, this does not appear to have a big impact on the overall fluxes. Thus, for this test case, the non-twisting flux tube is most advantageous if you are interested in local quantities away from $\chi = 0$. Alternatively, the non-twisting flux tube could provide a big advantage if the turbulent drive is not centered around $\chi = 0$ (as is indicated by the second test case). This is true in sheared slab geometry as all parallel locations are statistically identical and have equal turbulent drive [11]. It may also be the case in stellarator (see figure 7.12 of reference [29]) or tokamak pedestal simulations [30, 31, 32, 33], which often observe that the linear growth rate peaks for a non-zero k_x corresponding to the top and bottom of the flux surfaces. In nonlinear simulations, such turbulence would be very expensive to resolve with the conventional flux tube when the magnetic shear is large.

On the other hand, the non-twisting flux tube does have to make a sacrifice in order to better resolve the inboard midplane. Because the radial wavenumber grid always stays

centered around $K_x \approx 0$, the length of the linear modes is reduced when the binormal mode index $n = k_y/k_{y,\min}$ is large. This can be understood using figure 4(d), which has $n = 14$. We see that, in the conventional flux tube, all of the linear modes are at least one poloidal turn long. At higher values of n , this can only be ensured by sacrificing *ballooning angles* (i.e. the poloidal angle $\theta_0 \equiv k_x/(k_y\hat{s})$ at which the linear mode would have $K_x = 0$ if the local correction to the shear was zero). In the limit of very high n , all of the linear modes in the k_x grid have ballooning angles that are strongly clustered around the $\chi = 0$, but they remain one poloidal turn long. Hence, the linear modes that are present at high n in the conventional flux are well resolved in the parallel direction, but this comes at the cost of entirely omitting linear modes with ballooning angles away from $\chi = 0$. This can lead to problems like the oscillations discussed in the previous paragraph. In contrast, figure 4(d) shows that the non-twisting flux tube maintains the same number of linear modes at all values of n , but they can become very short in the parallel direction. In fact, if the k_y grid is large enough, the highest values of n will have linear modes that are each comprised of just a single parallel grid point. This can be challenging to implement properly (as discussed in Appendix B) and seems like a major concern for the non-twisting flux tube. Fortunately, this behavior appears physically appropriate because the parallel extent of linear modes is indeed reduced at higher values of k_y . This can be seen clearly in figure 3 of reference [34] and is a consequence of the fact that the magnetic shear acts in proportion to $k_y\hat{s}$ (e.g. equations (22) and (23)). To summarize, just as the conventional flux tube has difficulty resolving the inboard midplane, the non-twisting flux tube *may* have issues at large n , i.e. when using grids with many k_y modes. The non-twisting simulations shown in figure 9(a) did exhibit pile-up in the k_y spectrum. However, simply adding a small amount of hyperdiffusion in the binormal wavenumber $\epsilon_y = 0.2$ resolved this problem and did not affect the overall heat fluxes, as is shown in figure 9(b,d).

Test case 4

The fourth test case uses a realistic “D”-shaped numerical equilibrium based on an inductive EU-DEMO scenario [35] with a 50-50 mixture of deuterium-tritium fuel. It has a high magnetic shear of $\hat{s} = 2.4$ and strongly driven Ion Temperature Gradient (ITG) turbulence. Unlike the prior test cases, these simulations include kinetic electrons and local corrections to the magnetic shear. The equilibrium has $x_0/R_0 = 0.43$, $q_0 = 1.6$, $R_0/L_{Ti} = 11.7$, $R_0/L_{Te} = 12.0$, and $R_0/L_n = 0.85$. The parallel hyperdiffusion was increased to $\epsilon_\chi = 1.0$ to match the stronger drive, a binormal hyperdiffusion of $\epsilon_y = 0.2$ was used, and the gyroBohm heat flux is defined as $Q_{gB} \equiv (\rho_i/R_0)^2 n_e T_e c_S$ for this test case. In figure 10(a), we see that the computational cost of the two flux tubes is similar at all radial resolutions. Looking into the details, we actually find that the time step does not vary with N_x or the type of flux tube, which is very different from the previous test cases with *adiabatic* electrons. We believe that this is because including kinetic electrons changes the physics determining the time step. Specifically, as the electron

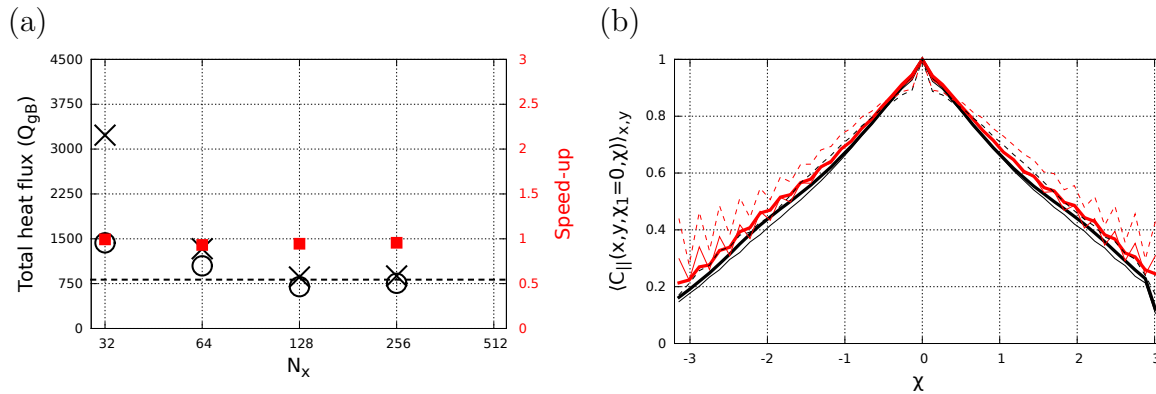


Figure 10. (a) The total heat flux and speed-up for the non-twisting (black circles) and conventional (black crosses) flux tube using kinetic electrons, a realistic shaped DEMO equilibrium with $\hat{s} = 2.4$, $N_{\text{pol}} = 1$, and the resolutions of table 2. (b) The corresponding parallel correlation function for the non-twisting (black) and conventional (red) flux tube with radial resolutions of $N_x = 256$ (thick solid), $N_x = 128$ (thin solid), and $N_x = 64$ (thin dashed). All simulations were run on 5120 CPUs.

plasma $\beta_e \rightarrow 0$, the electrostatic shear Alfvén wave becomes very high frequency, thereby limiting the time step (see reference [36] and section 2.1 of reference [37]). This can be mitigated by using a small, but finite value of the plasma β . However, for these simulations the value was limited to $\beta = 10^{-4}$ because higher values were found to modify the heat flux. Unfortunately, at $\beta = 10^{-4}$ it appears that the electrostatic shear Alfvén wave still determines the time step. Nevertheless, we would like to note that this issue *may* be a particular consequence of the strong turbulent drive. Perhaps for more typical simulations, β could be increased further without affecting the fluxes and the smallest radial spatial scale would again determine the time step as in previous adiabatic electron test cases.

Looking at the convergence properties of figure 10(a), we find that the non-twisting flux tube converges at a slightly lower value of N_x , but there is not a big difference. However, figure 10(b) shows that there is a more significant difference in the parallel correlation between the outboard and inboard midplanes (i.e. equation (61)). Therefore, although figures 9(c,d) and 10(b) show different quantities, they reinforce the same conclusion — the non-twisting flux tube better resolves the inboard midplane at high values of the magnetic shear.

Test case 5

The fifth test case is designed to most clearly demonstrate the advantage of the non-twisting flux tube. It is a repeat of the third test case, except the simulation domain is lengthened to three poloidal turns, so that it has $N_{\text{pol}} = 3$, $\hat{s} = 4.0$, and adiabatic electrons. As before, due to the larger magnetic shear, the background temperature gradient had to be increased to $a/L_{Ti} = 6.5$ to ensure instability. To optimally resolve these new conditions, the number of parallel grid points was increased to $N_x = 40N_{\text{pol}}$,

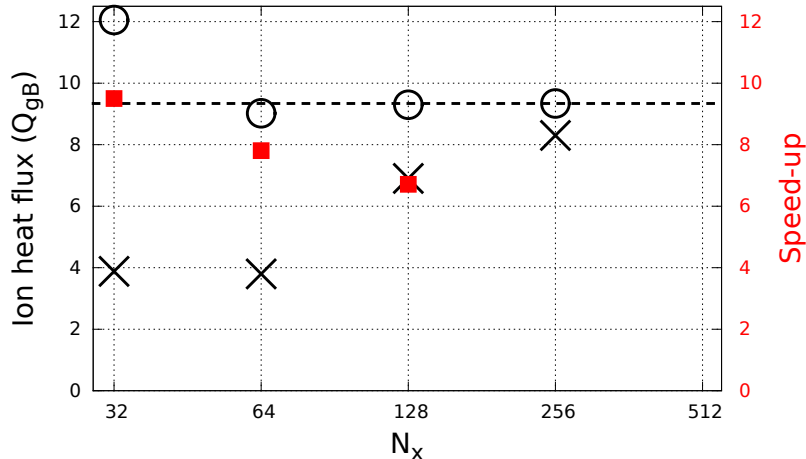


Figure 11. A resolution study for the non-twisting (black circles) and conventional (black crosses) flux tube using adiabatic electrons, $N_{\text{pol}} = 3$, CBC parameters with $\hat{s} = 4.0$, and the parameters and resolutions of tables 1 and 2 (with the modifications detailed in the text). Each simulation was run on $20N_x$ CPUs, but the conventional flux-tube run with $N_x = 256$ was not fully completed due to its computational cost.

the radial domain width was decreased to $L_x = 80\rho_i$, the binormal domain width was halved to $L_y = 63\rho_i$ along with a corresponding decrease in the number of binormal grid points to $N_y = 32$, and the parallel hyperdiffusion was lowered to $\epsilon_\chi = 0.02$ to match the linear growth rate. Additionally, the time average over the nonlinearly saturated state was performed from $t = 1000$ to $6000a/c_s$ and the hyperdiffusion in y was set to $\epsilon_y = 0.2$.

