Zonally dominated dynamics and the transition to strong turbulence in ion-scale plasma turbulence

Plamen G. Ivanov
St John's College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy in Theoretical Physics

Trinity 2021

To my parents

Zonally dominated dynamics and the transition to strong turbulence in ion-scale plasma turbulence

Plamen G. Ivanov
St John's College, University of Oxford

Submitted for the degree of Doctor of Philosophy in Theoretical Physics

Trinity 2021

We present a study of the low-transport Dimits state (Dimits *et al.*, 2000) and the transition to a high-transport saturated state in plasma turbulence driven by the ion-temperature-gradient instability. This thesis focuses on a fluid model derived in the cold-ion, long-wavelength asymptotic limit of the ion gyrokinetic equation in a magnetic field with constant curvature (a Z-pinch) and in the presence of an equilibrium temperature gradient. Numerical simulations reveal that the Dimits state is dominated by a quasi-static staircase-like structure of the temperature gradient intertwined with zonal flows which have patch-wise constant shear. Such a structure is reminiscent of the so-called "$\boldsymbol{E} \times \boldsymbol{B}$ staircase" observed in global gyrokinetic numerical simulations (Dif-Pradalier *et al.*, 2010). It suppresses turbulence in two complementary ways: first, by shearing turbulent eddies in the regions of strong zonal shear, and, secondly, by flattening the background temperature gradient at the turning points of the zonal flow, where the shear vanishes. The turbulent heat flux in the low-collisionality, near-marginal state is dominated by turbulent bursts, triggered by coherent long-lived structures closely resembling those found in gyrokinetic simulations with imposed equilibrium flow shear (van Wyk *et al.*, 2016). The breakup of the low-transport Dimits regime is linked to a competition between the two different sources of poloidal momentum in the system — the Reynolds stress and the advection of the diamagnetic flow by the $\boldsymbol{E} \times \boldsymbol{B}$ flow. The former acts to support the staircase by providing a net negative turbulent viscosity for the zonal flows and is opposed by the latter. The winner of this competition decides the type of saturated state. When the Reynolds stress dominates, the system enters the Dimits regime which is characterised by the aforementioned zonal staircase. Otherwise, if the diamagnetic stress prevails, strong turbulence-suppressing zonal flows cannot be maintained and turbulence reigns supreme. We show that the

transition from low to high transport can be understood by analysing the linearly unstable ion-temperature-gradient modes. This is demonstrated by a semi-analytic model for the Dimits threshold in 2D and at large collisionality. In 3D, unless the system is restricted in the magnetic-field direction, a Dimits state arises for all values of the equilibrium parameters. This is explained by the existence of a “parasitic” small-scale slab-ITG instability which is driven by the gradients of large-scale 2D perturbation. The modes of this parasitic instability provide an effective thermal diffusion at large scales and act to move energy from large scales to small viscous scales where dissipation takes place, thus providing a mechanism for saturation. Although such a saturation mechanism was investigated as early as (Cowley *et al.*, 1991), it is not part of the conventional discourse on strong ITG turbulence, which often follows simpler scenarios (e.g., critical balance, see Barnes *et al.*, 2011).

Acknowledgements

First and foremost, I would like to thank my parents for everything they have done for me throughout my life. I fully appreciate that I would not be the person I am today without their dedication. They were there to answer my constant barrage of questions about life, the Universe, and everything when I was a kid. They decided to move out of my home town in order for me to study in one of the best high schools in the country. It was with their support that I was able to leave Bulgaria nine years ago to pursue my dreams. For all they have done and continue to do for me, I am forever indebted.

One's life is a chaotic story of seemingly random events where every perturbation, no matter how insignificant it may seem at the time, can make a great difference down the timeline. Nevertheless, some events do stand out. I would like to thank my high-school classmates and maths brothers-in-arms Ivan Mechkov, Maria Gorinova, and Alex Petkov, together with whom we fought in many maths competitions (and drank even more beers). I am additionally grateful to Maria for convincing me to apply to the University of Cambridge in spite of my reluctance. That truly was a life-changing decision.

Despite nearly failing STEP II and missing the conditions of my offer, I was accepted by St John's College, Cambridge. Several years would pass before I learnt that there is a single person whom I have to thank for the opportunity to study at Cambridge. With the argument that I would be the perfect supervision partner for Luka Skorić (who, unlike me, met his offer), Matthias Dörrzapf, the Director of Studies in Applied Mathematics,

persuaded the college to take me. It goes without saying that I am deeply grateful to Matthias for this.

As it turned out, Luka and I really did click. I am often nostalgic for the time when we spent hours at the blackboard challenging each other with maths and physics. I would also like to thank Namir Asmar for all the fantastic conversations which, to the frustration of our neighbours, entertained us well beyond midnight.

When I started my DPhil back in 2016, I knew absolutely nothing about plasma physics. But the excellent courses by Alex Schekochihin, Felix Parra Diaz, and Paul Dellar, combined with the enthusiastic and motivational environment of the Oxford Plasma Theory Group quickly got me up to speed. We had a lot of fun and many useful discussions on maths, physics, and whatever else exists in the Universe with the members of the group. I cannot imagine a better environment to study, research, and discuss. In particular, I would like to thank Jason Parisi, Yohei Kawazura, Denis St-Onge, Michael Hardman, Michael Barnes, Alessandro Geraldini, and Toby Adkins for both answering and raising interesting plasma-physics questions.

Special thanks go to my housemates Alex Hall-Chen, Valerian Hall-Chen, and Javier Maurino. Our evenings spent mostly on politics, history, cooking, and wine tasting were definitely a highlight of my non-academic time in Oxford. I hope that our paths will cross again some time in the future.

Outside of Oxford, I would like to thank Bill Dorland for all the good ideas and guidance that he is always so enthusiastic to provide. I would also like to thank Ilya Dodin and Hongxuan Zhu for the insightful discussions and their fantastic publications. My co-supervisor Anthony Field has always been quick to point out how and why my theoretical ramblings were or weren't relevant to the real world, and I thank him for all the experimental knowledge that he has provided.

More recently, I would like to thank Zlatina Dimitrova for being the best pandemic partner one can ask for. Her invaluable and never-ending support and her fantastic meals have eased my struggles for more than a year now. She was always there for me, encouraging me to work when I had to, and persuading me to rest when I needed to. I could not ask for more. Together with Nasko Stamenov, she inspired me to take part in the FameLab competition, which was both enjoyable and extremely useful for developing my skills in public speaking and science communication. In addition, I thank Nasko for being my largest reservoir of general knowledge and quirky facts.

Last, but not least, I would not be writing this were it not for my supervisor, Alex Schekochihin. When my research was stuck and I was banging my head against the wall, he encouraged me to keep calm and carry on. Eventually, the wall yielded. He has always managed to provide a balanced mix of enthusiasm, criticism, and insightfulness; his energy is both inspiring and motivating. Anyone who has ever attended one of our group meetings knows exactly what I mean. Outside of research, he has always been a great conversation partner regardless of the topic. I feel lucky and privileged to have had the opportunity to work with him.

Contents

List of Figures

Chapter 1

Introduction

1.1 Fusion energy

With hydrogen fuel that is ubiquitous on Earth (and beyond) and no dangerous waste, fusion's promise to satiate humanity's ever-growing energy demands may sound too good to be true. And it certainly would have been too good to be true were it not for the exotic conditions that it requires. The Coulomb repulsion between like electric charges sets up a kind of barrier that prevents the positively charged nuclei from getting close enough to fuse unless their kinetic energy is appropriately huge. This implies that nuclear fusion becomes a viable source of energy only when the fuel is heated up to well above one hundred million degrees Celsius. It is unlikely that we will ever discover a material that can withstand such temperatures. Fortunately, we do not have to. When it is this hot, hydrogen ionises and turns into a mixture of ions and electrons, i.e., a plasma — and plasma can be manipulated by magnetic and electric fields. This is precisely the idea behind the tokamak device: trap hydrogen plasma in a magnetic cage, then heat it enough for fusion to take place. If we managed to contain the energy released by the nuclear reactions, the plasma would continuously heat itself and thus support the ongoing reactions — we would have "ignited" our fusion device. Having said all this, the absence of such a reactor down the road should be indicative of the complications and challenges that magnetic confinement fusion faces.[1] One of the most serious ones is trapping the heat inside the plasma.

---

[1] For a recent review, see Ongena *et al.* (2016).

1.2 Turbulence and heat transport

Due to the high plasma temperature required by any viable reactor design, it is clear that an in-depth understanding of the heat-transport properties of magnetically confined plasmas is crucial. Sadly for the fusion programme, tokamak transport is a result of turbulence — a chaotic and unpredictable mess of fluctuations (Hugill, 1983; Liewer, 1985; Wootton *et al.*, 1990). Turbulence is notoriously difficult to work with, both analytically and numerically. In fact, even with modern computational capabilities, we are still unable to simulate plasma turbulence throughout the volume of a fusion device. We are thus forced to adopt some simplifications when dealing with turbulent transport.

Since the characteristic correlation length of the turbulence is small compared to the size of the tokamak, one normally assumes that the local heat transport depends only on local conditions, such as density, temperature, magnetic field, and their gradients. This greatly simplifies the problem of turbulent transport by allowing us to separate the turbulent scale (typically, ion or electron gyroradius) from the scale of equilibrium variation (proportional to the size of the fusion device). This "local" approach has been challenged by the so-called "global" one where the variation of the plasma equilibrium is not ignored. Dif-Pradalier *et al.* (2010) showed that in global simulations of turbulence, the local heat transport depends on the local temperature gradient over an apparent "mesoscale" that sits between the turbulent and equilibrium scales. However, it is yet unclear why such a mesoscale should exist and what sets it. In fact, we shall see that our local simulations can produce structures of size tens, if not hundreds of times the turbulent scale (see §3.2.1.3). The evidence for such large structures further complicates the search for mesoscale evidence. So, despite the existence of some objections to this, the results in this thesis are based on the local approach.

Existing research suggests that the dominant contribution to the heat flux in tokamaks arises from turbulence driven by microinstabilities, the most prominent of which is the ion-temperature-gradient instability (Rudakov & Sagdeev, 1961; Coppi *et al.*, 1967; Pogutse, 1968; Waltz, 1988; Cowley *et al.*, 1991; Kotschenreuther *et al.*, 1995*a*). This instability is the topic of this thesis. We use "ITG" and "ITG turbulence" as shorthand terms for this instability and the turbulence driven by it, respectively. As the name suggests, it is controlled by the gradient of ion temperature, which is a source of free energy for unstable microscale perturbations. It is then natural to investigate the dependence of the heat flux carried by the perturbations on the temperature gradient that drives those perturbations. Knowing the relationship between them, one can invert this relationship and find the heating power needed to support a given temperature gradient.

1.3 The low-transport Dimits state and zonal flows

Strongly driven ITG turbulence, i.e., ITG turbulence with a temperature gradient far above the linear-instability threshold, is believed to saturate via a "critically balanced" turbulent cascade (Barnes *et al.*, 2011): free energy stored in the equilibrium gradient is injected into perturbations by the instability and nonlinearly transferred (cascaded) to smaller scales, where it is thermalised via collisions. This is governed by the same kind of processes as the Kolmogorov cascade in hydrodynamic turbulence (Frisch, 1995). This strongly turbulent saturated state supports vigorous turbulent transport of energy, so increasing the temperature gradient in such a system requires very substantial increases in heating power.

Naïvely, one expects strong turbulence and high levels of transport to set in as soon as the temperature gradient exceeds the linear-instability threshold. However, there is numerical evidence for a low-transport regime with low levels of turbulence at temperature gradients larger than the linear threshold for the ITG instability but smaller than some nonlinear threshold above which strong turbulence and a high-transport state set in (Dimits *et al.*, 2000). Simulations have shown that the low-transport state below this threshold (to which we refer as the "Dimits state" and "Dimits threshold", respectively) is dominated by strong zonal flows (ZFs) — Larmor-scale shear flows in the poloidal direction. In the Dimits state, the ZFs regulate turbulence by shearing the heat-carrying turbulent eddies: the radial wavenumber of a sheared eddie is increased, which leads to an increase in damping and a reduction of eddie amplitude. Our goal shall be to explain how the Dimits state is created and maintained, as well as the processes that lead to its collapse.

Despite being fairly well studied, many aspects of ZF physics, e.g., generation of ZFs from turbulence, stability of zonal fields, dependence of experimentally important quantities, like the heat flux, on basic plasma parameters (density, temperature, magnetic field and their gradients) in zonally dominated plasmas, remain far from being settled. One of the established paradigms is the primary-secondary-tertiary instability scenario (Rogers *et al.*, 2000; Rogers & Dorland, 2005). Let us outline it here. The primary ITG instability feeds energy into a spectrum of linearly unstable modes that become nonlinearly unstable to zonal perturbations: this is the "secondary instability". Saturation is reached when the energy injection into ZFs is balanced by their slow viscous damping. Increasing the temperature gradient increases the primary drive, hence the secondary drive, hence the amplitude of ZFs. However, ZFs of large enough amplitude become nonlinearly unstable to a "tertiary instability", so they break up, transferring energy back into the ITG modes. The suppression due to zonal shear having been lost, fully developed turbulence ensues. In this scenario, the Dimits threshold is given by the threshold of the tertiary instability.

One of the important ideas of this thesis is that one may have to rethink the role of the tertiary instability for the Dimits regime. Even though we show that it determines important properties of the saturated state (e.g., poloidal spectra), it turns out that the Dimits threshold is not directly determined by the tertiary instability itself. The latter only works to excite turbulent perturbations that coexist with the ZFs. The way these perturbations interact with the ZFs via a mechanism akin to a generalised nonlinear secondary instability is what determines the Dimits threshold. In the low-collisionality regime, the interactions between turbulent perturbations and ZFs give rise to predator-prey-like oscillations familiar from past studies of ZF physics (see, e.g., Diamond *et al.*, 2005; Ricci *et al.*, 2006; Kobayashi & Rogers, 2012).

Recent progress has suggested that an entirely different scenario might need to be developed for turbulence with imposed background flow shear, applicable to tokamak plasmas made to rotate differentially. The work by van Wyk *et al.* (2017) has shown that close to marginality (i.e., near the threshold for the instability drive below which the system cannot sustain a turbulent state), the effect of the self-generated zonal shear is negligible compared to the equilibrium flow shear. The heat flux in this near-marginal state is dominated not by space-filling turbulence, but by localised, long-time-coherent, finite temperature and density perturbations travelling through the plasma (van Wyk *et al.*, 2016). We call these structures “ferdinons”, after Ferdinand van Wyk’s name. As the temperature-gradient drive is increased, the number of ferdinons increases, they begin overlapping and interacting strongly, and the system enters a fully developed turbulent state.