Figure 11 shows that non-twisting flux tube performs dramatically better than the conventional flux tube. At the same value of N_x , the non-twisting flux tube is more than six times faster, due to having a much longer time step. Moreover, we see that the non-twisting flux tube is well converged at $N_x = 64$, while the conventional flux tube requires many more radial grid points. In fact, we do *not* believe that the conventional flux tube is properly converged even at $N_x = 256$ because the radial spectrum of ϕ still exhibits significant pile-up. Instead we believe the heat flux is oscillating around the converged result in a similar fashion to figure 8(a). All of this means that a fully converged simulation is made at least 30 times less expensive by using the non-twisting flux tube, likely more.

While we deliberately chose the parameters of this test case to increase the computational savings, similar results likely hold in physically motivated simulations. The essential features are high magnetic shear and more than one region of turbulent drive. These could be found in pedestal simulations, which have high magnetic shear due to the proximity to the separatrix and may have two regions of turbulence drive (i.e. the top and bottom) [30, 31, 32, 33]. Alternatively, stellarator simulations generally require more than one poloidal turn [22, 28], though they often have small *global* magnetic shear. Nevertheless, the non-twisting flux tube could still benefit stellarators with high

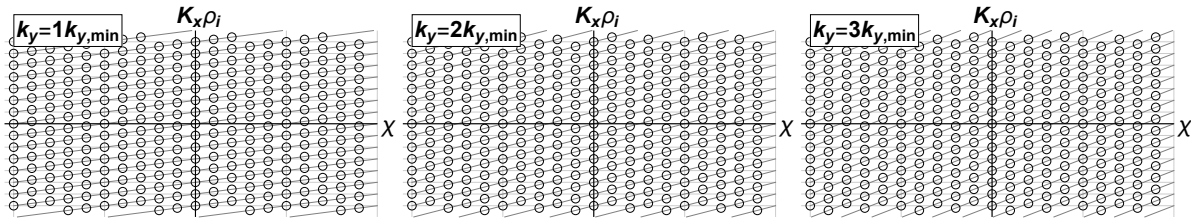


Figure 12. Cuts of the full three-dimensional lattice of allowed spatial grid points (black circles) at the three lowest allowed values of k_y (columns). The linear modes (diagonal gray lines) are indicated, $N_{\text{exc}} = 4$, $k_{y,\text{min}} = 2\pi/L_y$ is the minimum value of k_y anywhere on the spatial grid, and a regularly spaced χ grid is used.

local shear. Specifically, if local magnetic shear causes the conventional flux tube cross-section to twist strongly between two regions of turbulent drive, it doesn't matter if it eventually twists back to rectangular. The conventional flux tube would still be strongly sheared for one of the regions of turbulent drive and struggle to resolve it efficiently.

6. Alternative flux tubes

In this section we hope to stimulate future progress by discussing what spatial grids are possible in a flux tube and what considerations appear to limit our options. The coordinate system transform of section 3 was deliberately chosen to exactly and entirely eliminate the twist of the flux tube. While this seems like a natural choice to explore, many other possibilities exist, some of which may give better performance. In fact, it may be feasible to optimize the grid for each individual simulation to most efficiently treat the corresponding turbulent conditions. Therefore, we will imagine transforming the gyrokinetic model to a completely general set of coordinates $(\tilde{K}_x, \tilde{K}_y, \tilde{\chi})$ and seek constraints on their possible definitions.

The parallel coordinate is treated in real space, due to the variation of the geometric coefficients in this dimension. The $\tilde{\chi}$ grid is already known to be very flexible and the locations of the grid points can be chosen arbitrarily. By default GENE uses a regularly spaced grid in the straight-field line poloidal angle, but it also has the ability to concentrate the grid points around the outboard midplane (see Section 4.3 of reference [38]). GS2, on the other hand, has a routine that adapts the locations of the parallel grid points based on the geometric coefficients of the equilibrium being simulated [39]. For simplicity, in this section we will use a regularly spaced grid in the straight-field line poloidal angle $\tilde{\chi} = \chi$.

The other two spatial coordinates, binormal and radial, are represented using Fourier modes. This is well-motivated, given the periodic boundary conditions in these directions and the fact that linear modes are composed of Fourier modes (i.e. a linear mode has a fixed k_y and is a chain of linked K_x values in χ). By enforcing the boundary conditions, we discretize the allowed K_x and k_y Fourier modes into an infinite two-dimensional lattice at each value of χ . Given L_x , L_y , and $\vec{\nabla}x \cdot \vec{\nabla}y/|\vec{\nabla}x|^2$, we can

calculate all of these potential grid points, which are shown in figure 12 for an example simulation. Since L_x and L_y are numerical parameters, we can always increase them to make the grid spacing arbitrarily fine. Moreover, in theory, we have complete freedom to choose at will which of these grid points to include in the simulation. In practice, we will see that there are significant constraints that stem primarily from the physical effects that couple different wavenumbers — parallel streaming and the quadratic nonlinearity.

To understand these constraints, we will start by considering the binormal wavenumber. The most general possible definition of the binormal wavenumber is

$$\tilde{K}_y \equiv \tilde{f}_1(k_y, \chi), \quad (63)$$

as any dependence on \tilde{K}_x can be handled in its definition. Here \tilde{f}_j is any arbitrary function of its arguments and the subscript is simply to distinguish the many arbitrary functions that will appear in this section. How does equation (63) determine which grid locations a simulation includes from figure 12? As with the definition of K_x in the non-twisting flux tube, we lay down a grid in \tilde{K}_y that has a uniform spacing of $2\pi/L_y$ and is centered around $\tilde{K}_y \approx 0$. Each value of \tilde{K}_y that this produces must be rounded to the nearest grid point allowed by the boundary conditions. For the non-twisting flux tube, this was accomplished by including $m_0(k_y, \chi)$ in constructing the K_x grid of equation (45). As an example, we could define $\tilde{K}_y \equiv \sqrt{k_y 2\pi/L_y}$ to produce the grid shown in figure 13(a), which only includes the $n = \{0, 1, 4, 9, 16, \dots\}$ grid points from the traditional k_y grid defined by equation (14).

This makes clear that the form of equation (63) does *not* necessarily maintain an evenly spaced k_y grid, which makes it difficult to implement efficiently and realistically. Specifically, calculating the nonlinear term in real space according to equation (51) will become much more expensive because the Fourier transforms rely on computational libraries that assume a uniform grid. Instead, it would likely need to be directly calculated using a convolution in Fourier-space as given by equation (50), which makes it an $O(N_x^2 N_y^2)$ operation rather than $O(N_x N_y \log(N_x N_y))$. Additionally, if the \tilde{K}_y grid is non-uniform, then Fourier modes can couple via the quadratic nonlinearity to modes that do not exist *within* the grid. For example, this occurs for all nonlinear coupling involving different k_y modes in the grid of figure 13(a) because the grid spacing is quadratic. In the conventional flux tube, Fourier modes do nonlinearly couple to non-existing modes, but this only occurs beyond the boundaries of the grid (i.e. above the maximum value of k_y). Nonlinear coupling involving non-existing Fourier modes internal to the grid can be much more problematic. When a non-existing mode is surrounded by existing modes should it still be taken as having zero amplitude as is done beyond the grid boundaries? Should one somehow interpolate from the existing grid points (even though turbulent mode amplitudes are not continuous in Fourier-space) or use one grid point to represent multiple modes? The ultimate feasibility of non-uniform binormal grids is left for future work. It may be of interest for multi-scale simulations because one could perhaps retain some coupling between ion and electrons scales, but still coarsen the binormal grid at intermediate wavenumbers.

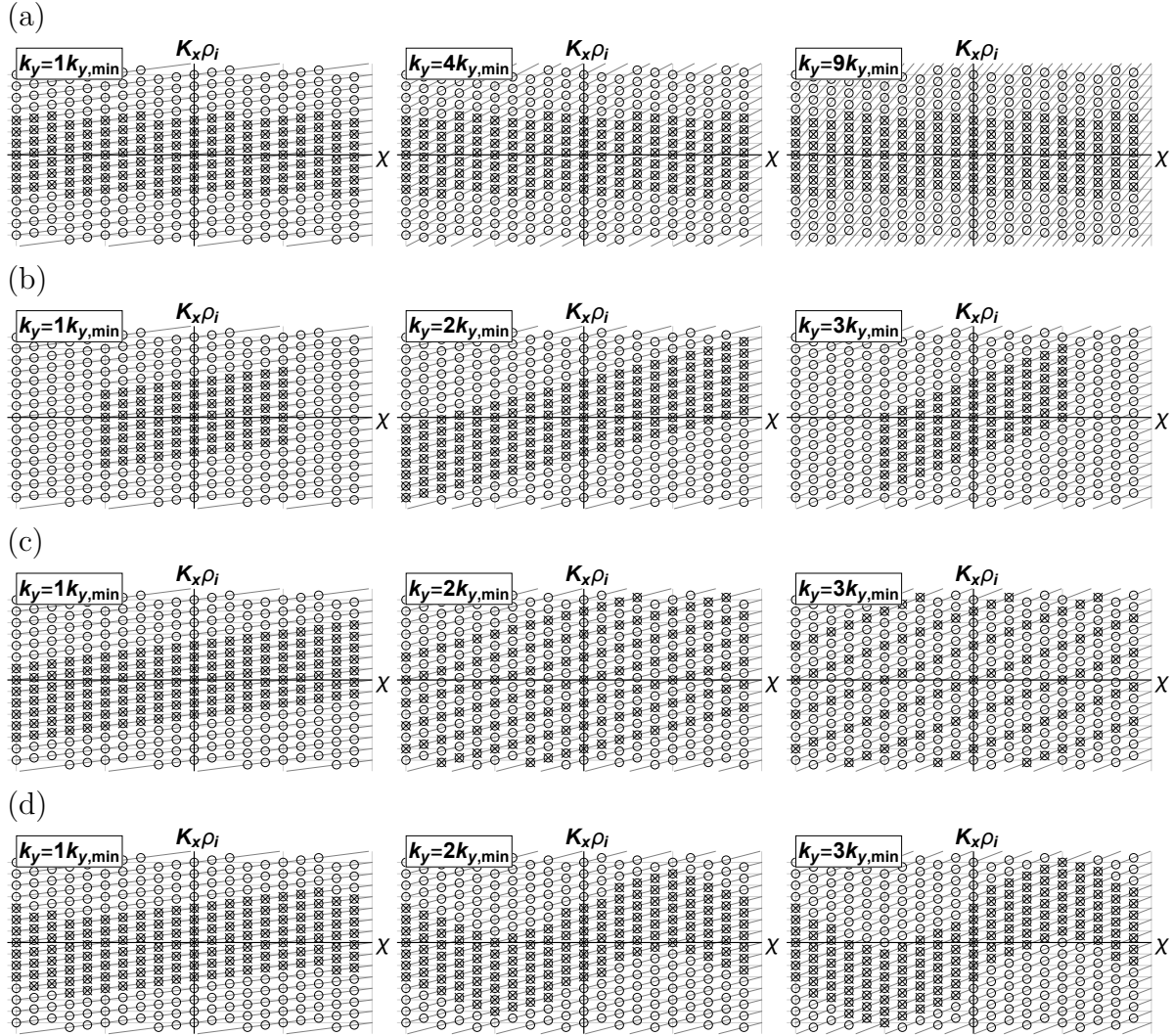


Figure 13. Cuts of the full three-dimensional lattice of the allowed (black circles) and included (black crosses) spatial grid points at the three lowest values of k_y included in the simulation (columns) using (a) the local radial wavenumber $\tilde{K}_x \equiv K_x$ with $\tilde{K}_y \equiv \sqrt{k_y 2\pi/L_y}$, (b) the ballooning radial wavenumber $\tilde{K}_x \equiv k_x$ with $\tilde{K}_y \equiv k_y/(1 + \text{NINT}[|\chi|/\pi])$, (c) $\tilde{K}_x \equiv k_x k_{y,\min}/k_y$ with $\tilde{K}_y \equiv k_y$, and (d) $\tilde{K}_x \equiv K_x + 2k_y s \sin(\chi)$ with $\tilde{K}_y \equiv k_y$. The linear modes (diagonal gray lines) are indicated, $k_{y,\min} = 2\pi/L_y$ is the minimum value of k_y anywhere on the spatial grid, $N_{\text{pol}} = 1$, and $N_x = 7$.

If we want to maintain an evenly spaced binormal grid, \tilde{K}_y must have a linear relationship with k_y according to

$$\tilde{K}_y \equiv \tilde{f}_2(\chi) k_y + \tilde{f}_3(\chi). \quad (64)$$

We can immediately set $\tilde{f}_3(\chi) = 0$ because the grid uses the reality condition to handle the negative binormal modes and we want the grid to retain the zonal modes at $k_y = 0$. Thus, the only difference from the conventional definition is that $\tilde{f}_2(\chi)$ can be used to vary the binormal domain width with parallel location. For example, figure 13(b) sets $\tilde{f}_2(\chi) = (1 + \text{NINT}[|\chi|/\pi])^{-1}$ in order to halve the binormal domain width on

the inboard side. Such a grid appears feasible to implement and would change the \tilde{K}_y grid spacing between neighboring parallel grid locations. Of course, figure 13(b) shows that doing so can cause many linear modes to end (i.e. couple to zero) in the interior of the (\tilde{K}_y, χ) grid. This is analogous to the issue with nonlinear coupling in non-uniform k_y grids, but is perhaps less conceptually concerning as the Fourier coefficients are continuous with χ . Therefore, one could conceivably interpolate individual missing grid points or simply ensure that there is negligible turbulent activity at the parallel locations where a grid transition takes place. Perhaps a domain size that changes in the parallel direction could be appropriate for strongly shaped stellarators that have several independent regions of turbulence divided by regions of strong local magnetic shear. However, if these regions are fully independent, one could just run individual simulations for each. For the remainder of this section we will use the standard binormal wavenumber $\tilde{K}_y = k_y$.

The radial wavenumber coordinate has even more possibilities. The most general possible form is

$$\tilde{K}_x \equiv \tilde{f}_4(k_x, k_y, \chi). \quad (65)$$

As with \tilde{K}_y , if we want to maintain an evenly spaced radial grid, \tilde{K}_x must have a linear relationship with k_x according to

$$\tilde{K}_x \equiv \tilde{f}_5(k_y, \chi) k_x + \tilde{f}_6(k_y, \chi). \quad (66)$$

Otherwise, calculating the nonlinear term using Fourier transforms to real space will become much more expensive.

Interestingly, it is unclear if the form of equation (66) is sufficient to guarantee an efficient treatment of the nonlinear term. This is because, while the values of \tilde{K}_x are evenly spaced for every value of k_y , the \tilde{K}_x grid spacing is *not* necessarily the same between different k_y values. Nevertheless, it may still be possible to efficiently calculate the nonlinear term in mixed (x, k_y) space on a non-uniform x grid. Alternatively, one may be able to efficiently treat grids composed of blocks of k_y modes with the same \tilde{K}_x spacing by Fourier transforming the blocks independently and then combining them in real space. The feasibility of grids with k_y -dependent \tilde{K}_x spacing is left for future work. We point out this category because it includes interesting possibilities such as $\tilde{K}_x \equiv k_x k_{y,\min}/k_y$ (see figure 13(c)), which is similar to the conventional flux tube except it prevents the linear modes from being closely clustered around zero ballooning angle at high values of k_y .

However, efficiency aside, if the \tilde{K}_x grid does not have the same spacing at all k_y , then Fourier modes can again nonlinearly couple to modes that do not exist *within* the grid. While this may be solvable in some cases, we will set this problem aside in this work. Instead we will ensure a consistent treatment of the nonlinear coupling by using the form

$$\tilde{K}_x \equiv \tilde{f}_7(\chi) k_x + \tilde{f}_8(k_y, \chi). \quad (67)$$

Varying the radial domain width using $\tilde{f}_7(\chi)$ appears feasible and has similar considerations and consequences as varying the binormal domain width in equation (64). Unlike with \tilde{K}_y , in this case the offset term $\tilde{f}_8(k_y, \chi)$ may be useful because the turbulence is not always centered around $\tilde{K}_x = 0$. Therefore, we will maintain the offset but choose to have a radial box width that is independent of χ by assuming the form

$$\tilde{K}_x \equiv k_x + \tilde{f}_8(k_y, \chi). \quad (68)$$

A radial coordinate defined by equation (68) has the same grid spacings as the conventional flux tube, but grid points can have a k_y and χ dependent offset. For example, the non-twisting flux tube chooses $\tilde{f}_8(k_y, \chi) = k_y \vec{\nabla}x \cdot \vec{\nabla}y / |\vec{\nabla}x|^2$. This linear shift in the center of the radial grid with k_y can be seen clearly in figure 8(b). It is the reason why the turbulent activity in the non-twisting flux tube remains centered in the perpendicular wavenumber grid at $\chi = 2\pi$, while it stretches diagonally across the grid in the conventional flux tube. It appears feasible to implement a choice of $\tilde{f}_8(k_y, \chi)$ that has a complicated dependence on k_y , but we do not see any particular physical motivation for this. Instead we expect a relatively simple dependence from the competition of the two physical effects discussed earlier in this paper — parallel streaming motivates turbulence centered around $k_x = 0$ and finite gyroradius damping motivates turbulence centered around $k_x = -k_y \vec{\nabla}x \cdot \vec{\nabla}y / |\vec{\nabla}x|^2$. To accommodate both of these k_y dependencies and any combination in between, we assume the form

$$\tilde{K}_x \equiv k_x + k_y \tilde{f}_{tw}(\chi). \quad (69)$$

As shown for an example in figure 13(d), this simply allows a shift in the included K_x modes that varies with parallel location. Notably, the form makes it possible to maintain a Fourier series analogous to equations (10) and (39) by defining a real space binormal coordinate

$$\tilde{Y}(x, y, \chi) \equiv y - \tilde{f}_{tw}(\chi) x \quad (70)$$

such that $\tilde{K}_x x + k_y \tilde{Y} = k_x x + k_y y$. Thus, we see a clear physical interpretation of the function $\tilde{f}_{tw}(\chi)$ — it controls the twist of the flux tube cross-section as a function of parallel location. This flexibility has the potential to be useful.