Here we do not investigate the case of imposed background flow shear, but we do find that locally generated zonal flows arrange themselves in regions of nearly constant shear. Structures closely resembling ferdinons and produced by a localised tertiary instability at the ZF maxima are seen drifting through these sheared regions. This suggests that the formation of localised structures is a robust feature of sheared ITG turbulence as they are seen both in our simplified model and in more realistic GK simulations. We shall see that it is precisely sheared turbulence and, in particular, its momentum-transport properties that determine the fate of the ZFs generated by the secondary instability.

1.4 Modelling ITG turbulence

In this thesis, we undertake a comprehensive study of the ZF-dominated Dimits state: how it is established, under what conditions it exists, and how it transitions into the strongly turbulent regime. A truly complete treatment of this problem requires the gyrokinetic (GK) framework in toroidal tokamak geometry (Frieman & Chen, 1982; Sugama *et al.*, 1996; Sugama & Horton, 1997, 1998; Abel *et al.*, 2013; Catto, 2019). However, its complexity makes it both analytically and numerically hard to treat. This is why we focus on the more realistic task of tackling the problem in a simplified model for the dynamical evolution of the perturbations of electrostatic potential (or, equally well, density), ion temperature and ion parallel flow in a tokamak plasma. The model is derived as an exact asymptotic limit of the underlying gyrokinetic equations in a physically realisable, if not necessarily most general, regime. The approximations used are chosen to ensure that our model has a number of features that we consider essential: 1) a curvature-driven ion-temperature-gradient (”curvature-ITG”) instability, characteristic of tokamak plasmas; 2) a “slab-ITG” instability driven by the parallel derivatives of pressure and flow; 3) an appropriate modified adiabatic electron response, which has been found to be crucial for capturing essential zonal-flow properties (e.g., the correct ITG secondary instability: see Hammett *et al.*, 1993; Rogers *et al.*, 2000). The resulting equations are a three-field fluid model linking the perturbations of the electrostatic potential, ion temperature, and ion parallel flow rather than a one-field drift-wave model of the Hasegawa & Mima (1978) variety. Such a model allows us to capture the two linear ITG instabilities driven by the gradient of the equilibrium ion temperature, viz., slab ITG and curvature-driven ITG, while keeping the equations simple enough to allow for an analytic treatment.

1.5 Turbulent saturation

Broadly speaking, there are two ways for turbulence to achieve saturation — it can either cascade injected energy down to dissipation scales or, if it is internally driven by an instability, it can assemble itself in a configuration that suppresses that instability, i.e., the initial unstable equilibrium evolves into a new equilibrium with weaker instabilities. As our model contains both ZF and zonal temperature perturbations, in principle it can accommodate the physics of two possible instability-suppression mechanisms: shearing of the turbulence by ZFs and modifying the background temperature gradient by zonal temperature perturbations in order to cancel the ITG drive. Neither of these can be done uniformly across the entire domain because we impose periodic boundary conditions in the radial direction. Interestingly, we find that in the Dimits regime, the zonal perturbations arrange themselves in alternating wide regions of nearly-constant zonal shear, strong enough to suppress turbulence, and narrow regions of strong zonal temperature gradient, which flattens the background temperature gradient and quenches the ITG instability. A “staircase”-like radial temperature profile emerges. This is reminiscent of the zonal structures seen in some global and local flux-driven gyrokinetic simulations (Dif-Pradalier *et al.*, 2010, 2017; Villard *et al.*, 2013;

Rath *et al.*, 2016) and reported in experimental data (Dif-Pradalier *et al.*, 2015). The resulting turbulent heat flux is significantly suppressed. The zonal staircase slowly decays due to collisional viscosity, and this viscous decay results in recurrent turbulent bursts that are triggered by localised travelling structures emerging from the ZF maxima, where they are created by a local tertiary instability of the ZF profile. The turbulence that develops during a burst is sheared by the ZFs. Locally, the zonal shear breaks a fundamental symmetry of GK turbulence (Parra *et al.*, 2011) and gives rise to radial flux of poloidal momentum whose sign is controlled by the sign of the zonal shear. This momentum flux is made up of two parts — the usual Reynolds stress of the $\boldsymbol{E} \times \boldsymbol{B}$ flow that is known to generate strong ZFs, and a diamagnetic contribution that opposes the Reynolds stress. The distinguishing feature of the Dimits regime is that the Reynolds stress overcomes the diamagnetic one. The zonal staircase is stable to turbulent bursts because ZF-sheared turbulence provides an effective negative viscosity for the ZFs.

The Dimits transition to higher turbulent transport occurs when the diamagnetic stress wins over the Reynolds one, so the effective turbulent viscosity flips its sign and becomes positive. In this regime, the coherent ZFs that support the Dimits state become nonlinearly unstable. A strongly turbulent state is established where saturation is facilitated by "parasitic" small-scale slab-ITG modes. These modes are driven not by the equilibrium temperature gradient but rather by the gradients of large-scale perturbations, themselves generated by the equilibrium-driven curvature-ITG instability. The parasitic modes extract energy from the large-scale perturbations and transfer it to smaller perpendicular scales whence it cascades to viscous scales and is dissipated. The idea of such parasitic modes is hardly original (Drake *et al.*, 1988; Cowley *et al.*, 1991; Rath & Sridhar, 1992). Here we back their existence both by analytical arguments and by numerical results, and show that they provide positive effective thermal diffusion for the large-scale perturbations.

1.6 Thesis outline

In the following chapter 2, we describe our fluid model for ITG turbulence and explain the choices we made in deriving it. The detailed derivation is contained in appendix A.

We then proceed to investigate the 2D version of the model. Chapter 3 contains a comprehensive study of the 2D Dimits regime, its characteristic zonal staircase, and the mechanism of momentum transport that generates and sustains the ZFs. Based on this, we derive a simple analytic prediction for the threshold of the transition to strong turbulence. Beyond the Dimits state, the 2D model fails to reach a saturated state — perturbations grow exponentially and box-sized radially elongated structures eventually dominate.[2] We argue that this is a fundamental limitation of the unphysical 2D approximation (cf. the inverse cascade in 2D hydrodynamics). This is corroborated by the existence of a well-defined strongly turbulent saturated state in the 3D version of the model.

In chapter 4, we relax the 2D approximation and turn to the full 3D fluid model. We find that the physics of the 3D Dimits regime is virtually the same as in 2D. The qualitative difference between the 3D and 2D models is the existence of small-scale slab-ITG instability in the former. This instability is driven by the gradients associated with large-scale 2D perturbations. We provide both analytical arguments and numerical results on the small-scale slab-ITG modes and show that their influence on the large-scale perturbations is to provide effective thermal diffusion. This allows the 3D system to have a finite-amplitude strongly turbulent saturated state, unlike its 2D counterpart, which suffers from a blow up beyond the Dimits regime.

Finally, chapter 5 contains a detailed summary of our results. We also discuss how this work fits in the context of existing research, and suggest directions for future exploration.

---

[2] This is somewhat similar to the blow up observed in ITG fluid simulations with exact adiabatic response for the electrons (Hammett *et al.*, 1993) (or in electron-scale turbulence, where the ion response is naturally adiabatic). In both cases, the system requires ZFs in order to saturate at finite amplitudes. However, in our system, the inability to generate ZFs is not due to the electron (or ion) response, but is rather due to a change of the properties of momentum transport of sheared turbulence, see §3.3.

Chapter 2

ITG-driven dynamics in a Z-pinch

2.1 Model equations

2.1.1 Magnetic geometry

The magnetic geometry of constant magnetic curvature is chosen because it is the simplest one that enables the curvature-driven ITG instability by coupling the electrostatic potential and the temperature perturbations via the magnetic drift. In this work, we shall ignore any electromagnetic effects. This is justified if the plasma pressure is much smaller than the magnetic pressure (i.e., the plasma beta is small). The integration domain is positioned in the magnetic field of a line of current (Z-pinch)[1] at radial distance L_B from the current line: see figure 2.1. We define the x and y axes as pointing radially outwards and antiparallel to the Z-pinch current, respectively. We assume $L_B \gg L_x, L_y$, where L_x and L_y are the "radial" (x) and "poloidal" (y) sizes of the domain, respectively. Here we use the terms "radial" and "poloidal" to reflect the intended similarity of the domain to one positioned at the outboard midplane in a tokamak geometry. In that sense, we can think of the radial x coordinate as perpendicular to flux surfaces. These surfaces are parametrised by the poloidal y and field-parallel z coordinates. Here $\boldsymbol{B} = B\hat{\boldsymbol{z}}$ is the magnetic field and the unit vectors $\{\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}}\}$ form a right-handed basis. The magnetic field of the Z-pinch with total current I is azimuthal around the current line (as shown in figure 2.1) and has magnitude $B(x) = 2I/cx$. The

---

[1] This simplification as a route to a minimal model of ion-scale turbulence goes back at least to Ricci *et al.* (2006).

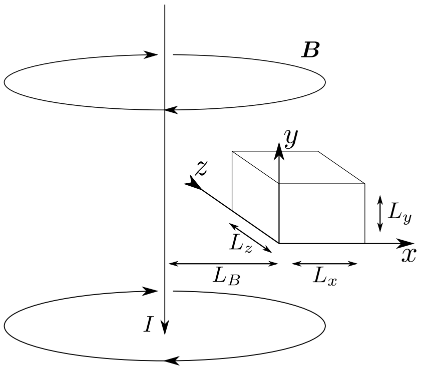


Figure 2.1: Illustration of the Z-pinch magnetic geometry.

radial gradient of this field is then

$$\frac{1}{B}\frac{dB}{dx} = -\frac{1}{L_B}. \tag{2.1}$$

This value is constant across the domain to lowest order in $L_x/L_B \ll 1$. We define L_B to be the magnetic scale length. Similarly, we can define the ITG scale length

$$\frac{1}{L_T} \equiv -\frac{1}{T_i}\frac{dT_i}{dx}. \tag{2.2}$$

In a tokamak, L_B scales with the major radius of the device, while L_T scales with the minor radius. Here we will take the limit

$$L_B \gg L_T, \tag{2.3}$$

equivalent to a large-aspect-ratio approximation in a tokamak geometry. We do this in order to ensure that the magnetic drift in the density equation is of the appropriate order, see (2.9) and its derivation in appendix A.1.3. This drift is essential for the linear curvature-driven ITG instability that we aim to capture.

2.1.2 Electron response

The electron density is assumed to follow the modified adiabatic response

$$\frac{\delta n_e}{n_e} = \frac{e(\phi - \overline{\phi})}{T_e}. \tag{2.4}$$

This is essentially an equation of state for the electrons and can be rigorously derived in the asymptotic limit of small electron mass (Abel & Cowley, 2013). Intuitively, it is a consequence of the electrons' ability to stream quickly along the magnetic field lines, thus "shorting out" electric fields within a flux surface, but not across different flux surfaces. Here δn_e and n_e are the perturbed and equilibrium electron density, respectively, ϕ is the electrostatic potential, T_e is the electron temperature, and

$$\overline{\phi}(x) \equiv \frac{1}{L_y L_\parallel} \int dydz\ \phi(x,y,z) \tag{2.5}$$

is the zonal (flux-surface) spatial average of the perturbed electric potential ϕ, where the parallel size of the domain is $L_\parallel$. We refer to zonally averaged fields as being "zonal". The difference

$$\phi' \equiv \phi - \overline{\phi} \tag{2.6}$$

is the "nonzonal" part of the field. The "standard" adiabatic response would be $\delta n_e \propto \phi$, instead of $\delta n_e \propto \phi'$. The subtraction of the zonal average $\overline{\phi}$ reflects the fact that the electrons can stream only within a flux surface. The correct form of the electron response (2.4) has been found to be crucial for capturing the ZF-dominated state of ITG turbulence (Hammett *et al.*, 1993; Dorland & Hammett, 1993).

A cautious reader has spotted that there is no way to define flux surfaces in the magnetic geometry of a Z-pinch, as the magnetic field lines do not describe 2D surfaces, but rather close on themselves after one turn around the current axis. This problem can be fixed by demanding that the magnetic field be, in fact, sheared: $\boldsymbol{B} = B_0(\hat{\boldsymbol{z}} + x\hat{\boldsymbol{y}}/L_s)$, where L_s is the characteristic scale length of the magnetic shear. This is the magnetic field of a helimak (Gentle & He, 2008). The field lines define cylindrically symmetric concentric flux surfaces and the electron parallel streaming mixes the azimuthal (z) and poloidal (y) directions. We can then take the limit $L_s \to \infty$ *after* performing the small-mass-ratio ($m_e/m_i \to 0$) expansion and eliminate magnetic shear from the ion equations, while retaining the flux-surface effect in the electron response.

2.1.3 Cold-ion limit

The cold-ion limit allows us to simplify the gyroaveraging operator that appears in gyrokinetics. Its corresponding Fourier-space operator is multiplication by the Bessel function

$$J_0\left(\frac{k_\perp v_\perp}{\Omega_i}\right) = 1 - \frac{1}{4} k_\perp^2 \rho_i^2 \frac{v_\perp^2}{v_{ti}^2} + \mathcal{O}\big(k_\perp^4 \rho_i^4\big). \tag{2.7}$$

The square of the ion gyroradius $\rho_i = v_{ti} m_i c/ZeB$ and the ion temperature $T_i = m_i v_{ti}^2/2$ are both proportional to the square of the ion thermal speed v_{ti} (here m_i and Z are the ion mass and charge in units of e, respectively). Thus, the cold-ion limit is equivalent to a long-wavelength expansion $\mathcal{O}(k_\perp \rho_i) \ll 1$ with a finite sound radius $\mathcal{O}(k_\perp \rho_s) \sim 1$, where $k_\perp$ is the perpendicular (to the magnetic field) wavenumber, $\rho_s = \rho_i/\sqrt{2\tau}$ is the sound radius, and $\tau = T_i/ZT_e$ is the temperature ratio. The sound radius ρ_s is the natural normalisation for the microphysical length scales in the problem: see (2.9)–(2.11) below and their derivation in appendix A.1.