To implement this, we must first determine the boundary conditions. Repeating the derivations from sections 2 and 3 for \tilde{K}_x and \tilde{Y} , we find

$$\tilde{\Phi}(x + L_x, \tilde{Y} - \tilde{f}_{tw}(\chi) L_x, \chi) = \tilde{\Phi}(x, \tilde{Y}, \chi) \quad (71)$$

$$\tilde{\Phi}(x, \tilde{Y} + L_y, \chi) = \tilde{\Phi}(x, \tilde{Y}, \chi) \quad (72)$$

$$\tilde{\Phi}(x, \tilde{Y} \mp 2\pi N_{\text{pol}} \hat{s}x - (\tilde{f}_{tw}(\chi + 2\pi N_{\text{pol}}) - \tilde{f}_{tw}(\chi)) x, \chi + 2\pi N_{\text{pol}}) = \tilde{\Phi}(x, \tilde{Y}, \chi) \quad (73)$$

in real space. In (\tilde{K}_x, k_y) Fourier-space, we find equation (14) and

$$\tilde{\Phi} \left(\tilde{K}_x \pm 2\pi N_{\text{pol}} k_y \hat{s} + k_y \left(\tilde{f}_{\text{tw}}(\chi + 2\pi N_{\text{pol}}) - \tilde{f}_{\text{tw}}(\chi) \right), k_y, \chi + 2\pi N_{\text{pol}} \right) = \tilde{\Phi} \left(\tilde{K}_x, k_y, \chi \right) \quad (74)$$

$$\tilde{K}_x \in \frac{2\pi}{L_x} (m + \tilde{m}_0(k_y, \chi)) + k_y \tilde{f}_{\text{tw}}(\chi) \quad \text{for all } m \in \left[-\frac{N_x - 1}{2}, \frac{N_x - 1}{2} \right], \quad (75)$$

where

$$\tilde{m}_0(k_y, \chi) = -\text{NINT} \left[\frac{L_x}{2\pi} k_y \tilde{f}_{\text{tw}}(\chi) \right]. \quad (76)$$

Note that the functions $\bar{\bar{\Phi}}$ and $\tilde{\Phi}$ for the electrostatic potential are defined analogously to equations (29) and (43) as

$$\bar{\bar{\Phi}}(x, \tilde{Y}, \chi) = \bar{\phi}(x, \tilde{Y} + \tilde{f}_{\text{tw}}(\chi)x, \chi) = \bar{\phi}(x, y(x, \tilde{Y}, \chi), \chi) \quad (77)$$

$$\tilde{\Phi}(\tilde{K}_x, k_y, \chi) = \phi(\tilde{K}_x - k_y \tilde{f}_{\text{tw}}(\chi), k_y, \chi) = \phi(k_x(\tilde{K}_x, k_y, \chi), k_y, \chi). \quad (78)$$

Like before, we can combine the real space radial and parallel boundary conditions of equations (71) and (73) (as well as equation (72)) to find constraints on the domain. However, here we actually find two conditions. First, we find the standard discretization of the domain aspect ratio (i.e. equation (9)) by evaluating the parallel boundary condition at $x \rightarrow x + L_x$, applying the radial boundary condition to both sides, and then applying the parallel boundary condition to the left side. Second, by evaluating the radial boundary condition at $\chi \rightarrow \chi + 2\pi N_{\text{pol}}$, applying the parallel boundary condition to both sides, applying the radial boundary condition to the left side, and then using equation (9), we find

$$\frac{\tilde{f}_{\text{tw}}(\chi + 2\pi N_{\text{pol}}) - \tilde{f}_{\text{tw}}(\chi)}{2\pi N_{\text{pol}} |\hat{s}|} = \frac{N_{\mathbb{Z}}}{N_{\text{asp}}}, \quad (79)$$

where $N_{\mathbb{Z}} \in \mathbb{Z}$ is some integer. Equation (79) is a novel result and is interesting. We see that any $2\pi N_{\text{pol}}$ -periodic function can be added to $\tilde{f}_{\text{tw}}(\chi)$ without causing concern. However, the global twist of the simulation domain is discretized by the magnetic shear just like the domain aspect ratio. Finishing the calculation, we find that, despite the generality of this transformation, the geometric coefficients and equilibrium quantities

retain simple forms of

$$\vec{\nabla}x \cdot \vec{\nabla}\tilde{Y} = \vec{\nabla}x \cdot \vec{\nabla}y - \tilde{f}_{\text{tw}}(\chi) \left| \vec{\nabla}x \right|^2 \quad (80)$$

$$\left| \vec{\nabla}\tilde{Y} \right|^2 = \left| \vec{\nabla}y \right|^2 - 2\tilde{f}_{\text{tw}}(\chi) \vec{\nabla}x \cdot \vec{\nabla}y + \tilde{f}_{\text{tw}}^2(\chi) \left| \vec{\nabla}x \right|^2 \quad (81)$$

$$\vec{\nabla}\tilde{Y} \cdot \vec{\nabla}\chi = \vec{\nabla}y \cdot \vec{\nabla}\chi - \tilde{f}_{\text{tw}}(\chi) \vec{\nabla}x \cdot \vec{\nabla}\chi \quad (82)$$

$$J^{-1} \equiv \mp \left(\vec{\nabla}x \times \vec{\nabla}\tilde{Y} \right) \cdot \hat{b} = \mp \left(\vec{\nabla}x \times \vec{\nabla}y \right) \cdot \hat{b} \quad (83)$$

$$\vec{B} = \frac{1}{C_y} \frac{d\psi}{dx} \vec{\nabla}x \times \vec{\nabla}y = \frac{1}{C_y} \frac{d\psi}{dx} \vec{\nabla}x \times \vec{\nabla}\tilde{Y} \quad (84)$$

$$\vec{\nabla}B = \left(\frac{\partial B}{\partial x} + \tilde{f}_{\text{tw}}(\chi) \frac{\partial B}{\partial y} \right) \vec{\nabla}x + \frac{\partial B}{\partial y} \vec{\nabla}\tilde{Y} + \frac{\partial B}{\partial \chi} \vec{\nabla}\chi. \quad (85)$$

In summary, we see that it is possible to implement a wide variety of flux tubes apart from just the conventional and non-twisting versions. While it is currently unclear if certain exotic options are computationally feasible and useful, it does appear straightforward to achieve complete control of the twist of the flux tube along its length by defining the radial wavenumber according to equation (69). In fact, all local computational domains considered so far in the literature can be expressed through different choices of the function $\tilde{f}_{\text{tw}}(\chi)$. Specifically, the conventional flux tube, the non-twisting flux tube, the globally non-twisting flux tube (which only removes the twist from the *global* magnetic shear), and the flux tube train [13] are given by

$$\tilde{f}_{\text{tw}}(\chi) = 0 \quad (86)$$

$$\tilde{f}_{\text{tw}}(\chi) = \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{\left| \vec{\nabla}x \right|^2} \quad (87)$$

$$\tilde{f}_{\text{tw}}(\chi) = \mp \hat{s}\chi \quad (88)$$

$$\tilde{f}_{\text{tw}}(\chi) = \mp 2\pi \hat{s} \text{NINT} \left[\frac{\chi}{2\pi} \right] = \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{\left| \vec{\nabla}x \right|^2} \Bigg|_{\chi' = 2\pi \text{NINT}[\chi/(2\pi)]} \quad (89)$$

respectively. The first three have already been implemented in GENE and any choice of $\tilde{f}_{\text{tw}}(\chi)$ looks simple to add. For unusual turbulence, finding the best choice of $\tilde{f}_{\text{tw}}(\chi)$ may offer significant computational savings. More broadly, it may be possible to formulate an automated routine that performs linear tests to optimize $\tilde{f}_{\text{tw}}(\chi)$ before beginning a nonlinear simulation or adapts the choice of $\tilde{f}_{\text{tw}}(\chi)$ during a nonlinear simulation.

7. Conclusions

In this paper, we have presented a novel derivation of the “non-twisting flux tube” — a local simulation domain that does not twist due to the effects of magnetic shear. It

maintains a field-aligned spatial grid, but abandons having a field-aligned domain. In other words, grid points are placed in rows along magnetic field lines, but the boundaries of the domain do *not* follow field lines. This is made possible by allowing field lines to pass through the binormal boundary of the domain, but, whenever they exit, binormal periodicity is invoked so that they immediately enter back in on the opposite side. Importantly, this derivation shows that such a flux tube is consistent with a periodic radial boundary condition and the Fourier representation that is typically used in the perpendicular directions of local domains.

By comparing the spatial coordinate grids of the non-twisting and conventional flux tubes, we found that they are optimized to model different physical effects. The conventional flux tube grid prioritizes parallel streaming and is constructed to follow linear modes along field lines from the outboard midplane. On the other hand, the non-twisting flux tube grid prioritizes the Fourier modes that are least damped by finite gyroradius effects (i.e. those with a small *local* radial wavenumber K_x). Thus, although both simulation domains should give the same result in the limit of infinite radial resolution, one may be more computationally efficient and allow accurate simulations at lower grid resolution.

To investigate this, we successfully implemented and benchmarked the non-twisting flux tube in the gyrokinetic code GENE. Then, using five different test cases, we analyzed the runtime and convergence properties. We found *no* case where the non-twisting flux tube performed worse than the conventional flux tube. Moreover, we found cases where the non-twisting flux tube was seven, or even thirty times less expensive. This computational savings came partially from a longer time step and partially from enabling a lower radial resolution. The non-twisting flux tube was found to perform best when the magnetic shear was high and when there was more than one region of turbulent drive in the parallel direction. This makes it potentially useful for stellarator simulations, pedestal simulations, and tokamak simulations with several poloidal turns.

Lastly, we showed that the non-twisting flux tube is just one example of many possible flux tube formulations. By generalizing the non-twisting flux tube derivation, we explored the feasibility of some exotic flux tube grids. Moreover, we demonstrated that it is straightforward to fully control how the flux tube twists as it extends along the field line. Thus, it may be possible to tailor the flux tube cross-section to the particular turbulence being modeled. This could enable even more efficient spatial grids and, ultimately, significant computational savings.

Acknowledgments

The authors would like to thank Tobias Goerler, Daniel Told, and Ajay C. J. for useful discussions pertaining to this work. This work was supported by the EUROfusion - Theory and Advanced Simulation Coordination (E-TASC). This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant

agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission. We acknowledge the CINECA award under the ISCRA initiative, for the availability of high performance computing resources and support. This work was supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s863, s956, and 1050. This work was carried out using the JFRS-1 supercomputer system at Computational Simulation Centre of International Fusion Energy Research Centre (IFERC-CSC) in Rokkasho Fusion Institute of QST (Aomori, Japan).

Appendix A. The straight-field line and non-twisting poloidal angles

We are seeking a field line label y to construct a field-aligned coordinate system. We already have a radial coordinate x that selects the flux surface. The coordinate y will be used to select the field line and any arbitrary poloidal angle θ will determine the parallel position. Since y must be constant along a magnetic field line, we know it must respect the field line trajectory relation

$$\left. \frac{\partial \zeta}{\partial \theta} \right|_{x,y} = \frac{\vec{B} \cdot \vec{\nabla} \zeta}{\vec{B} \cdot \vec{\nabla} \theta}, \quad (\text{A.1})$$

where the vertical bar indicates the quantities that are held constant in the operation. This says that moving poloidally while staying on a field line (i.e. staying at constant x and y) involves moving toroidally an amount proportional to the field line pitch. Integrating this equation gives

$$\zeta = \int_0^\theta \left. d\theta' \frac{\vec{B} \cdot \vec{\nabla} \zeta}{\vec{B} \cdot \vec{\nabla} \theta'} \right|_{x,y} + f(x, y), \quad (\text{A.2})$$

where we see that y only appears in the integration constant. Choosing a form for $f(x, y)$ is unconstrained and reflects the flexibility we have in defining y . The conventional choice is $f(x, y) = \pm y/C_y$, which after rearranging gives the typical definition [18] of

$$y = \pm C_y \left(\zeta - \int_0^\theta \left. d\theta' \frac{\vec{B} \cdot \vec{\nabla} \zeta}{\vec{B} \cdot \vec{\nabla} \theta'} \right|_x \right). \quad (\text{A.3})$$

Here we have indicated that the integral is performed at constant x , but do not need to specify a second constant because of toroidal symmetry. Now this definition of y holds for any poloidal angle θ (including the geometric poloidal angle). However, if we chose to define the poloidal angle according to

$$\chi(x, \theta) \equiv \frac{1}{q(x)} \int_0^\theta \left. d\theta' \frac{\vec{B} \cdot \vec{\nabla} \zeta}{\vec{B} \cdot \vec{\nabla} \theta'} \right|_x, \quad (\text{A.4})$$

we can substitute it into equation (A.3) and arrive at the simpler definition of y given by equation (1). This poloidal angle χ is called a ‘‘straight-field line’’ poloidal angle

because we can invert equation (1) to directly calculate that

$$\left. \frac{\partial \zeta}{\partial \chi} \right|_{x,y} = q(x). \quad (\text{A.5})$$

In other words, the magnetic field lines appears as straight lines when plotted in the (χ, ζ) coordinate plane.

In this paper, we use the straight-field line poloidal angle χ because it is what is used by the GENE code. Moreover, it makes it straightforward to calculate the geometric coefficients for the non-twisting flux tube from the geometric coefficients that were already calculated for the conventional flux tube (i.e. according to equations (53) through (58)). However, if one was coding a non-twisting flux tube from scratch, there is another poloidal coordinate that may be more convenient. Analogously to the new non-twisting binormal coordinate Y , one can define a new “non-twisting poloidal angle” that has no minor radial variation and is a straight-field line poloidal angle only at $x = 0$. Such a non-twisting poloidal angle is defined by

$$\vartheta(\theta) \equiv \chi(x = 0, \theta) = \frac{1}{q_0} \int_0^\theta \bigg|_{x=0} d\theta' \frac{\vec{B} \cdot \vec{\nabla} \zeta}{\vec{B} \cdot \vec{\nabla} \theta'}. \quad (\text{A.6})$$

This is useful primarily because $\vec{\nabla} x \cdot \vec{\nabla} \vartheta = 0$, which will allow it to cleanly separate the effects of global magnetic shear from the local corrections. To see how, substitute the definition of ϑ and the radial derivative of the definition of

$$q(x) \equiv \frac{1}{2\pi} \oint_0^{2\pi} \bigg|_x d\theta' \frac{\vec{B} \cdot \vec{\nabla} \zeta}{\vec{B} \cdot \vec{\nabla} \theta'}, \quad (\text{A.7})$$

into a Taylor expansion of equation (A.3) about $x = 0$. This gives

$$y = \pm C_y \zeta \mp C_y q_0 \vartheta \mp \hat{s} \vartheta x \mp x \int_0^\vartheta d\vartheta' \hat{l}(\vartheta'), \quad (\text{A.8})$$

where

$$\hat{l}(\vartheta) \equiv C_y \frac{d\theta}{d\vartheta} \left. \frac{\partial}{\partial x} \right|_\theta \frac{\vec{B} \cdot \vec{\nabla} \zeta}{\vec{B} \cdot \vec{\nabla} \theta} - \frac{C_y}{2\pi} \oint_0^{2\pi} \bigg|_{x=0} d\theta' \left(\left. \frac{\partial}{\partial x} \right|_\theta \frac{\vec{B} \cdot \vec{\nabla} \zeta}{\vec{B} \cdot \vec{\nabla} \theta} \right) \quad (\text{A.9})$$

is the local correction to the magnetic shear (which is defined to be the total shear minus the global shear) and $\theta(\vartheta)$ can be computed by inverting equation (A.6).

Comparing equation (A.8) with equation (3), we see that a new term has appeared, which makes explicit the local correction to the magnetic shear. This is the advantage of the new poloidal angle. Previously, the local correction was hidden away within the definition of the poloidal angle χ , specifically in the second term of equation (3). In this new non-twisting poloidal angle, one can use equation (A.8) to show that the geometric

coefficients of the conventional flux tube are

$$\vec{\nabla}y = \pm C_y \vec{\nabla}\zeta \mp C_y q_0 \vec{\nabla}\vartheta \mp \hat{s}\vartheta \vec{\nabla}x \mp \vec{\nabla}x \int_0^\vartheta d\vartheta' \hat{l}(\vartheta') \quad (\text{A.10})$$

$$\frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2} = \mp \hat{s}\vartheta \mp \int_0^\vartheta d\vartheta' \hat{l}(\vartheta') \quad (\text{A.11})$$

$$|\vec{\nabla}y|^2 = C_y^2 |\vec{\nabla}\zeta|^2 + C_y^2 q_0^2 |\vec{\nabla}\vartheta|^2 + \left(\hat{s}\vartheta + \int_0^\vartheta d\vartheta' \hat{l}(\vartheta') \right)^2 |\vec{\nabla}x|^2, \quad (\text{A.12})$$

which look pretty similar to what they were using the typical straight-field line poloidal angle χ (e.g. equations (23) and (25)). However, using equation (24) we can calculate the geometric coefficients of the non-twisting flux tube to be

$$Y = \pm C_y \zeta \mp C_y q_0 \vartheta \quad (\text{A.13})$$

$$\vec{\nabla}Y = \pm C_y \vec{\nabla}\zeta \mp C_y q_0 \vec{\nabla}\vartheta \quad (\text{A.14})$$

$$\vec{\nabla}x \cdot \vec{\nabla}Y = 0 \quad (\text{A.15})$$

$$|\vec{\nabla}Y|^2 = C_y^2 |\vec{\nabla}\zeta|^2 + C_y^2 q_0^2 |\vec{\nabla}\vartheta|^2 \quad (\text{A.16})$$

$$\vec{\nabla}Y \cdot \vec{\nabla}\vartheta = \mp C_y q_0 |\vec{\nabla}\vartheta|^2, \quad (\text{A.17})$$

$$\vec{\nabla}x \cdot \vec{\nabla}\vartheta = 0, \quad (\text{A.18})$$

which are considerably simpler than equations (26), (48), (54), and (55).