2.1.4 Moments of the GK equation

We take the density, temperature, and parallel-velocity moments of the electrostatic ion gyrokinetic equation and adopt the high-collisionality, cold-ion, long-wavelength, large-aspect-ratio ordering

$$\frac{\partial_t}{\nu_i} \sim \tau \sim k_\perp^2 \rho_i^2 \sim \frac{L_T}{L_B} \ll 1 \sim \frac{\varphi}{T} \sim \frac{U_\parallel}{\tau c_s \varphi}, \tag{2.8}$$

where $\varphi \equiv Z_i e\phi/T_i$ is the normalised electric potential, $T = \delta T/T_i$ is the normalised ion-temperature perturbation, $U_\parallel$ is the perturbation of the parallel ion velocity, and ν_i is the ion-ion collision frequency. The equations that we obtain in appendix A.1 are

$$\begin{aligned}
&\frac{\partial}{\partial t}\left(\tau\varphi' - \frac{1}{2}\rho_i^2\nabla_\perp^2\varphi\right) + \frac{\partial U_\parallel}{\partial z} - \frac{\rho_i v_{ti}}{L_B}\frac{\partial}{\partial y}(\varphi + T) + \frac{\rho_i v_{ti}}{2L_T}\frac{\partial}{\partial y}\left(\frac{1}{2}\rho_i^2\nabla_\perp^2\varphi\right) \\
&\quad + \frac{1}{2}\rho_i v_{ti}\left(\left\{\varphi, \tau\varphi' - \frac{1}{2}\rho_i^2\nabla_\perp^2\varphi\right\} + \frac{1}{2}\rho_i^2\boldsymbol{\nabla}_\perp\cdot\{\boldsymbol{\nabla}_\perp\varphi, T\}\right) \\
&\quad = -\frac{1}{2}\chi\rho_i^2\nabla_\perp^4(a\varphi - bT),
\end{aligned} \tag{2.9}$$

$$\frac{\partial T}{\partial t} + \frac{\rho_i v_{ti}}{2L_T}\frac{\partial\varphi}{\partial y} + \frac{1}{2}\rho_i v_{ti}\{\varphi, T\} = \chi\nabla_\perp^2 T, \tag{2.10}$$

$$\frac{\partial U_\parallel}{\partial t} + \frac{v_{ti}^2}{2}\frac{\partial}{\partial z}(\varphi + T) + \frac{1}{2}\rho_i v_{ti}\{\varphi, U_\parallel\} = d\chi\nabla_\perp^2 U_\parallel, \tag{2.11}$$

where the Poisson bracket is defined by

$$\{f, g\} = \hat{\boldsymbol{b}}\cdot(\boldsymbol{\nabla}_\perp f \times \boldsymbol{\nabla}_\perp g) = \frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}, \tag{2.12}$$

and $\boldsymbol{\nabla}_\perp \equiv \hat{\boldsymbol{x}}\partial_x + \hat{\boldsymbol{y}}\partial_y$ denotes the gradient operator in the perpendicular plane. The thermal diffusivity is

$$\chi \equiv \frac{8}{9}\sqrt{\frac{2}{\pi}}\nu_i\rho_i^2 \tag{2.13}$$

The numerical factor in (2.13) and the constants $a = 9/40$, $b = 67/160$, $d = 9/10$ in (2.9) and (2.11) are specific to the Landau collision operator (see appendix A.1.6). These agree with more general calculations of the collisional perpendicular viscosity, gyroviscosity and collisional heat flux (Mikhailovskii & Tsypin, 1971; Catto & Simakov, 2004, 2005).

Let us discuss the physics content of (2.9)–(2.11). Equation (2.9) describes the time evolution of the sum of perturbed ion density and the vorticity of the $\boldsymbol{E} \times \boldsymbol{B}$ drift velocity:

$$\frac{\delta n_i}{n_i} + \frac{1}{n_i} \int d^3\boldsymbol{v} \, \big(\varphi - \langle\langle\varphi\rangle_R\rangle_r\big) F_i = \tau\varphi' - \frac{1}{2}\rho_i^2 \nabla_\perp^2 \varphi + \mathcal{O}\big(k_\perp^4 \rho_i^4 \varphi\big), \tag{2.14}$$

where n_i and δn_i are the equilibrium and perturbed ion densities, respectively (see appendix A.1 for the definitions of the equilibrium distribution function F_i and the gyroaveraging operators $\langle.\rangle_r$ and $\langle.\rangle_R$). Thus, (2.9) can be thought of as both a perturbed-ion-density equation and as the curl of the perpendicular-momentum equation. The equality in (2.14) follows from the quasineutrality condition $Z\delta n_i = \delta n_e$, the electron response (2.4), the approximation (2.7), and the ordering (2.8).

The first term of (2.9) is the time derivative of (2.14). The second term, $\partial_z U_\parallel$, describes the compressions and rarefactions due to the parallel ion flow. The next term, namely, $(\rho_i v_{ti}/L_B)\partial_y(\varphi + T)$, is the magnetic drift (both curvature and ∇B). This, or rather the $\partial_y T$ part of it, is essential for the curvature-driven ITG instability. It appears in the density equation because the magnetic drift creates charge separation, and hence electrostatic potential, which is then coupled to the perturbed density via quasineutrality. The fourth term, $(-\rho_i v_{ti}/4L_T)\partial_y(\rho_i^2\nabla_\perp^2\varphi)$, is a finite-Larmor-radius (FLR) term originating from the gyroaveraged $\boldsymbol{E} \times \boldsymbol{B}$ drift. It is the diamagnetic drift due to the equilibrium temperature gradient. The first of the nonlinear terms represents the advection of the quantity (2.14) by the $\boldsymbol{E}\times\boldsymbol{B}$ drift. The second nonlinear term $\boldsymbol{\nabla}_\perp \cdot \{\boldsymbol{\nabla}_\perp\varphi, T\}$ is another FLR effect, which describes the $\boldsymbol{E} \times \boldsymbol{B}$ advection of diamagnetic momentum. This term provides a crucial source of turbulent poloidal momentum flux that destabilises the ZF profiles, destroying the ZF-dominated Dimits regime. The nature of this term and its role in the Dimits transition are discussed in detail in §3.3.3. Note that the nonlinear terms in (2.9) and (2.10) are equivalent to those in the model analysed by Rogers *et al.* (2000) in the cold-ion limit (2.8). Finally, the collisional terms in (2.9) represent the viscous damping of the $\boldsymbol{E} \times \boldsymbol{B}$ flow and also couple the density and temperature perturbations. The latter coupling does not appear to be important for the results of this thesis, but has been kept for consistency.

The temperature equation (2.10) describes the advection of the total temperature (perturbations plus equilibrium) by the $\boldsymbol{E} \times \boldsymbol{B}$ drift $\boldsymbol{V}_E = c\hat{\boldsymbol{b}} \times \boldsymbol{\nabla}_\perp \phi / B$, and the thermal diffusion perpendicular to the magnetic field. Indeed, (2.10) can be rewritten as

$$\frac{d}{dt}(\delta T + T_i) = \chi \nabla_\perp^2 \delta T, \tag{2.15}$$

where the advective time derivative is

$$\frac{d}{dt} \equiv \frac{\partial}{\partial t} + \boldsymbol{V}_E \cdot \boldsymbol{\nabla}_\perp. \tag{2.16}$$

The advection of the equilbrium temperature, $\boldsymbol{V}_E \cdot \boldsymbol{\nabla}_\perp T_i$, which is the second term on the left-hand side of (2.10), is responsible for the injection of free energy (see §2.2), causing the ITG instability.

Similarly, (2.11) has a straightforward interpretation — the parallel flow is driven by the parallel gradient of the pressure $p = \varphi + T$, advected by the $\boldsymbol{E} \times \boldsymbol{B}$ flow, and damped by the collisional viscosity $d\chi$.

To prepare (2.9)–(2.11) for numerical analysis and distil important parameters, we introduce normalised variables and fields

$$\begin{gathered} \hat{t} \equiv \frac{2\rho_s \Omega_i}{L_B} t, \qquad \hat{x} \equiv \frac{x}{\rho_s}, \qquad \hat{y} \equiv \frac{y}{\rho_s}, \qquad \hat{z} \equiv \frac{2z}{L_B} \\ \hat{\varphi} \equiv \frac{\tau L_B \varphi}{2\rho_s} = \frac{\tau L_B}{2\rho_s} \frac{Z_i e \phi}{T_i}, \qquad \hat{T} \equiv \frac{\tau L_B T}{2\rho_s} = \frac{\tau L_B}{2\rho_s} \frac{\delta T}{T_i}, \qquad \hat{u} \equiv \frac{U_\parallel}{\rho_s \Omega_i \tau} \\ \hat{\kappa}_T \equiv \frac{\tau L_B}{2 L_T}, \qquad \hat{\chi} \equiv \frac{L_B}{2\rho_s} \frac{\chi}{\Omega_i \rho_s^2}. \end{gathered} \tag{2.17}$$

All hatted quantities are ordered as $\mathcal{O}(1)$. Dropping hats, we obtain from (2.9)–(2.11) the following equations in normalised units:

$$\begin{aligned}\partial_t\big(\varphi' - \nabla_\perp^2\varphi\big) + \partial_\parallel u - \partial_y(\varphi + T) + \kappa_T\partial_y\nabla_\perp^2\varphi \\ +\big\{\varphi, \varphi' - \nabla_\perp^2\varphi\big\} + \boldsymbol{\nabla}_\perp\cdot\{\boldsymbol{\nabla}_\perp\varphi, T\} = -\chi\nabla_\perp^4(a\varphi - bT),\end{aligned} \tag{2.18}$$

$$\partial_t T + \kappa_T\partial_y\varphi + \{\varphi, T\} = \chi\nabla_\perp^2 T, \tag{2.19}$$

$$\partial_t u + \partial_\parallel(\varphi + T) + \{\varphi, u\} = d\chi\nabla_\perp^2 u, \tag{2.20}$$

where we use $\partial_\parallel \equiv \partial_z$. These equations have two independent parameters: the normalised equilibrium temperature gradient, κ_T, and the normalised collisionality, χ.[2] There are three other parameters — L_x, L_y, and $L_\parallel$ that are the domain sizes in x, y, and z. We expect that any physically relevant results must be independent of the perpendicular sizes L_x and L_y if our equations are indeed a valid local model of the plasma. On the other hand, $L_\parallel$ can be thought of as analogous to the connection length in tokamak geometry and so we expect the 3D saturation to depend on it (Barnes *et al.*, 2011). In chapter 4, we shall see that small $L_\parallel$ can effectively push the 3D system into the 2D regime, while for large enough values of $L_\parallel$, the saturated state becomes independent of the parallel size of the box.

We shall at first focus heavily on the simplified 2D version of (2.18)–(2.20) that is obtained by setting $\partial_\parallel = 0$. In this case, the dynamics of φ and T become independent of u. In fact, the latter necessarily decays exponentially to zero: multiplying (2.20) by u and integrating over space yields

$$\partial_t \int d^3\boldsymbol{r}\ \frac{1}{2}u^2 = -d\chi\int d^3\boldsymbol{r}\ (\boldsymbol{\nabla}_\perp u)^2 \le 0, \tag{2.21}$$

where equality is satisfied only if $u = 0$. Therefore, the 2D limit of (2.18)–(2.20) is

$$\begin{aligned}\partial_t\big(\varphi' - \nabla_\perp^2\varphi\big) - \partial_y(\varphi + T) + \kappa_T\partial_y\nabla_\perp^2\varphi \\ +\big\{\varphi, \varphi' - \nabla_\perp^2\varphi\big\} + \boldsymbol{\nabla}_\perp\cdot\{\boldsymbol{\nabla}_\perp\varphi, T\} = -\chi\nabla_\perp^4(a\varphi - bT),\end{aligned} \tag{2.22}$$

$$\partial_t T + \kappa_T\partial_y\varphi + \{\varphi, T\} = \chi\nabla_\perp^2 T, \tag{2.23}$$

where the fields φ and T depend only on x and y. Chapter 3 shall concentrate entirely on the nonlinear saturation of (2.22) and (2.23).

---

[2] The reader might wonder what the experimentally relevant values of χ are. Using the data from Abel & Cowley (2013), we find $\chi \approx 6\times10^{-4}$ for a deuterium plasma in JET. Thus, the low-collisionality regime is the one we expect to be of greater interest.

2.1.5 Relationship to Hasegawa-Mima equation and related models

It is easy to see that setting $\kappa_T = 0$ in the 2D equations (2.22) and (2.23) effectively decouples φ and T — taking the initial condition $T(t = 0) = 0$ leads to the trivial solution $T(t) = 0$. In that case, (2.22) reduces to

$$\partial_t\big(\varphi' - \nabla_\perp^2\varphi\big) - \partial_y\varphi + \big\{\varphi, \varphi' - \nabla_\perp^2\varphi\big\} = -a\chi\nabla^4\varphi, \tag{2.24}$$

which is the well-known (modified) Charney-Hasegawa-Mima equation (Hasegawa & Mima, 1978) that includes the appropriate modified adiabatic electron response, with viscous damping. Even though we have considered a situation with no equilibrium density gradient, the magnetic drift provides a term identical to the one that would have arisen from the $\boldsymbol{E}\times\boldsymbol{B}$ advection of an inhomogeneous equilibrium density profile. This puts the model considered here in the same class of systems as those proposed by Hasegawa & Wakatani (1983), Terry & Horton (1983) and others — all essentially extensions of the Hasegawa-Mima (HM) equation with additional physics to account for microinstabilities in the plasma.[3] Our model and the HM equation share the same φ–φ nonlinear terms in (2.18). These terms are responsible for the secondary instability and generation of strong ZFs in both systems (see §3.1), as well as the Kelvin-Helmholtz tertiary instability of the ZFs (see appendix B.2). However, the HM model does not capture the effects of temperature perturbations, which are expected to play a crucial role in the saturated state of tokamak turbulence since the latter is believed to be driven predominantly by temperature-gradient instabilities. Indeed, in §3.3, we shall find that both in 2D and 3D, the break up of the Dimits state of (2.18)–(2.20) is, in an essential way, governed by the behaviour of the temperature perturbations, which are absent from (2.24).

---

[3] There has recently been a significant effort to advance the understanding of (2.24) and its relatives (Parker & Krommes, 2013, 2014; Parker, 2016; Ruiz *et al.*, 2016, 2019; Zhu *et al.*, 2018*b*, 2019, 2020*a*,*b*; Zhou *et al.*, 2019; Plunk & Bañón Navarro, 2017; St-Onge, 2017; Majda *et al.*, 2018; Qi *et al.*, 2019).

2.2 Conservation laws

2.2.1 Conservation in 2D

The 2D equations (2.22) and (2.23) have three conservations laws describing the time evolution of quantities that would be conserved in the absence of equilibrium gradients and dissipation:

$$\partial_t \int dxdy\ \frac{1}{2}T^2 = -\kappa_T \int dxdy\ T\partial_y\varphi - \chi \int dxdy\ (\boldsymbol{\nabla}_\perp T)^2, \tag{2.25}$$

$$\begin{aligned}
&\partial_t \int dxdy\ \frac{1}{2}\left[\varphi'^2 + (\boldsymbol{\nabla}_\perp\varphi)^2\right] \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (2.26)\\
&= -\int dxdy\ T\partial_y\varphi - \chi \int dxdy\ (\nabla^2\varphi)\big(a\nabla^2\varphi - b\nabla^2 T\big),\\
&\partial_t \int dxdy\ \left[\frac{1}{2}\varphi'^2 + T\varphi' + \frac{1}{2}(\boldsymbol{\nabla}_\perp T + \boldsymbol{\nabla}_\perp\varphi)^2\right]\\
&= -\chi \int dxdy\ \bigg[(\boldsymbol{\nabla}_\perp\varphi')\cdot(\boldsymbol{\nabla}_\perp T) + a\big(\nabla^2\varphi\big)^2\\
&\quad + (a+1-b)\big(\nabla^2\varphi\big)\big(\nabla^2 T\big) + (1-b)\big(\nabla^2 T\big)^2\bigg].
\end{aligned} \tag{2.27}$$

These conservation laws can be deduced directly from (2.22) and (2.23): e.g., (2.25) is obtained by multiplying (2.23) by T and integrating over x and y. They are also particular cases of the conservation laws of the gyrokinetic equation. The conservation of the variance of T, given by (2.25), is the lowest-order version of the gyrokinetic free-energy budget. The other two conservation laws, (2.26) and (2.27), can be derived from the conservation of the two-dimensional gyrokinetic invariant (see Schekochihin *et al.*, 2009; Plunk *et al.*, 2010). This invariant is a function of velocity in the GK formalism. The 2D equations (2.22) and (2.23) are based only on two velocity moments of the distribution function, namely density and temperature, and so the two-dimensional invariant yields two independent conservation laws. More specifically, (2.26) is a generalisation of the "electrostatic gyrokinetic invariant". The derivations of the three invariants of our system directly from the corresponding GK invariants can be found in appendix A.2.