Appendix B. Computational implementation

In this appendix we will outline the practicalities of implementing the non-twisting flux tube in the gyrokinetic code GENE [17, 18]. If we compare the equations for the non-twisting flux tube (i.e. equations (14), (17), (42), (45), (46), (49), and (51)) with the equations for the conventional flux tube (i.e. equations (13), (14), (15), (16), (17), (18), and (20)) we see that relatively few changes are required. The binormal k_y grid requires no changes. The geometric quantities must be transformed, but this is simple to accomplish using equations (53) through (58).

However, the modification of the radial grid is more substantial. It should be constructed according to equation (45), which has different values of K_x at each parallel location χ and for each binormal wavenumber k_y . Changing the radial grid to a three dimensional array would have ramifications all throughout the code, so we will make some adjustments to the analytic derivation to avoid it. To do this, we notice that, even though the physical meanings of the ballooning radial wavenumber k_x and the local radial wavenumber K_x are quite different, the numerical values of the grid are very similar. Equation (13) is a rectangular grid centered around $k_x = 0$ with a regular grid spacing of $2\pi/L_x$ and equation (45) is a rectangular grid centered around $K_x \approx 0$ with a regular spacing of $2\pi/L_x$. We see that the only difference in the numerical values is a

small shift of less than a grid point that comes from being centered at $K_x \approx 0$ instead of $K_x = 0$. Thus, we will write the local radial wavenumber as a sum of two contributions

$$K_x = \mathcal{K}_x + \Delta K_x, \quad (\text{B.1})$$

where

$$\Delta K_x \equiv k_y \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2} - \frac{2\pi}{L_x} \text{NINT} \left[\frac{L_x}{2\pi} k_y \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2} \right] \quad (\text{B.2})$$

is the small shift that is the only difference between the K_x and k_x modes in the non-twisting and conventional flux tubes. By substituting equations (B.1) and (B.2) into equation (45), we find that the grid for \mathcal{K}_x is

$$\mathcal{K}_x \in \frac{2\pi}{L_x} m \quad \text{for all } m \in \left[-\frac{N_x - 1}{2}, \frac{N_x - 1}{2} \right], \quad (\text{B.3})$$

which no longer depends on k_y nor χ and is numerically identical to the k_x grid of the conventional flux tube (i.e. equation (13)). Thus, if we used \mathcal{K}_x as the radial wavenumber, the radial grid would not require any modification. To do so, we substitute equation (B.1) into the gyrokinetic model for the non-twisting flux tube (i.e. equation (49)) to find

$$\begin{aligned} \frac{\partial \mathcal{H}_s}{\partial t} + v_{\parallel} \hat{b} \cdot \vec{\nabla} \chi \frac{\partial \mathcal{H}_s}{\partial \chi} \Big|_{\mathcal{K}_x + \Delta K_x - k_y \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2}, k_y} + i \vec{v}_{ds} \cdot \left(\mathcal{K}_x \vec{\nabla}x + \Delta K_x \vec{\nabla}x + k_y \vec{\nabla}Y \right) \mathcal{H}_s + a_{s\parallel} \frac{\partial \mathcal{H}_s}{\partial v_{\parallel}} \\ \mp \frac{1}{JB} \{ \mathcal{H}_s, \varphi J_0(\mathcal{K}_{\perp} \rho_s) \} = \frac{Z_s e F_{Ms}}{T_s} \frac{\partial \varphi}{\partial t} J_0(\mathcal{K}_{\perp} \rho_s) \mp i \frac{k_y}{JB} \varphi J_0(\mathcal{K}_{\perp} \rho_s) \frac{\partial F_{Ms}}{\partial x}, \end{aligned} \quad (\text{B.4})$$

where the perpendicular wavenumber is

$$\mathcal{K}_{\perp} = \sqrt{(\mathcal{K}_x + \Delta K_x)^2 |\vec{\nabla}x|^2 + k_y^2 |\vec{\nabla}Y|^2} \quad (\text{B.5})$$

and the quasineutrality equation remains unchanged except for the new form of the perpendicular wavenumber, the electrostatic potential becomes a function of \mathcal{K}_x according to

$$\varphi(\mathcal{K}_x, k_y, \chi) = \phi \left(\mathcal{K}_x + \Delta K_x - k_y \frac{\vec{\nabla}x \cdot \vec{\nabla}y}{|\vec{\nabla}x|^2}, k_y, \chi \right) = \phi(k_x(\mathcal{K}_x, k_y, \chi), k_y, \chi), \quad (\text{B.6})$$

and the distribution function \mathcal{H}_s is defined analogously. The nonlinear term is computed

in real space according to

$$\begin{aligned}
 \{\mathcal{H}_s, \varphi J_0(\mathcal{K}_\perp \rho_s)\} &= \frac{1}{N_x(2N_y - 1)} \sum_{x,Y} \left[\left(\sum_{\mathcal{K}'_x, k'_y} (\mathcal{K}'_x + \Delta K'_x) \mathcal{H}'_s e^{i\Delta K'_x x} e^{i\mathcal{K}'_x x + ik'_y Y} \right) \right. \\
 &\quad \times \left(\sum_{\mathcal{K}''_x, k''_y} k''_y \varphi'' J_0(\mathcal{K}_\perp \rho_s) e^{i\Delta K''_x x} e^{i\mathcal{K}''_x x + ik''_y Y} \right) \\
 &\quad - \left(\sum_{\mathcal{K}'_x, k'_y} k'_y \mathcal{H}'_s e^{i\Delta K'_x x} e^{i\mathcal{K}'_x x + ik'_y Y} \right) \\
 &\quad \left. \times \left(\sum_{\mathcal{K}''_x, k''_y} (\mathcal{K}''_x + \Delta K''_x) \varphi'' J_0(\mathcal{K}_\perp \rho_s) e^{i\Delta K''_x x} e^{i\mathcal{K}''_x x + ik''_y Y} \right) \right] e^{-i\Delta K_x x} e^{-i\mathcal{K}_x x - ik_y Y}
 \end{aligned} \tag{B.7}$$

using the three-halves rule as before and the modified Fourier transforms

$$\bar{\varphi}(x, Y, \chi) = \sum_{\mathcal{K}_x, k_y} \varphi(\mathcal{K}_x, k_y, \chi) e^{i\Delta K_x x} e^{i\mathcal{K}_x x + ik_y Y} \tag{B.8}$$

$$\varphi(\mathcal{K}_x, k_y, \chi) = \frac{1}{N_x(2N_y - 1)} \sum_{x,Y} \bar{\varphi}(x, Y, \chi) e^{-i\Delta K_x x} e^{-i\mathcal{K}_x x - ik_y Y}. \tag{B.9}$$

This eliminates the need to modify the radial wavenumber grid because the preexisting k_x grid is identical to the new \mathcal{K}_x grid. Instead all we must do is add the small ΔK_x correction terms in the perpendicular wavenumber within the argument to the Bessel functions, the magnetic drift term, and the nonlinear term (including the exponential phase factors in the Fourier transforms).

Additionally, there is a numerical subtlety related to the nearest integer rounding used in constructing the new radial grid. The NINT function that appears in equations (44), (B.2), etc. behaves unpredictably when its argument evaluates to half-integer values. This is because any incidental round-off errors will then determine whether the function rounds up or down. This is particularly important at the $\chi = -\pi N_{\text{pol}}$ grid point, as can be seen from equation (60). Note that GENE does *not* put a grid point at $\chi = \pi N_{\text{pol}}$, so we do not need to worry about it. To ensure well-defined and consistent rounding, instead of calculating the argument of NINT at a given grid location χ , it is actually approximated at a slightly larger value using the linear interpolation

$$\text{NINT} \left[\frac{L_x k_y}{2\pi} \frac{\vec{\nabla} x \cdot \vec{\nabla} y}{|\vec{\nabla} x|^2} \Big|_\chi \right] \rightarrow \text{NINT} \left[\frac{L_x k_y}{2\pi} \left((1 - \delta) \frac{\vec{\nabla} x \cdot \vec{\nabla} y}{|\vec{\nabla} x|^2} \Big|_\chi + \delta \frac{\vec{\nabla} x \cdot \vec{\nabla} y}{|\vec{\nabla} x|^2} \Big|_{\chi + \Delta\chi} \right) \right], \tag{B.10}$$

where $\Delta\chi$ is the parallel grid spacing. This means that, for typical simulations, the $\chi = -\pi N_{\text{pol}}$ grid point will be rounded in the same way as the $\chi = -\pi N_{\text{pol}} + \Delta\chi$ grid point. The constant $\delta = 0.01313$ was chosen fairly arbitrarily to be small enough to

be negligible from a physical perspective, but large enough to overcome the significant numerical errors that can be present when using geometric coefficients calculated by external MHD equilibrium codes. It is also important that δ is not a rational number with a small denominator, otherwise such a round-off problem may inadvertently occur at another parallel grid location.