Physically, (2.26) can be thought of as conservation of the kinetic energy of $\boldsymbol{E}\times\boldsymbol{B}$ flows. The third conservation law (2.27) has some peculiar properties. First, neither the conserved quantity on the left-hand side nor the dissipation rate on the right-hand side is sign-definite. Secondly, all of the evolution is dissipative, i.e., this invariant is not injected by any equilibrium gradients and is constant in time if $\chi = 0$. We are not aware of a physical interpretation of this conservation law.

2.2.2 Conservation in 3D

The conservation law of free energy for the 3D equations (2.18)–(2.20) is equivalent (modulo the integration domain) to the 2D one (2.25). It reads

$$\partial_t W \equiv \partial_t \int d^3\boldsymbol{r}\ \frac{1}{2}T^2 = -\kappa_T \int d^3\boldsymbol{r}\ T\partial_y\varphi - \chi \int d^3\boldsymbol{r}\ (\boldsymbol{\nabla}_\perp T)^2. \tag{2.28}$$

The first term on the right-hand side of (2.28) is proportional to the nondimensional heat flux[4]

$$Q = -\frac{1}{L_x L_y L_\parallel} \int d^3\boldsymbol{r}\ T\partial_y\varphi, \tag{2.29}$$

whereas the second one is the collisional thermalisation. Equation (2.28) implies that a steady saturated state, i.e., $\partial_t W = 0$, can be achieved only if appropriate balance between injection and dissipation terms is established:

$$Q = \frac{\chi}{\kappa_T} \int \frac{dxdy}{L_x L_y}\ (\boldsymbol{\nabla}_\perp T)^2. \tag{2.30}$$

Thus, a saturated state would necessarily have a net positive turbulent (or “anomalous”) heat flux $Q > 0$. Note that the first term on the right-hand side of (2.28), which represents injection of free energy, is $\kappa_T Q$. The turbulent heat flux is enabled by the turbulence excited by the ITG instability.

Surprisingly, going from 2D to 3D does not eliminate both of the other two 2D invariants. One of them survives, and the following conservation law holds even in 3D:

$$\begin{aligned}\partial_t I \equiv \partial_t \int d^3\boldsymbol{r}\ &\left[\frac{1}{2}(\varphi' + T')^2 + \frac{1}{2}\overline{T}^2 + \frac{1}{2}(\boldsymbol{\nabla}_\perp T + \boldsymbol{\nabla}_\perp \varphi)^2 + \frac{1}{2}u^2\right] \\ = -\kappa_T \int d^3\boldsymbol{r}\ T\partial_y\varphi - \chi \int d^3\boldsymbol{r}\ &\Big[(\boldsymbol{\nabla}_\perp\varphi')\cdot(\boldsymbol{\nabla}_\perp T) + (\boldsymbol{\nabla}_\perp T)^2 + \\ &a\big(\nabla_\perp^2\varphi\big)^2 + (a+1-b)\big(\nabla_\perp^2\varphi\big)\big(\nabla_\perp^2 T\big) + (1-b)\big(\nabla_\perp^2 T\big)^2 + d(\boldsymbol{\nabla}_\perp u)^2\Big].\end{aligned} \tag{2.31}$$

---

[4] The dimensional ion heat flux $Q_i = V^{-1}\int d^3\boldsymbol{r} \int d^3\boldsymbol{v}(\boldsymbol{V}_E \cdot \hat{\boldsymbol{x}})(m_i v^2/2)\delta f_i$ (Barnes *et al.*, 2011), where V is the volume of integration and δf_i is the perturbed ion distribution function, see (A.1) in appendix A.1.1, is related to Q via $Q_i/Q = 3n_i T_i v_{ti}(\rho_i/L_B)^2/\tau^{5/2}\sqrt{2}$.

As expected, one recovers a corresponding 2D conservation law by setting $u = 0$ and excluding z from the integration. In particular, (2.31) is then the sum of (2.26) and (2.27). Again, we are not aware of the physical meaning of (2.31).

In chapter 4, it will prove useful to discuss the spectra of W and I. For this, we write $W = \sum_{\boldsymbol{k}} W_{\boldsymbol{k}}$ and $I = \sum_{\boldsymbol{k}} I_{\boldsymbol{k}}$, where we have defined

$$W_{\boldsymbol{k}} \equiv \frac{1}{2}|T_{\boldsymbol{k}}|^2, \tag{2.32}$$

$$I_{\boldsymbol{k}} \equiv \frac{1}{2}\big(|\varphi'_{\boldsymbol{k}} + T'_{\boldsymbol{k}}|^2 + |\overline{T}_{k_x}|^2 + k_\perp^2|\varphi_{\boldsymbol{k}} + T_{\boldsymbol{k}}|^2 + |u_{\boldsymbol{k}}|^2\big). \tag{2.33}$$

2.3 Linear physics of the ITG instability

Equations (2.18)–(2.20) support two distinct types of linear instability: the curvature-driven ITG and the slab-ITG instabilities. The former is essentially 2D and is driven by the magnetic drift $-\partial_y(\varphi + T)$ in (2.18) and the $\boldsymbol{E} \times \boldsymbol{B}$ advection of equilibrium temperature $\kappa_T \partial_y \varphi$ in (2.19). The latter instability is 3D and relies on the parallel compression $\partial_\parallel u$ in (2.18), the $\boldsymbol{E} \times \boldsymbol{B}$ advection of equilibrium temperature $\kappa_T \partial_y \varphi$ in (2.19), and the parallel pressure gradient $\partial_\parallel(\varphi + T)$ in (2.20). In order to investigate the linear instability of (2.18)–(2.20), we drop the nonlinear terms and look for Fourier modes $\varphi, T, u \propto \exp(-i\omega_{\boldsymbol{k}} t + i\boldsymbol{k} \cdot \boldsymbol{r})$, where $\mathrm{Re}(\omega_{\boldsymbol{k}})$, $\mathrm{Im}(\omega_{\boldsymbol{k}})$, and $\boldsymbol{k} = (k_x, k_y, k_\parallel)$ are the real frequency, growth rate, and wavenumber of the mode, respectively.[5] We also define the perpendicular wavenumber $k_\perp \equiv \sqrt{k_x^2 + k_y^2}$. The dispersion relation can be written as

$$(-i\omega_{\boldsymbol{k}} + dk_\perp^2)D_{2D} + \frac{k_\parallel^2}{1 + k_\perp^2}\big(-i\omega_{\boldsymbol{k}} + \chi k_\perp^2 - i\kappa_T k_y\big) = 0, \tag{2.34}$$

where the 2D dispersion relation is given by

$$D_{2D} = (-i\omega_{\boldsymbol{k}} + A)(-i\omega_{\boldsymbol{k}} + B - iC) - fAB + igAC = 0, \tag{2.35}$$

$$A = \chi k_\perp^2, \quad B = \frac{a k_\perp^4}{1 + k_\perp^2}, \quad C = k_y \frac{1 + \kappa_T k_\perp^2}{1 + k_\perp^2}, \quad f = \frac{\kappa_T k_y^2}{a^2 k_\perp^6}, \quad g = \frac{b \kappa_T k_\perp^2}{1 + \kappa_T k_\perp^2}. \tag{2.36}$$

We first discuss the curvature-driven ITG modes, viz., the solutions to (2.35), before turning our attention to the 3D slab-ITG linear instability.

---

[5] We remind the reader that perpendicular distances (hence k_x and k_y) are normalised to the sound radius ρ_s and parallel distances (hence $k_\parallel$) are normalised to the magnetic scale $L_B/2$.

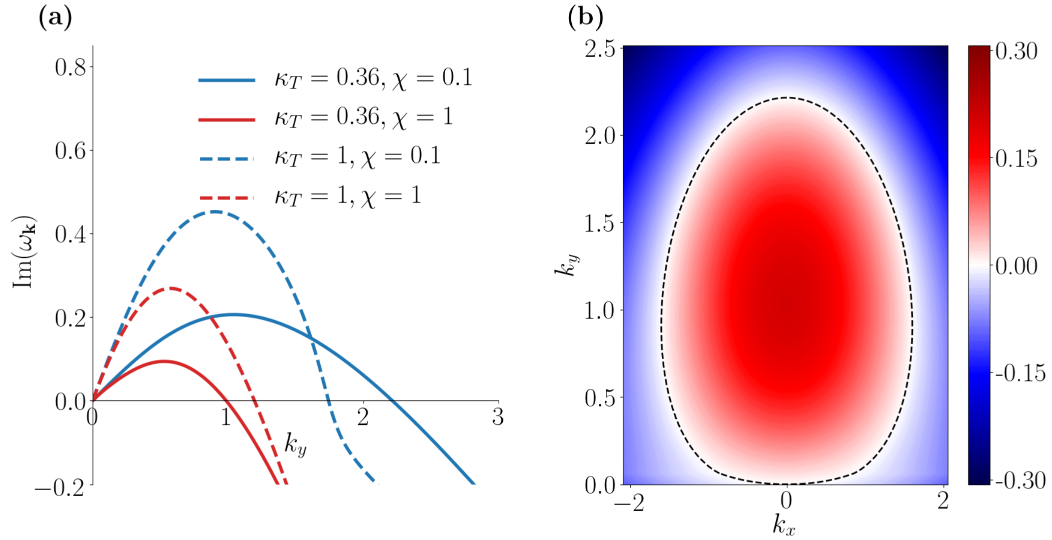


Figure 2.2: **(a)** Dependence of the growth rate $\mathrm{Im}(\omega_{\boldsymbol{k}})$ on k_y for the streamer modes ($k_x = 0$). **(b)** Dependence of $\mathrm{Im}(\omega_{\boldsymbol{k}})$ on k_x and k_y for $\kappa_T = 0.36, \chi = 0.1$. The dashed line is the boundary between stable and unstable modes, viz., $\mathrm{Im}(\omega_{\boldsymbol{k}}) = 0$. In §3.2, we will consider nonlinear simulations with these same parameters.

2.3.1 Curvature-driven ITG instability

Figure 2.2 shows the growth rate $\mathrm{Im}(\omega_{\boldsymbol{k}})$ of the solutions to (2.35) as a function of the wavenumber $\boldsymbol{k}$. Qualitatively, it resembles the growth rate of toroidal ITG modes in tokamaks (Horton *et al.*, 1981). This is expected because the mechanism of the toroidal ITG instability is similar to that of the 2D curvature-driven ITG instability. We find that the fastest growing linear modes are radially extended across the entire box, i.e., they have $k_x = 0$. Such modes are sometimes called “streamers”.

We can get a good qualitative idea of the properties of the instability by setting $\chi = 0$.[6] Then the solution of (2.35) is

$$\omega_{\boldsymbol{k}} = \frac{-k_y(1+\kappa_T k_\perp^2) \pm i k_y \sqrt{4\kappa_T(1+k_\perp^2) - (1+\kappa_T k_\perp^2)^2}}{2(1+k_\perp^2)}, \tag{2.37}$$

so the growth rate of the unstable mode is

$$\mathrm{Im}(\omega_{\boldsymbol{k}}) = \frac{k_y \sqrt{(2\sqrt{\kappa_T} - 1 + \kappa_T k_\perp^2)(2\sqrt{\kappa_T} + 1 - \kappa_T k_\perp^2)}}{2(1+k_\perp^2)}. \tag{2.38}$$

---

[6] The reader might be worried that we are setting $\chi = 0$ in a highly collisional asymptotic limit. This is a question of non-interchangeability of limits. One should think of the “collisionless” case as the limit of $\chi \to 0$, but χ much larger than the small parameters in (2.8); e.g., $\chi \gg \tau$, but both $\chi \to 0$ and $\tau \to 0$.

To simplify further, consider $\kappa_T \gg 1 \gg \kappa_T^{-1/4} \gg k_\perp$. Then

$$\mathrm{Im}(\omega_{\boldsymbol{k}}) \approx k_y \sqrt{\kappa_T}. \tag{2.39}$$

This expression, with the normalisations (2.17) undone, is the well-known "bad-curvature-instability" growth rate (Beer, 1995):

$$\mathrm{Im}(\omega_{\boldsymbol{k}}) = \Omega_i \frac{\rho_i^2 k_y}{\sqrt{2\tau L_B L_T}}. \tag{2.40}$$

Now let us return to the general 2D dispersion (2.35). An important feature of the modes described by it is the boundedness of the region of unstable wavenumbers in the $\boldsymbol{k}$ plane (figure 2.2b). There are both collisionless and collisional mechanisms that lead to the suppression of the ITG instability at small perpendicular scales. Let us consider these mechanisms.

2.3.1.1 Collisionless bounds on unstable wavenumbers

It is easy to see that, in order to be positive, the collisionless growth rate (2.38) requires $k_\perp < k_{\mathrm{max,FLR}}$, where

$$k^2_{\mathrm{max,FLR}} = \frac{1 + 2\sqrt{\kappa_T}}{\kappa_T}. \tag{2.41}$$

This cut-off of the ITG instability is due to FLR effects. For $\kappa_T < 1/4$, (2.38) also gives a lower bound on the wavenumbers k of the unstable collisionless modes, viz., $k_\perp > k_{\mathrm{min,FLR}}$, where

$$k^2_{\mathrm{min,FLR}} = \frac{1 - 2\sqrt{\kappa_T}}{\kappa_T}. \tag{2.42}$$

Adding collisions re-establishes the instability at low k. We deem this to be an unimportant peculiarity of our model, thus we shall only consider $\kappa_T > 1/4$.

2.3.1.2 Collisional bounds on unstable wavenumbers

For nonzero ($\chi > 0$) collisionality, the collisional term AB in (2.35) dominates over the ITG term fAB when $k_\perp$ is large enough and gives strictly damped modes. To show this, note that the instability threshold is given by $\mathrm{Im}(\omega_{\boldsymbol{k}}) = 0$. The real and imaginary parts of (2.35) for $\mathrm{Im}(\omega_{\boldsymbol{k}}) = 0$ are

$$-\mathrm{Re}(\omega_{\boldsymbol{k}})^2 - \mathrm{Re}(\omega_{\boldsymbol{k}})C + (1-f)AB = 0, \tag{2.43}$$

$$-\mathrm{Re}(\omega_{\boldsymbol{k}})(A+B) - (1-g)AC = 0. \tag{2.44}$$

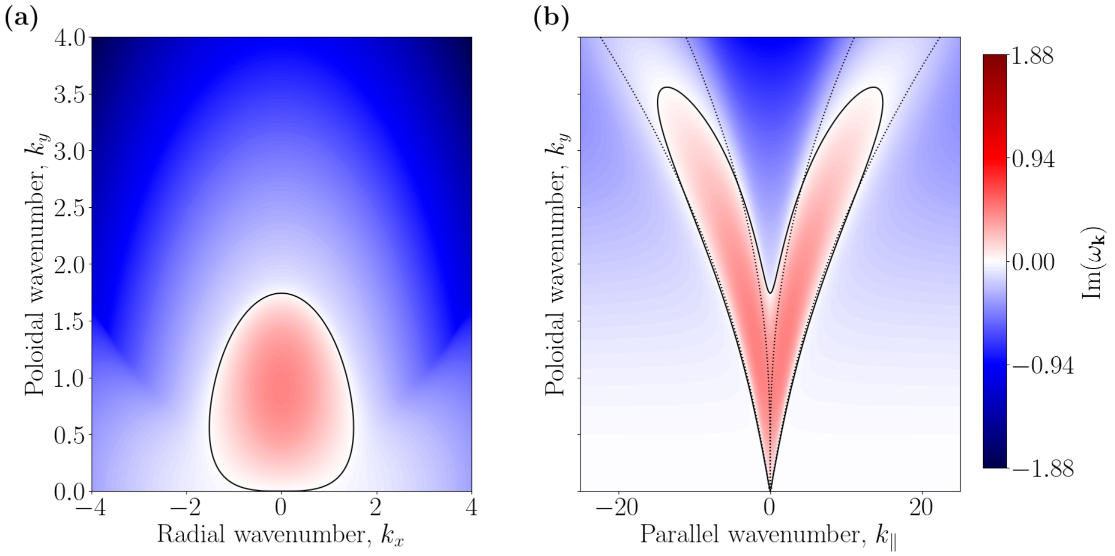


Figure 2.3: A visualisation of the linear growth rate, $\mathrm{Im}(\omega_{\boldsymbol{k}})$, given by (2.34) for $\kappa_T = 1$ and $\chi = 0.1$. **(a)** The linear growth rate in the $k_\parallel = 0$ plane. **(b)** The linear growth rate in the $k_x = 0$ plane (where it is largest). The black lines denote the marginal modes with $\mathrm{Im}(\omega_{\boldsymbol{k}}) = 0$. The dotted lines outline the region of unstable collisionless ($\chi = 0$), pure-slab ($L_B^{-1} = 0$) modes, given by (2.50).

Substituting into (2.43) the value of $\mathrm{Re}(\omega_{\boldsymbol{k}})$ derived from (2.44), and using $A \neq 0$, we find

$$g(1-g)AC^2 + BC^2(1-g) + (1-f)B(A+B)^2 = 0. \tag{2.45}$$

Since $g \in (0,1)$,[7] a necessary condition for instability is

$$f > 1 \implies a\chi^2 k_\perp^6 < \kappa_T k_y^2. \tag{2.46}$$

Thus, the region of unstable modes is bounded by $k_\perp < k_{\max,\chi}$, where

$$k_{\max,\chi}^2 = \sqrt{\frac{\kappa_T}{a\chi^2}}. \tag{2.47}$$

2.3.2 Slab-ITG instability

We now turn to the general 3D dispersion relation (2.34). An example of its solutions is given in figure 2.3. It is evident that (2.34) is too complicated for a general analytical solution. Thus, we will limit our discussion here to several important limits.

---

[7] Note that for $b < 0$ or $b > 1$, there would be a collisional instability in the absence of magnetic curvature, i.e., without the $-\partial_y(\varphi + T)$ term in (2.18). No such instability exists in our model because the Landau collision operator gives $g \in (0,1)$.

We expect that at sufficiently large $k_\parallel$, the curvature-ITG mechanism, viz., the term in (2.18) containing the magnetic-field gradient, will become subdominant. Let us investigate the linear instability of (2.18)–(2.20) without the magnetic-gradient term $-\partial_y(\varphi + T)$ in (2.18); we shall shortly see when this is appropriate. We also limit ourselves to the collisionless ($\chi = 0$) regime in order to make analytical progress. Under these assumptions, (2.34) becomes

$$\left(\hat{\omega}_{\boldsymbol{k}}^2 - \frac{\hat{k}_\parallel^2}{1 + k_\perp^2}\right)(\hat{\omega}_{\boldsymbol{k}} + 1) = \frac{2k_\perp^2 \hat{\gamma}_{\boldsymbol{k}}^2 \hat{\omega}_{\boldsymbol{k}}^2}{1 + k_\perp^2}, \tag{2.48}$$

where we have defined $\omega_{\boldsymbol{k}} \equiv \kappa_T k_y \hat{\omega}_{\boldsymbol{k}}$, $k_\parallel \equiv \kappa_T k_y \hat{k}_\parallel$, and $\hat{\gamma}_{\boldsymbol{k}}^2 \equiv 1/2k_\perp^2$.[8] The last of these may seem like an inconvenience now, but will make the following analysis more easily generalisable to our needs in chapter 4. Since (2.48) is a real cubic in $\hat{\omega}_{\boldsymbol{k}}$, it either has three real solutions, so all linear modes are stable waves, or one real and two complex solutions, in which case one of the complex solutions has a positive imaginary part and thus corresponds to a linearly unstable mode. It can be shown (see appendix C.1) that (2.48) has complex solutions if and only if $\hat{\gamma}_{\boldsymbol{k}}^2 > 0$ and $\hat{k}_\parallel^2 \in (\hat{k}_{\parallel,-}^2, \hat{k}_{\parallel,+}^2)$, where

$$\hat{k}_{\parallel,\pm}^2 = \frac{k_\perp^4 + 10k_\perp^2\hat{\gamma}_{\boldsymbol{k}}^2(1 + k_\perp^2) + k_\perp^2(4 - k_\perp^2\hat{\gamma}_{\boldsymbol{k}}^2) + 2 \pm k_\perp\hat{\gamma}_{\boldsymbol{k}}[4(1 + k_\perp^2) + k_\perp^2\hat{\gamma}_{\boldsymbol{k}}^2]^{3/2}}{2(1 + k_\perp^2)}. \tag{2.49}$$

Substituting $\hat{\gamma}_{\boldsymbol{k}}^2 = 1/2k_\perp^2$ yields

$$\hat{k}_{\parallel,\pm}^2 = \frac{9}{8(1 + k_\perp^2)}\left[\frac{8k_\perp^4}{9} + 4k_\perp^2 + 3 \pm 3\left(1 + \frac{8k_\perp^2}{9}\right)^{3/2}\right]. \tag{2.50}$$

The marginal modes, i.e., those on the boundary between unstable and oscillatory modes, are given by $\hat{k}_\parallel^2 = \hat{k}_{\parallel,\pm}^2$; these are shown in figure 2.4. We now consider two distinct asymptotic limits of (2.50): $k_\perp \ll 1$ and $k_\perp \gg 1$.

2.3.2.1 $k_\perp \ll 1$ modes

To lowest order in $k_\perp \ll 1$, (2.48) simplifies to

$$\hat{\omega}_{\boldsymbol{k}}^3 - \hat{k}_\parallel^2 \hat{\omega}_{\boldsymbol{k}} - \hat{k}_\parallel^2 = 0. \tag{2.51}$$

---

[8] This maps onto equation (49) of Cowley *et al.* (1991) for $Q = \Gamma = 0$ under the following change of notation: $\hat{k}_\parallel \mapsto k_z/k_y$, $\hat{\omega}_{\boldsymbol{k}} \mapsto -\Omega$.

This is the well-known slab-ITG dispersion without FLR effects (Cowley *et al.*, 1991). The instability boundaries (2.50) are

$$\hat{k}_{\parallel,-} = \frac{2}{3\sqrt{3}} k_\perp^3 + \mathcal{O}\big(k_\perp^5\big), \quad \hat{k}_{\parallel,+} = \frac{3\sqrt{3}}{2} + \frac{\sqrt{3}}{4} k_\perp^2 + \mathcal{O}\big(k_\perp^4\big). \tag{2.52}$$

For small $\hat{k}_\parallel$, the linearly unstable solution of (2.51) is $\hat{\omega}_{\boldsymbol{k}} \approx |\hat{k}_\parallel^{2/3}|(-1 + i\sqrt{3})/2$. Thus, the linear growth rate for small $\hat{k}_\parallel$, or $k_\parallel \ll \kappa_T k_y$, is

$$\mathrm{Im}(\omega_{\boldsymbol{k}}) \approx \frac{\sqrt{3}}{2} \big(\kappa_T k_y k_\parallel^2\big)^{1/3}. \tag{2.53}$$

The 2D curvature-ITG growth rate and poloidal wavenumber scale as $\omega_{\boldsymbol{k},\mathrm{2D}} \sim \kappa_T^{1/4}$ and $k_y \sim \kappa_T^{-1/4}$ in the strongly driven limit $\kappa_T \gg 1$ (see §2.3.1). Using (2.53), we can estimate that the 3D modes will dominate the 2D ones, and thus it is appropriate to ignore the magnetic drift if

$$\big(\kappa_T k_y k_\parallel^2\big)^{1/3} \gg \omega_{\boldsymbol{k},\mathrm{2D}} \sim \kappa_T^{1/4} \implies k_\parallel \gg 1. \tag{2.54}$$

Thus, in the $\kappa_T \gg 1$ limit, $k_\parallel \lesssim 1$ modes are essentially curvature-ITG modes with finite-$k_\parallel$ modification, while the $k_\parallel \gg 1$ modes are slab-ITG modes.

2.3.2.2 $k_\perp \gg 1$ modes

Expanding (2.49) for $k_\perp \gg 1$ and using $\hat{\gamma}_{\boldsymbol{k}} \sim \mathcal{O}\big(k_\perp^{-1}\big)$, we find

$$\hat{k}_{\parallel,\pm} = k_\perp \pm 2\hat{\gamma}_{\boldsymbol{k}} k_\perp + \mathcal{O}\big(k_\perp^{-1}\big). \tag{2.55}$$

Therefore, for $k_\perp \gg 1$, the slab-ITG instability is localised at $\hat{k}_\parallel = \pm k_\perp$, or, equivalently,

$$k_\parallel \approx \pm \kappa_T k_y k_\perp. \tag{2.56}$$

For $\hat{\gamma}_{\boldsymbol{k}}^2 = 1/2k_\perp^2$, (2.55) is

$$\hat{k}_{\parallel,\pm} = k_\perp \pm \sqrt{2} + \mathcal{O}\big(k_\perp^{-1}\big), \tag{2.57}$$

which implies

$$\left|\hat{k}_\parallel - k_\perp\right| = \left|\frac{k_\parallel}{\kappa_T k_y} - k_\perp\right| < \sqrt{2}. \tag{2.58}$$

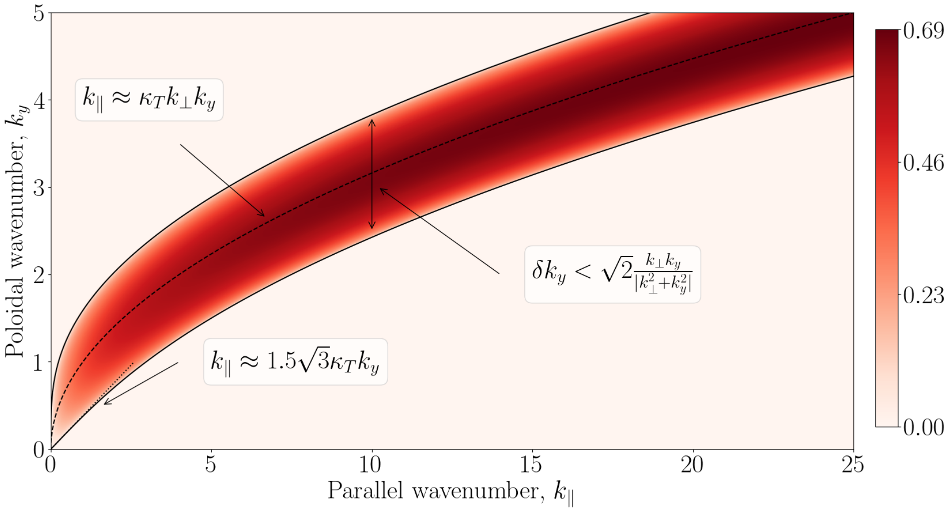


Figure 2.4: Linear growth rates for the slab-ITG instability (without magnetic drift) as a function of parallel ($k_\parallel$) and poloidal (k_y) wavenumbers for $k_x = 0$, $\kappa_T = 1$, $\chi = 0$. The growth rate along the $k_\parallel = \kappa_T k_\perp k_y$ line converges to $\kappa_T/\sqrt{2} \approx 0.7$ for large $k_\parallel$. The solid black lines are given by (2.50).