The last and most significant challenge to implement the non-twisting flux tube is the parallel derivative because it is crucial to maintain an exactly field-aligned grid. In a conventional flux tube this is straightforward to accomplish because taking a derivative exactly along the field line corresponds to taking the derivative at constant Fourier mode *index*. To see this, note that the parallel derivative in equation (16) is taken at constant k_x and k_y , which is the same as holding the Fourier mode indexes m and n constant because of equations (13) and (14). In the code, these mode numbers directly correspond to the array indexes of the variables for h_s and ϕ , so standard finite difference schemes along χ naturally accomplish what we want. For example, the fourth-order centered finite difference scheme typically used by GENE calculates the parallel derivative at a point using the value of h_s from the nearest two parallel locations on each side. Writing this at constant m and n is simply

$$\left. \frac{\partial h_s}{\partial \chi} \right|_{k_x, k_y} [m, n, l] \approx \frac{1}{12\Delta\chi} (-h_s [m, n, l - 2] + 8h_s [m, n, l - 1] - 8h_s [m, n, l + 1] + h_s [m, n, l + 2]), \quad (\text{B.11})$$

where we are treating h_s as an array with the index m in the k_x direction, n in the k_y direction, and l in the χ direction (assuming uniform grid spacing in χ). The primary challenge for the conventional flux tube is taking the parallel derivative across the parallel boundary because it requires correctly coupling k_x modes according to equation (15). In the non-twisting flux tube, this complicated mode coupling gets spread out over the interior of the domain.

Thus, for the non-twisting flux tube, we must be careful to select the proper radial wavenumber index at the various parallel locations in our parallel finite difference. This is accomplished, in accordance with equation (B.4), by selecting the indexes that have the same value of $\mathcal{K}_x + \Delta K_x - k_y \vec{\nabla} x \cdot \vec{\nabla} y / |\vec{\nabla} x|^2$. By substituting equations (B.2) and (B.3), we see that

$$\mathcal{K}_x + \Delta K_x - k_y \frac{\vec{\nabla} x \cdot \vec{\nabla} y}{|\vec{\nabla} x|^2} = \frac{2\pi}{L_x} (m + m_0(k_y, \chi)), \quad (\text{B.12})$$

where $m_0(k_y, \chi)$ is given by equation (44). Thus, instead of holding m and n constant as in the conventional flux tube, we hold $m + m_0(k_y, \chi)$ and n constant in the non-twisting flux tube. While this may seem like a minor change, it is fairly significant because m_0 has a piecewise dependence on k_y and χ . In other words, we are now taking the parallel derivative diagonally across the (\mathcal{K}_x, k_y) grid. However, because the grid is composed

of discrete points, it is really a diagonal line that is everywhere rounded to the closest \mathcal{K}_x grid location. Thus, the fourth-order centered finite difference scheme becomes

$$\begin{aligned} \left. \frac{\partial \mathcal{H}_s}{\partial \chi} \right|_{\mathcal{K}_x + \Delta K_x - k_y \frac{\vec{\nabla}_x \cdot \vec{\nabla}_y}{|\vec{\nabla}_x|^2}, k_y} [m, n, l] &\approx \frac{1}{12\Delta\chi} (-\mathcal{H}_s [m + m_0 [n, l] - m_0 [n, l - 2], n, l - 2] \\ &+ 8\mathcal{H}_s [m + m_0 [n, l] - m_0 [n, l - 1], n, l - 1] \\ &- 8\mathcal{H}_s [m + m_0 [n, l] - m_0 [n, l + 1], n, l + 1] \\ &+ \mathcal{H}_s [m + m_0 [n, l] - m_0 [n, l + 2], n, l + 2]) \end{aligned} \quad (\text{B.13})$$

because the radial wavenumber index m' at a different location l' must satisfy the relationship $m' + m_0 [n, l'] = m + m_0 [n, l]$.

This is implemented together with the parallel boundary condition, which the derivation of section 3 indicates is simpler than for the conventional flux tube. However, we still must derive its precise form for the slightly modified local radial wavenumber \mathcal{K}_x . Equation (42) shows it should be taken holding K_x and k_y constant according to

$$\varphi(\mathcal{K}_x(K_x, k_y, \chi + 2\pi N_{\text{pol}}), k_y, \chi + 2\pi N_{\text{pol}}) = \varphi(\mathcal{K}_x(K_x, k_y, \chi), k_y, \chi). \quad (\text{B.14})$$

Using equations (9), (34), (B.1), and (B.2), we see that $\mathcal{K}_x(K_x, k_y, \chi + 2\pi N_{\text{pol}}) = \mathcal{K}_x(K_x, k_y, \chi)$. Therefore, the parallel boundary condition is

$$\varphi(\mathcal{K}_x, k_y, \chi + 2\pi N_{\text{pol}}) = \varphi(\mathcal{K}_x, k_y, \chi), \quad (\text{B.15})$$

which can be shown to be consistent with the conventional parallel boundary condition of equation (15) by using equations (34), (41), (B.1), and (B.6). Thus, we can see that the parallel boundary condition for coding the non-twisting flux tube is actually a subset of its prior form. This makes it easy to implement — we simply treat all modes in the same way that zonal modes were treated in the conventional flux tube.

This completely specifies the GENE implementation of the non-twisting flux tube, which is defined by equations (14), (17) (using φ , \mathcal{H}_s , and \mathcal{K}_\perp), (B.2), (B.3), (B.4), (B.7), (B.5), (B.13), and (B.15). However, there are still a few important numerical details. As a result of the numerical scheme of equation (B.11) or (B.13), grid points with odd parallel indexes are closely tied together, as are grid points with even parallel indexes, but the connection between these two subsets is weak (see section 3.1.2 of reference [29]). This can cause problems when the total number of parallel grid points in a linear mode is odd, because the two subsets will not have an equal number of points. Thus, unless the linear mode is very long, the smaller subset will be more affected by the Dirichlet boundary condition at the ends of the linear mode (i.e. $\varphi = 0$). This can lead to grid-scale oscillations in the parallel direction that significantly affect accuracy. In the conventional flux tube, this is typically addressed by simply choosing N_χ to be even. However, in the non-twisting flux tube, linear modes are not constrained to span an integer number of poloidal turns. In fact, even at the same value of k_y , the number

of χ grid points in the individual linear modes (with different ballooning angles) can differ by one. This makes it practically impossible to ensure that all linear modes are composed of an even number of χ grid points. To solve this, we first tried changing the boundary condition at the ends of the linear modes to be periodic, instead of Dirichlet. Specifically, whenever you would like to use the mode $m = N_x/2$, which is outside the grid given by equation (B.3), you use the $m = -(N_x - 1)/2$ mode at the same χ location instead. This works, but is observed to cause an overestimate of the linear growth rates when the linear modes are short. Instead, a better solution was to retain the Dirichlet boundary conditions, but to calculate the number of grid points in each linear mode. If it is odd, you artificially remove a grid point from that mode (i.e. the point with the lowest value of \mathcal{K}_x). This is especially important for the highest k_y modes in nonlinear simulations because the linear modes can have a very short extent in χ . In fact, when using large k_y grids, small \mathcal{K}_x grids, and a coarse parallel resolution, the highest values of k_y can have linear modes with just one parallel grid point. If such linear modes are not omitted from the simulation, they can cause fictitiously large heat fluxes. To test that the code treats such modes properly without running full nonlinear simulations, it is helpful to use linear simulations with large N_{asp} compared to N_x .

While implementing the parallel derivative seems complicated in the non-twisting flux tube, in practice it can be done at little computational cost. This is because the coordinate system grids stay fixed throughout the simulation, so all related work can be done when the simulation is initialized and does not need repeating. In practice, it was implemented in GENE by calculating a four-dimensional coupling matrix that is a function of \mathcal{K}_x , k_y , χ , and the points in the finite difference stencil (e.g. four in equation (B.13)). This coupling matrix simply holds the radial mode indexes that should be used in every finite difference computation possible in the simulation. Note that it is necessary for the coupling matrix to be a function of \mathcal{K}_x because, near the edges of the grid, the finite difference will extend off the radial grid. Thus, the entries of the coupling matrix near the boundaries in \mathcal{K}_x are unique as they implement the Dirichlet boundary condition.

Because coupling between different radial modes now occurs all throughout the domain (instead of just at the parallel boundaries), one might worry about the computational cost of data communication between different processors when running on a supercomputer. Fortunately, while GENE has a very flexible parallelization scheme to best distribute the five dimensional gyrokinetic equation across many processors, it does not allow any parallelization in the k_x dimension. This was because the conventional flux tube still requires a lot of data communication between k_x modes for the parallel boundary condition and computing the Fourier transforms in the nonlinear term. Thus, the changes to the parallel derivative needed for the non-twisting flux tube are not expected to significantly increase the data communication between processors. However, the fact that the parallel derivative uses array values that are no longer contiguous in memory likely increases the number of cache misses within each processor. We believe that this carries the most significant computational cost of all the changes required to

implement the non-twisting flux tube.

References

- [1] G.R. McKee, C.C. Petty, R.E. Waltz, C. Fenzi, R.J. Fonck, J.E. Kinsey, T.C. Luce, K.H. Burrell, D.R. Baker, E.J. Doyle, and others. Non-dimensional scaling of turbulence characteristics and turbulent diffusivity. *Nucl. Fusion*, 41(9):1235, 2001.
- [2] B. Geiger, Th. Wegner, C.D. Beidler, R. Burhenn, B. Buttenschön, R. Dux, A. Langenberg, N.A. Pablant, T. Pütterich, Y. Turkin, et al. Observation of anomalous impurity transport during low-density experiments in W7-X with laser blow-off injections of iron. *Nucl. Fusion*, 59(4):046009, 2019.
- [3] P.J. Catto. Linearized gyro-kinetics. *Plasma Phys.*, 20(7):719, 1978.
- [4] E.A. Frieman and L. Chen. Nonlinear gyrokinetic equations for low-frequency electromagnetic waves in general plasma equilibria. *Phys. Fluids*, 25:502, 1982.
- [5] A.J. Brizard and T.S. Hahm. Foundations of nonlinear gyrokinetic theory. *Rev. Mod. Phys.*, 79(2):421, 2007.
- [6] I.G. Abel, G.G. Plunk, E. Wang, M.A. Barnes, S.C. Cowley, W. Dorland, and A.A. Schekochihin. Multiscale gyrokinetics for rotating tokamak plasmas: Fluctuations, transport, and energy flows. *Rep. Prog. Phys.*, 76:116201, 2013.
- [7] J.L. Luxon. A design retrospective of the DIII-D tokamak. *Nucl. Fusion*, 42(5):614, 2002.
- [8] F. Najmabadi, A. Abdou, L. Bromberg, T. Brown, V.C. Chan, M.C. Chu, F. Dahlgren, L. El-Guebaly, P. Heitzenroeder, D. Henderson, and others. The ARIES-AT advanced tokamak, advanced technology fusion power plant. *Fusion Eng. Des.*, 80(1):3, 2006.
- [9] M.A. Beer, S.C. Cowley, and G.W. Hammett. Field-aligned coordinates for nonlinear simulations of tokamak turbulence. *Phys. Plasmas*, 2(7):2687, 1995.
- [10] S.A. Orszag. On the elimination of aliasing in finite-difference schemes by filtering high-wavenumber components. *J. Atmos. Sci.*, 28(6):1074, 1971.
- [11] D. Told and F. Jenko. Applicability of different geometry approaches to simulations of turbulence in highly sheared magnetic fields. *Phys. Plasmas*, 17(4):042302, 2010.
- [12] J. Ball, S. Brunner, and Ajay C.J. Eliminating turbulent self-interaction through the parallel boundary condition in local gyrokinetic simulations. *J. Plasma Phys.*, 86(2):905860207, 2020.
- [13] T.-H. Watanabe, H. Sugama, A. Ishizawa, and M. Nunami. Flux tube train model for local turbulence simulation of toroidal plasmas. *Phys. Plasmas*, 22(2):022507, 2015.
- [14] B. Scott. Shifted metric procedure for flux tube treatments of toroidal geometry: Avoiding grid deformation. *Phys. Plasmas*, 8(2):447, 2001.
- [15] G.W. Hammett, W. Dorland, N.F. Loureiro, and T. Tatsuno. Implementation of Large Scale ExB Shear Flow in the GS2 Gyrokinetic Turbulence Code. *48th Annual Meeting of the APS Division of Plasma Physics*, VP1, 2006.
- [16] B.F. McMillan, J. Ball, and S. Brunner. Simulating background shear flow in local gyrokinetic simulations. *Plasma Phys. Control. Fusion*, 61(5):055006, 2019.
- [17] F. Jenko, W. Dorland, M. Kotschenreuther, and B.N. Rogers. Electron temperature gradient driven turbulence. *Phys. Plasmas*, 7(5):1904, 2000.
- [18] T. Goerler, X. Lapillonne, S. Brunner, T. Dannert, F. Jenko, F. Merz, and D. Told. The global version of the gyrokinetic turbulence code GENE. *J. Comput. Phys.*, 230(18):7053, 2011.
- [19] M. Kotschenreuther, G. Rewoldt, and W.M. Tang. Comparison of initial value and eigenvalue codes for kinetic toroidal plasma instabilities. *Comput. Phys. Commun.*, 88(2):128, 1995.
- [20] W. Dorland, F. Jenko, M. Kotschenreuther, and B.N. Rogers. Electron temperature gradient turbulence. *Phys. Rev. Lett.*, 85(26):5579, 2000.
- [21] F.I. Parra, M. Barnes, and A.G. Peeters. Up-down symmetry of the turbulent transport of toroidal angular momentum in tokamaks. *Phys. Plasmas*, 18(6):062501, 2011.
- [22] M.F. Martin, M. Landreman, P. Xanthopoulos, N.R. Mandell, and W. Dorland. The parallel

- boundary condition for turbulence simulations in low magnetic shear devices. *Plasma Phys. Control. Fusion*, 60(9):095008, 2018.
- [23] M.J. Pueschel, T. Dannert, and F. Jenko. On the role of numerical dissipation in gyrokinetic Vlasov simulations of plasma microturbulence. *Comput. Phys. Commun.*, 181(8):1428, 2010.
- [24] A.M. Dimits, G. Bateman, M.A. Beer, B.I. Cohen, W. Dorland, G.W. Hammett, C. Kim, J.E. Kinsey, M. Kotschenreuther, A.H. Kritz, and others. Comparisons and physics basis of tokamak transport models and turbulence simulations. *Phys. Plasmas*, 7:969, 2000.
- [25] J.M. Greene and M.S. Chance. The second region of stability against ballooning modes. *Nucl. Fusion*, 21(4):453, 1981.
- [26] R. Courant, K. Friedrichs, and H. Lewy. On the partial difference equations of mathematical physics. *IBM Journal of Research and Development*, 11(2):215–234, 1967.
- [27] Ajay C. J., S. Brunner, B. McMillan, J. Ball, J. Dominski, and G. Merlo. How eigenmode self-interaction affects zonal flows and convergence of tokamak core turbulence with toroidal system size. *J. Plasma Phys.*, 86(5):905860504, 2020.
- [28] B.J. Faber, M.J. Pueschel, P.W. Terry, C.C. Hegna, and J.E. Roman. Stellarator microinstabilities and turbulence at low magnetic shear. *J. Plasma Phys.*, 84(5):905840503, 2018.
- [29] F. Merz. *Gyrokinetic simulation of multimode plasma turbulence*. PhD thesis, Universität Münster, 2008.
- [30] D. Told, F. Jenko, P. Xanthopoulos, L.D. Horton, E. Wolfrum, and ASDEX Upgrade Team. Gyrokinetic microinstabilities in ASDEX upgrade edge plasmas. *Phys. Plasmas*, 15(10):102306, 2008.
- [31] F. Jenko, D. Told, P. Xanthopoulos, F. Merz, and L.D. Horton. Gyrokinetic turbulence under near-separatrix or nonaxisymmetric conditions. *Phys. Plasmas*, 16(5):055901, 2009.
- [32] D.P. Fulton, Z. Lin, I. Holod, and Y. Xiao. Microturbulence in DIII-D tokamak pedestal. i. electrostatic instabilities. *Phys. Plasmas*, 21(4):042110, 2014.
- [33] J.F. Parisi, F.I. Parra, C.M. Roach, C. Giroud, W. Dorland, D.R. Hatch, M. Barnes, J.C. Hillesheim, N. Aiba, J. Ball, and P.G. Ivanov. Toroidal and slab ETG instability dominance in the linear spectrum of JET-ILW pedestals. *Nucl. Fusion*, 60(12):126045, 2020.
- [34] M. Barnes, F.I. Parra, and A.A. Schekochihin. Critically balanced ion temperature gradient turbulence in fusion plasmas. *Phys. Rev. Lett.*, 107(11):115003, 2011.
- [35] E. Fable. private communication, 2019.
- [36] W.W. Lee, J.L.V. Lewandowski, T.S. Hahm, and Z. Lin. Shear-Alfvén waves in gyrokinetic plasmas. *Phys. Plasmas*, 8(10):4435, 2001.
- [37] P.B. Snyder. *Gyrofluid theory and simulation of electromagnetic turbulence and transport in tokamak plasmas*. PhD thesis, Princeton University, 1999.
- [38] D. Told. *Gyrokinetic microturbulence in transport barriers*. PhD thesis, Universität Ulm, 2012.
- [39] J.A. Baumgaertel. *Simulating the effects of stellarator geometry on gyrokinetic drift-wave turbulence*. PhD thesis, Princeton University, 2012.