We can also find the k_y width of the region of instability at fixed $k_\parallel$. Substituting $k_y = k_\parallel/\kappa_T k_\perp + \delta k_y$ in (2.58) and expanding for $\delta k_y \ll k_y$, we find

$$|\delta k_y| < \sqrt{2}\frac{k_\perp k_y}{|k_\perp^2 + k_y^2|} \leq \frac{\sqrt{2}}{2}. \tag{2.59}$$

To find the growth rate, consider the $k_\perp \gg 1$ limit of (2.48) and set $\hat{\omega}_{\boldsymbol{k}} = -1 + \delta\hat{\omega}_{\boldsymbol{k}}$, $\hat{k}_\parallel = \pm k_\perp + \delta\hat{k}_\parallel$, where $\delta\hat{\omega}_{\boldsymbol{k}} \sim \delta\hat{k}_\parallel/k_\perp \sim \mathcal{O}(k_\perp^{-1}) \ll 1$. Keeping terms of order up to $\mathcal{O}(k_\perp^{-2})$, (2.48) becomes

$$\left(\delta\hat{\omega}_{\boldsymbol{k}} \pm \frac{\delta\hat{k}_\parallel}{k_\perp}\right)\delta\hat{\omega}_{\boldsymbol{k}} + \hat{\gamma}_{\boldsymbol{k}}^2 \approx 0 \implies \delta\hat{\omega}_{\boldsymbol{k}} \approx -\frac{\delta\hat{k}_\parallel}{2k_\perp} \pm \sqrt{\frac{\delta\hat{k}_\parallel^2}{4k_\perp^2} - \hat{\gamma}_{\boldsymbol{k}}^2}. \tag{2.60}$$

Thus, in agreement with (2.58), the instability exists only for $|\delta\hat{k}_\parallel| < 2k_\perp\hat{\gamma}_{\boldsymbol{k}}$, and its growth rate is

$$\mathrm{Im}(\hat{\omega}_{\boldsymbol{k}}) \approx \sqrt{\hat{\gamma}_{\boldsymbol{k}}^2 - \frac{\delta\hat{k}_\parallel^2}{4k_\perp^2}}. \tag{2.61}$$

The maximum growth rate is then achieved for $\delta\hat{k}_\parallel = 0$ and is given by $\mathrm{Im}(\hat{\omega}_{\boldsymbol{k}}) \approx \hat{\gamma}_{\boldsymbol{k}}$. We can substitute $\hat{\gamma}_{\boldsymbol{k}}^2 = 1/2k_\perp^2$ to find

$$\mathrm{Im}(\omega_{\boldsymbol{k}}) \approx \frac{\kappa_T k_y}{\sqrt{2}k_\perp} \leq \frac{\kappa_T}{\sqrt{2}}. \tag{2.62}$$

The characteristics of the slab-ITG instability are summarised in figure 2.4.

Note that for $\kappa_T \gg 1$, the linear growth rate of the small-scale slab-ITG modes scales as $\mathcal{O}(\kappa_T)$ and, therefore, dominates the curvature-ITG modes, whose linear time scale is $\mathcal{O}(\kappa_T^{1/4})$. This time-scale separation will prove crucial for the saturation of strongly driven turbulence: see §4.2.2.

Finally, an important feature of these $k_\parallel \gg 1$ slab-ITG modes is the approximate relation $T \approx -\varphi$, or equivalently, $p/\varphi \ll 1$, where $p = \varphi + T$ is the perturbed pressure. Indeed, using (2.19) and (2.60), we find

$$\frac{T_{\boldsymbol{k}}}{\varphi_{\boldsymbol{k}}} = \frac{\kappa_T k_y}{\omega_{\boldsymbol{k}}} = \frac{1}{\hat{\omega}_{\boldsymbol{k}}} = -1 + \frac{\delta\hat{k}_\parallel}{2k_\perp} - i\sqrt{\hat{\gamma}_{\boldsymbol{k}}^2 - \frac{\delta\hat{k}_\parallel^2}{4k_\perp^2}} + \mathcal{O}\big(k_\perp^{-2}\big) \tag{2.63}$$

for the modes with $\mathrm{Im}(\hat{\omega}_{\boldsymbol{k}}) > 0$. Thus, these modes have $p_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}} \sim \mathcal{O}(k_\perp^{-1}) \ll 1$, and the most unstable ones ($\delta\hat{k}_\parallel = 0$) satisfy $p_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}} \sim \mathcal{O}(k_\perp^{-2})$ and $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) = -1 + \mathcal{O}(k_\perp^{-2})$. This relationship of $T_{\boldsymbol{k}}$ and $\varphi_{\boldsymbol{k}}$ will allow us to identify the slab-ITG modes in the saturated state, and will prove useful in understanding their role in maintaining the 3D Dimits state, see §4.2.1 and §4.2.4.

Chapter 3

Saturation in 2D: shear zones and turbulent momentum flux

As a first step towards understanding the behaviour of our cold-ion ITG system (2.18)–(2.20), we shall focus on the 2D regime that is described by (2.22) and (2.23). There are at least three reasons to do this. First, the 2D system is significantly simpler to understand both analytically and numerically, and so we will be able to make substantial progress. Secondly, it allows us to put our work in context: as mentioned in §2.1.5, (2.22) and (2.23) are part of the Hasegawa-Mima family of 2D models that are very popular in the community. Last but not least, we shall discover that the 2D model contains rich ZF physics, much of which will carry over naturally to the 3D model in chapter 4. In particular, the mechanisms which establish the ZF-dominated Dimits state of (2.18)–(2.20) are contained within (2.22) and (2.23). Therefore, the rest of chapter 3 will focus entirely on (2.22) and (2.23).

Before we delve into the study of nonlinear saturation, let us show how ZFs can be generated from the linearly unstable ITG modes. As the curvature-ITG driving terms, viz., the magnetic drift $-\partial_y(\varphi + T)$ in (2.22) and the advection of equilibrium temperature $\kappa_T \partial_y \varphi$ in (2.23), are both proportional to k_y, any zonal, i.e., $k_y = 0$, mode is linearly stable. Thus, if ZFs are to be generated, the linearly unstable $k_y \neq 0$ must themselves be unstable to developing ZFs. This is known as the "secondary" instability.

3.1 Secondary instability

Consider the stability of a streamer mode with $k_x = 0, k_y = q$ (the "primary" mode) to infinitesimal "secondary" perturbations:

$$\varphi = (\varphi_q e^{iqy} + \text{c.c.}) + \delta\varphi(x, y), \tag{3.1}$$

$$T = (T_q e^{iqy} + \text{c.c.}) + \delta T(x, y). \tag{3.2}$$

A common way of analysing the secondary instability is to take a Galerkin truncation by considering only four Fourier modes $(k_x, k_y) = \{(0, q), (p, 0), (p, \pm q)\}$ and their complex conjugates: the $(0, q)$ mode is the primary streamer in (3.1) and (3.2) and the others are

$$\delta\varphi = \big(\delta\varphi_+ e^{iqy} + \delta\varphi_- e^{-iqy} + \delta\varphi_0\big) e^{ipx} e^{\gamma_2 t} + \text{c.c.}, \tag{3.3}$$

$$\delta T = \big(\delta T_+ e^{iqy} + \delta T_- e^{-iqy} + \delta T_0\big) e^{ipx} e^{\gamma_2 t} + \text{c.c.}, \tag{3.4}$$

where p is the radial wavenumber of the secondary perturbations, $\delta\varphi_0$ and δT_0 are the zonal flow and temperature, and $\delta\varphi_\pm$ and $\delta T_\pm$ are known as "sidebands". Substituting all this into (2.22) and (2.23) and linearising the nonlinear terms for $\delta\varphi \ll \varphi_q$ and $\delta T \ll T_q$, we obtain a closed set of equations. In order to keep things simple, we drop the linear terms in (2.22) and (2.23) — this is valid when the amplitude of the primary mode is large enough, so that interactions with it are more important for the evolution of $\delta\varphi$ and δT than the effects of the equilibrium gradients and collisions. Observe that, due to the structure of the Poisson bracket (2.12), all nonlinear terms are proportional to pq. Defining for convenience $\gamma_2 \equiv \sqrt{2} pq|\varphi_q|\hat{\gamma}_2$, we obtain the following equation for $\hat{\gamma}_2$:

$$\big(\hat{\gamma}_2^2 + U\big)\big(\hat{\gamma}_2^2 + V\big) = W, \tag{3.5}$$

where

$$\begin{aligned}
U &= 1 + \frac{q^2 \text{Re}(T_q/\varphi_q)}{1 + p^2 + q^2}, \\
V &= \frac{p^2 \text{Im}(T_q/\varphi_q)^2 + p^2[1 + \text{Re}(T_q/\varphi_q)]^2 - (1 + q^2)[1 + \text{Re}(T_q/\varphi_q)]}{1 + p^2 + q^2}, \\
W &= \frac{p^2 q^2}{(1 + p^2 + q^2)^2}\Big[|T_q|^2/|\varphi_q|^2 + 2\text{Re}(T_q/\varphi_q)\Big]\Big[1 + \text{Re}(T_q/\varphi_q)\Big].
\end{aligned} \tag{3.6}$$

We see that the growth rate γ_2 of the secondary instability depends both on the amplitudes of the primary fields φ_q and T_q, and on their relative phase.

3.1.1 No temperature perturbation

If we set $T_q = 0$ in (3.5), i.e., ignore the temperature perturbation of the primary streamer, we obtain the well-known dispersion relation for the secondary instability of the modified Hasegawa-Mima model (Rogers *et al.*, 2000; Strintzi & Jenko, 2007):

$$\gamma_2^{\mathrm{HM}} = pq|\varphi_q|\sqrt{\frac{2(1+q^2-p^2)}{1+p^2+q^2}}. \tag{3.7}$$

This form of the secondary instability has long been associated with the strong ZFs observed numerically in ITG turbulence (Hammett *et al.*, 1993).[1]We will show that the inclusion of the temperature perturbations can introduce qualitative and quantitative changes, and even suppress the secondary instability completely.

3.1.2 Long-wavelength limit

To simplify (3.5), we can consider the long-wavelength limit $p \ll 1$. Then (3.5) gives

$$\left[\hat{\gamma}_2^2 + 1 + \frac{q^2\mathrm{Re}(T_q/\varphi_q)}{1+q^2}\right]\left[\hat{\gamma}_2^2 - 1 - \mathrm{Re}(T_q/\varphi_q)\right] = \mathcal{O}\big(p^2\big) \approx 0. \tag{3.8}$$

Thus, there are two independent branches of the secondary instability with instability conditions given by $\mathrm{Re}(T_q/\varphi_q) + 1/q^2 < -1$ and $\mathrm{Re}(T_q/\varphi_q) > -1$, respectively. The second branch is a modified form of the long-wavelength Hasegawa-Mima secondary instability (3.7):[2]

$$\gamma_2 = \sqrt{2}pq|\varphi_q|\sqrt{1+\mathrm{Re}\left(\frac{T_q}{\varphi_q}\right)}. \tag{3.9}$$

We observe that (3.9) relies only on a handful of the nonlinear terms in (2.22) and (2.23). Substituting (3.1) and (3.3) into (2.22) and taking the limit $p \ll 1$ gives us the

---

[1] Especially in contrast with the much weaker ZFs observed in electron-temperature-gradient-driven (ETG) turbulence on electron scales (Jenko *et al.*, 2000; Strintzi & Jenko, 2007). However, this distinction between ITG and ETG turbulence has recently been challenged by Colyer *et al.* (2017), who found that the long-time saturated state of ETG turbulence is also dominated by ZFs, although the system does go through a streamer-dominated quasi-saturated state at earlier times.

[2] Plunk & Bañón Navarro (2017) found the same expression in the context of the "warm-ion" approximation, i.e., dropping the FLR terms in the nonzonal part of (2.22).

following equations for $\delta\varphi_0, \delta\varphi_+$ and $\delta\varphi_-$:

$$\hat{\gamma}_2 \delta\varphi_0 = \frac{1}{|\varphi_q|\sqrt{2}}\big[\delta\varphi_+(\varphi_q^* + T_q^*) - \delta\varphi_-(\varphi_q + T_q)\big], \tag{3.10}$$

$$\hat{\gamma}_2 \delta\varphi_+ = \frac{1}{|\varphi_q|\sqrt{2}}\varphi_q \delta\varphi_0, \tag{3.11}$$

$$\hat{\gamma}_2 \delta\varphi_- = -\frac{1}{|\varphi_q|\sqrt{2}}\varphi_q^* \delta\varphi_0. \tag{3.12}$$

Substituting (3.11) and (3.12) into (3.10) yields precisely (3.9). We do not consider the equations for the temperature perturbations because $\delta T = 0$ is a consistent solution and it corresponds to (3.9). The terms on the right-hand side of (3.11) and (3.12) arise from the zonal advection term $\{\overline{\varphi}, \varphi' - \nabla^2\varphi'\}$ in (2.22) and represent the tilting of the primary streamer by the ZF. The terms on the right-hand side of (3.10) are the poloidal $\boldsymbol{E} \times \boldsymbol{B}$ and diamagnetic flows caused by the interaction of the primary mode and the two sidebands $(k_x, k_y) = (p, \pm q)$. The quantity $\mathrm{Re}(T_q/\varphi_q)$ controls the response of the primary mode to the zonal perturbation: $\mathrm{Re}(T_q/\varphi_q) > -1$ yields an unstable ZF, while $\mathrm{Re}(T_q/\varphi_q) < -1$ results in a stable, oscillatory perturbation.

Let us consider the collisionless case ($\chi = 0$), where analytical progress is possible, and ask for what values of κ_T the two modes described by (3.8) are unstable. Let us take $(k_x, k_y) = (0, q)$ to be the linear mode with the largest growth rate. We then define the critical gradient κ_T^{sec} for the long-wavelength secondary instability of the fastest-growing streamer as the value of κ_T at which $\mathrm{Re}(T_q/\varphi_q) = -1$. The ratio of the temperature and potential perturbations can be obtained from (2.23):

$$\frac{T_{\boldsymbol{k}}}{\varphi_{\boldsymbol{k}}} = \frac{-ik_y\kappa_T}{-i\omega_{\boldsymbol{k}} + \chi k_\perp^2}, \tag{3.13}$$

where $\omega_{\boldsymbol{k}}$ is the solution of the dispersion relation (2.35). Thus, using the collisionless dispersion relation (2.37), we obtain

$$\mathrm{Re}\left(\frac{T_q}{\varphi_q}\right) = -\frac{1 + \kappa_T q^2}{2}. \tag{3.14}$$

To determine q, we seek the maximum of $\mathrm{Im}(\omega_{\boldsymbol{k}})$, as given by (2.38) for $\boldsymbol{k} = (0, q)$. We find

$$\frac{\partial \mathrm{Im}(\omega_{\boldsymbol{k}})}{\partial q} \propto \kappa_T^2 q^6 + 3\kappa_T^2 q^4 - q^2 - 4\kappa_T + 1 = 0, \tag{3.15}$$

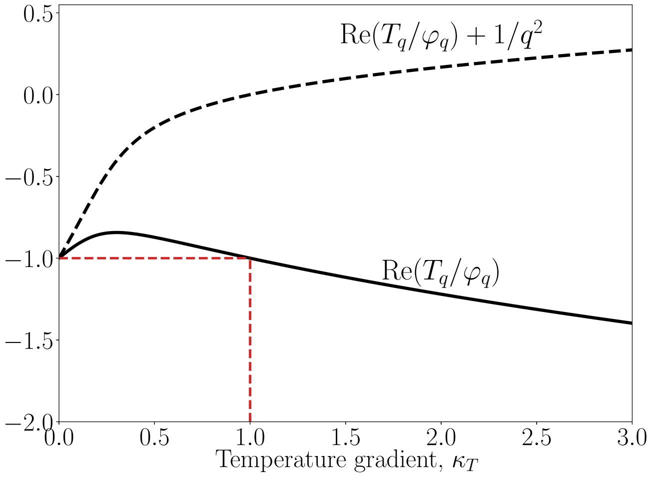


Figure 3.1: Temperature-gradient dependence of $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})$ (solid) and $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) + 1/q^2$ (dashed) for the most unstable collisionless ($\chi = 0$) mode. We find that $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) < -1$ for $\kappa_T > 1$ and $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) + 1/q^2 > -1$ for all κ_T. The secondary instability is present only for $\kappa_T < 1$.

where the equality holds for the most unstable mode. As an equation for q^2, (3.15) is a cubic with only one positive solution for $\kappa_T > 1/4$. Substituting that solution into (3.14), we find $\mathrm{Re}(T_q/\varphi_q)$ as a function of κ_T. This relationship is given in figure 3.1. In particular, we obtain that $\mathrm{Re}(T_q/\varphi_q) = -1$ at $\kappa_T = 1$, as can indeed be verified analytically from (3.15) and (3.14), and $\mathrm{Re}(T_q/\varphi_q) < -1$ for $\kappa_T > 1$. We also find that $\mathrm{Re}(T_q/\varphi_q) + 1/q^2 > -1$ always. Thus, for $\kappa_T > 1$, the most unstable collisionless ITG mode is stable to the secondary perturbations. Note that (3.14) depends crucially on the diamagnetic drift $\kappa_T \partial_y \nabla_\perp^2 \varphi$ in (2.22). If we do not include the diamagnetic drift, we find that $\mathrm{Re}(T_q/\varphi_q) = -1/2$ regardless of κ_T and q, and thus the collisionless secondary instability is never quenched.

3.1.3 General case

Let us go back to the general secondary dispersion relation (3.5). Its solution is

$$\hat{\gamma}_2^2 = \frac{-(U+V) \pm \sqrt{(U-V)^2 + 4W}}{2}, \tag{3.16}$$

where

$$U+V=\frac{p^2\big[1+\mathrm{Im}(T_q/\varphi_q)^2\big]+p^2[1+\mathrm{Re}(T_q/\varphi_q)]^2-\mathrm{Re}(T_q/\varphi_q)}{1+p^2+q^2}. \tag{3.17}$$

We can use the primary dispersion relation (2.35) to show that $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})<0$, and hence $U+V>0$, for any unstable primary mode with wavenumber $\boldsymbol{k}$. Indeed, the real part of (3.13) is

$$\mathrm{Re}\left(\frac{T_{\boldsymbol{k}}}{\varphi_{\boldsymbol{k}}}\right)=\frac{k_y\kappa_T\mathrm{Re}(\omega_{\boldsymbol{k}})}{|-i\omega_{\boldsymbol{k}}+\chi k_\perp^2|^2}<0 \tag{3.18}$$

if $k_y\mathrm{Re}(\omega_{\boldsymbol{k}})<0$. Let us show that this is true. For $k\ll 1$ and $\kappa_T>1/4$ (the reasons for the latter are discussed at the end of §2.3.1.1), the dispersion relation (2.35) gives simply $k_y\mathrm{Re}(\omega_{\boldsymbol{k}})=-k_y^2/2<0$. Since the solutions to (2.35) are continuous functions of $\boldsymbol{k}$, if $k_y\mathrm{Re}(\omega_{\boldsymbol{k}})$ changes sign and becomes positive, then $\mathrm{Re}(\omega_{\boldsymbol{k}})=0$ somewhere. However, if we set $\mathrm{Re}(\omega_{\boldsymbol{k}})=0$, the imaginary part of (2.35) gives $\mathrm{Im}(\omega_{\boldsymbol{k}})=(g-1)A<0$. Therefore, $k_y\mathrm{Re}(\omega_{\boldsymbol{k}})$ cannot change sign within the region of linear instability and so $k_y\mathrm{Re}(\omega_{\boldsymbol{k}})<0$ for all linearly unstable modes.

We now consider the solution (3.16) assuming that the relationship between φ_q and T_q is given by (3.13) with $k_y=q$ and $\omega_{\boldsymbol{k}}$ corresponding to the most unstable mode. This gives us γ_2 as a function of κ_T, χ and p. Figure 3.2 shows the real part of γ_2 maximised over p for each pair of equilibrium parameters κ_T and χ, and the wavenumber $p_{\max}$ at which that maximum is attained. Let us discuss this figure. There are three distinct regions:

1. $\kappa_T<\kappa_T^{\mathrm{sec}}$, where κ_T^{sec} is defined as the value of κ_T where $\mathrm{Re}(T_q/\varphi_q)=-1$; in this region, $\mathrm{Re}(T_q/\varphi_q)>-1$. Additionally, $UV-W<0$ for $p=p_{\max}$, so $\hat{\gamma}_2^2$ given by (3.16) is real and positive. The instability exists for arbitrarily small values of p (i.e., for an arbitrarily long wavelength of the ZF). Increasing the temperature gradient κ_T towards κ_T^{sec} has a dramatic effect on the secondary instability of the most unstable mode: it diminishes both the growth rate and the region of zonal wavenumbers that go unstable. On the line $\kappa_T=\kappa_T^{\mathrm{sec}}$, $\hat{\gamma}_2$ is purely imaginary and there are no growing secondary modes, just like in the long-wavelength analysis of §3.1.2. Indeed, substituting $\mathrm{Re}(T_q/\varphi_q)=-1$ in (3.6), we obtain $W=0$ and $U,V>0$. Then, by (3.16), $\hat{\gamma}_2^2=-U$ or $-V$. Figure 3.3 ($\kappa_T=0.7,1.1,1.5$) shows γ_2 vs. p in region (i).

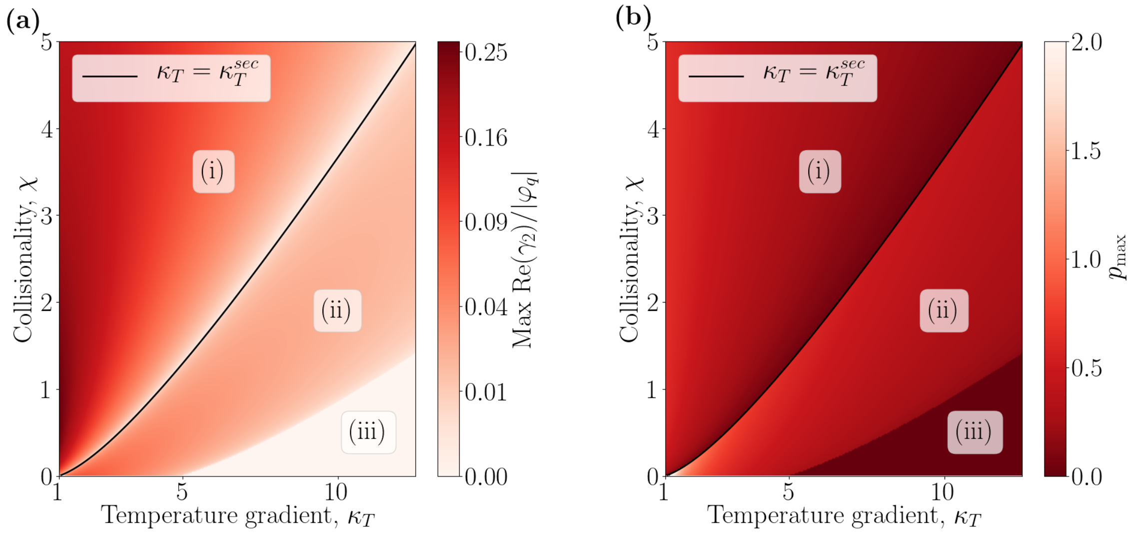


Figure 3.2: **(a)** Secondary-instability growth rate (3.16) of the most unstable streamer mode, maximised over all values of p. The growth rate vanishes on the $\mathrm{Re}(T_q/\varphi_q) = -1$ curve (shown in black). **(b)** Radial wavenumber $p_{\max}$ at which the maximum growth rate shown in (a) is attained. A discontinuity in the most unstable wavenumber across the $\mathrm{Re}(T_q/\varphi_q) = -1$ curve is evident. The absolute-stability region (iii), visible in the bottom right of both (a) and (b), where $\gamma_2 = 0$, is the region where $(U-V)^2 + 4W$ is always positive and γ_2 is purely imaginary.

2. $\kappa_T > \kappa_T^{\mathrm{sec}}$. Increasing κ_T past κ_T^{sec} changes the fastest-growing secondary mode discontinuously. The fastest-growing secondary mode now has $UV - W > 0$ and $(U-V)^2 + 4W < 0$. Hence the $\hat{\gamma}_2^2$ given by (3.16) is complex. In this region, there is always $\hat{\gamma}_2$ with a positive real part. The peak-growth wavenumber $p_{\max}$ changes discontinuously across the $\mathrm{Re}(T_q/\varphi_q) = -1$ line. For $\kappa_T > \kappa_T^{\mathrm{sec}}$, the secondary instability does not extend to arbitrarily small p (figure 3.3, $\kappa_T = 1.9, 2.3$), consistent with the discussion of the long-wavelength secondary instability in §3.1.2.

3. $\kappa_T > \kappa_T^{\mathrm{sec}}$, but now $(U-V)^2 + 4W > 0$ for all values of p, so $\hat{\gamma}_2^2$ given by (3.16) is real and negative. The location of this region of stability depends on the value of $\mathrm{Im}(T_q/\varphi_q)$, as well as $\mathrm{Re}(T_q/\varphi_q)$, and does not have a simple analytic form like the boundary between regions (i) and (ii).

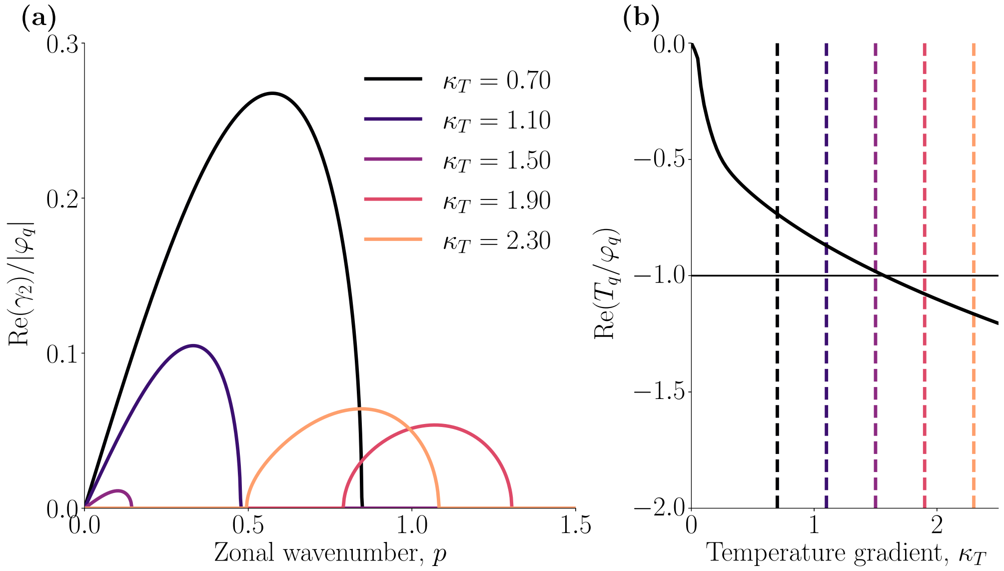


Figure 3.3: **(a)** Secondary-instability growth rate for $\chi = 0.1$ and a number of values of κ_T versus zonal (radial) wavenumber, as given by (3.5). The primary mode φ_q, T_q is taken to be the most unstable one in every case. **(b)** $\mathrm{Re}(T_k/\varphi_k)$ for the fastest-growing linear mode $\boldsymbol{k} = (0, q)$ for $\chi = 0.1$ as a function of κ_T. The dashed lines correspond to the same values of κ_T as in (a).

This analysis of the secondary instability suggests that the system will fail to generate ZFs at a high enough κ_T. In what follows, we will indeed find that the zonally dominated Dimits regime ceases to exist when the temperature gradient exceeds a certain threshold, $\kappa_T > \kappa_T^{c,\mathrm{2D}}$. However, the naïve guess $\kappa_T^{c,\mathrm{2D}} \approx \kappa_T^{\mathrm{sec}}$, as given by the secondary-instability threshold of the most unstable streamer, does not yield satisfactory agreement with the observed threshold for the 2D Dimits regime (see §3.3.4). The secondary-instability picture is incomplete because we must take into account not only whether ZFs can be generated by the ITG modes, but also whether the strong ZFs that support the Dimits regime are resilient to nonzonal perturbations. We shall pick up this topic in §3.3.

3.2 Nonlinear saturation and zonal staircase

We now proceed to investigate the saturated state of (2.22) and (2.23) numerically and semi-analytically. We solve these equations numerically in a doubly periodic box

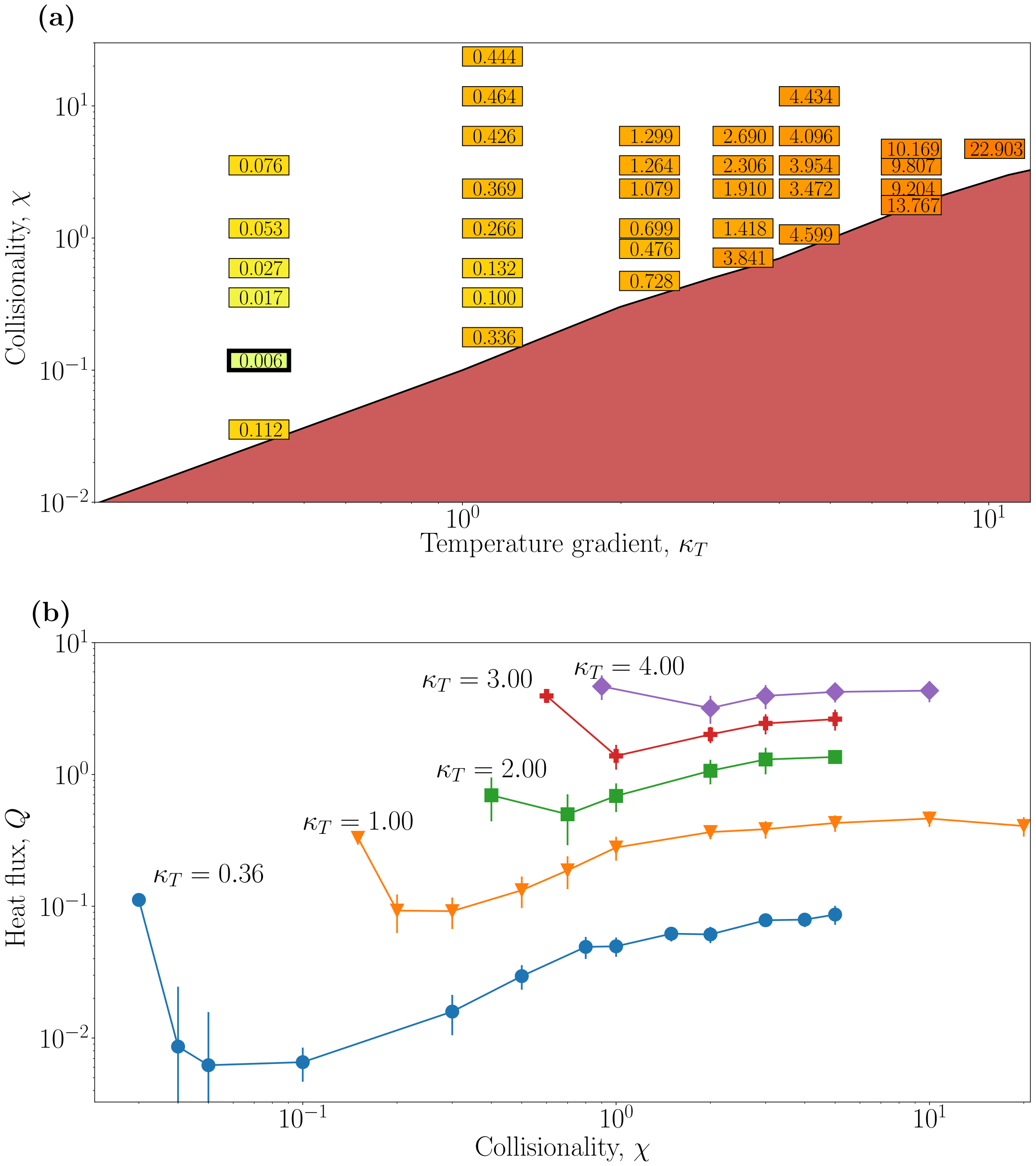


Figure 3.4: **(a)** Box-averaged heat flux Q as a function of κ_T and χ. Q is defined in (2.29) and given here in units of $3n_iT_iv_{ti}(\rho_i/L_B)^2/\tau^{5/2}\sqrt{2}$; see also the footnote on page 31. The shaded (in red) region is beyond the Dimits threshold, where strong turbulence resides (see §3.3.5). Bold-framed is the parameter point corresponding to SimL and SimH, viz., $\kappa_T = 0.36, \chi = 0.1$. **(b)** Box-averaged heat flux Q in the saturated state versus χ for various κ_T.

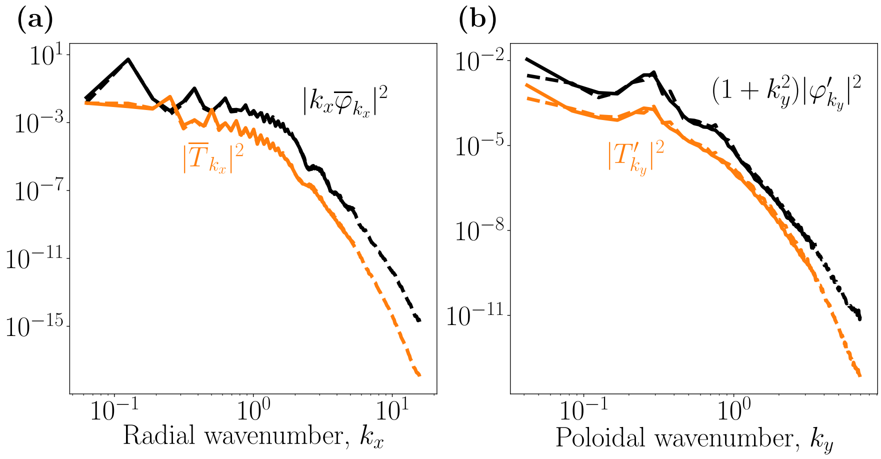


Figure 3.5: Comparison of the spectra of turbulence for SimL (solid) and SimH (dashed), described in §3.2. **(a)** ZF velocity $|k_x\overline{\varphi}_{k_x}|^2$ (black) and zonal temperature (orange) $|\overline{T}_{k_x}|^2$ spectra. **(b)** Streamer ($k_x = 0$) contribution to (2.26), viz., $(1+k_y^2)|\varphi'_{k_y}|^2$, (black) and temperature $|T'_{k_y}|^2$ (orange). A clear peak at $k_y \approx 0.25$ is seen. This corresponds to the dominant poloidal wavenumber in the ZF minima (see also figure B.1 in appendix B). The fastest linearly growing streamer has $k_y \approx 1$.

of size L_x and L_y using a pseudo-spectral algorithm. Numerical integration is done in Fourier space with the nonlinear term calculated in real space using the 2/3 rule for dealiasing (Orszag, 1971). We integrate the linear terms implicitly in time using the Crank–Nicolson method, while the nonlinear terms are integrated explicitly using the Adams-Bashforth three-step method. This integration scheme is similar to the one implemented in the popular gyrokinetic code GS2 (Kotschenreuther *et al.*, 1995*b*; Dorland *et al.*, 2000).

A finite-amplitude saturated state is found only for temperature gradients below a critical gradient $\kappa_T < \kappa_T^{c,2\mathrm{D}}$, where $\kappa_T^{c,2\mathrm{D}}$ is an increasing function of collisionality χ (see figure 3.4). The saturated state is always dominated by strong zonal flows and exhibits levels of turbulent transport that are low compared to the equilibrium diffusive transport ($Q \lesssim \chi\kappa_T$). We will refer to this state as the Dimits state. The critical gradient $\kappa_T^{c,2\mathrm{D}}$ is then the nonlinear critical gradient that marks the break up of the zonally dominated state and the onset of fully developed ITG turbulence. We will relate $\kappa_T^{c,2\mathrm{D}}$ to the resilience of the zonal profiles in the face of nonzonal perturbations, which is in turn determined by the behaviour of turbulence in the presence of strong (comparable to the ITG growth rate) zonal shear. For $\kappa_T < \kappa_T^{c,2\mathrm{D}}$, zonally sheared

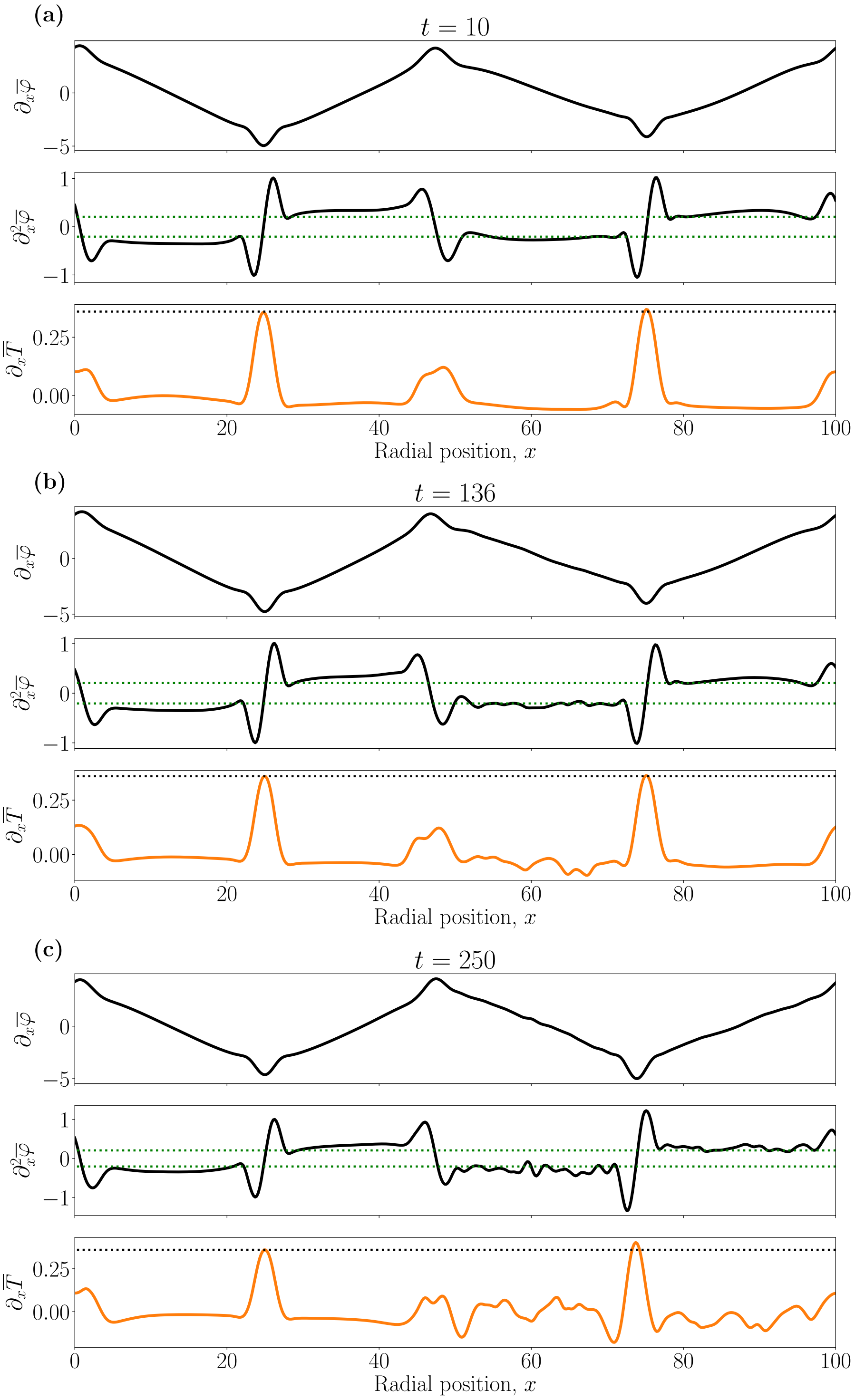


Figure 3.6: Radial profiles of ZF ($\partial_x\overline{\varphi}$), zonal shear ($\partial_x^2\overline{\varphi}$), and zonal temperature gradient ($\partial_x\overline{T}$) **(a)** in the quiescent period, **(b)** at the onset of a turbulent burst, **(c)** during a turbulent burst. The dotted green and black lines are the largest linear ITG growth rate and the equilibrium temperature gradient κ_T, respectively. The turbulence that develops in the shear zones does not disturb the ZF and zonal shear significantly. See figure 3.7 for 2D snapshots at these same times. The data is from SimH.

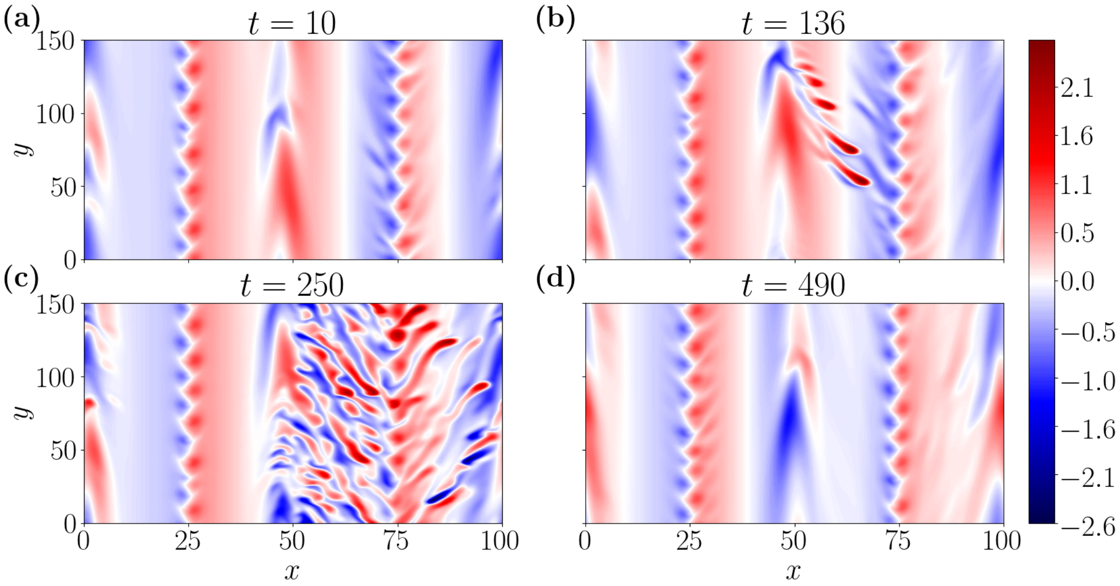


Figure 3.7: Snapshots of temperature perturbations in the Dimits state. The data is from SimH. **(a)** Quiescent, $t = 10$; there is a ZF minimum at $x = 24.5$ and a ZF maximum at $x = 47$. **(b)** Ferdinons visible around $x = 65$, $t = 136$. A zoomed-in version can be found in figure 3.17. **(c)** Turbulent burst, $t = 250$. **(d)** Relaxation back to the zonal staircase after the burst, $t = 490$. The full time history of SimH can be found in figure 3.8 and the radial profiles of the zonal fields and heat flux are shown in figure 3.6.

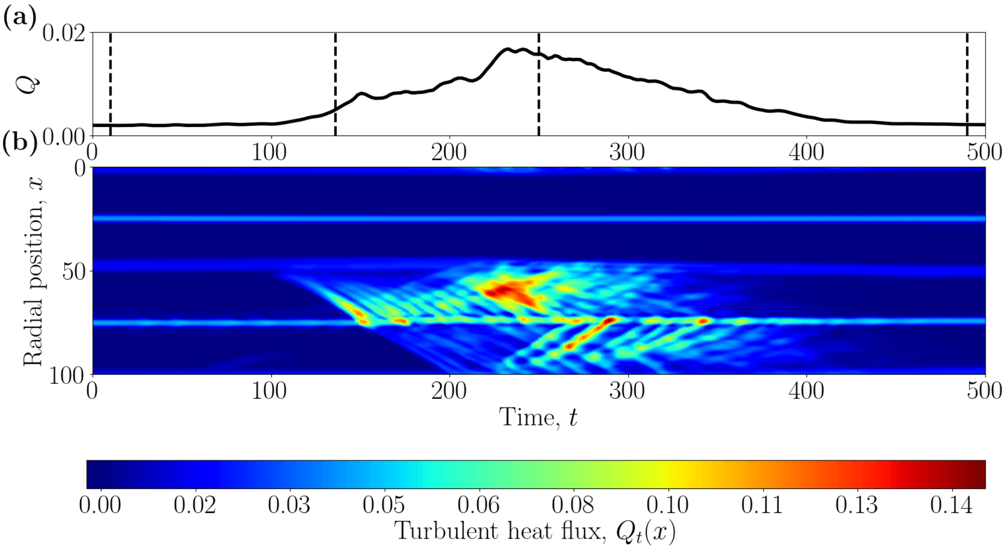


Figure 3.8: **(a)** Time evolution of the total heat flux Q during a turbulent burst for SimH. The dashed lines correspond to the times used for figures 3.6 and 3.7. **(b)** Time trace of the local (poloidally averaged) radial turbulent heat flux $Q_t(x) = -\overline{T\partial_y\varphi}$ as a function of radial position. A turbulent burst in the right half of the domain is clearly visible for $t \in [100, 400]$. The linear streaks correspond to radially drifting ferdinons (see §3.2.3).

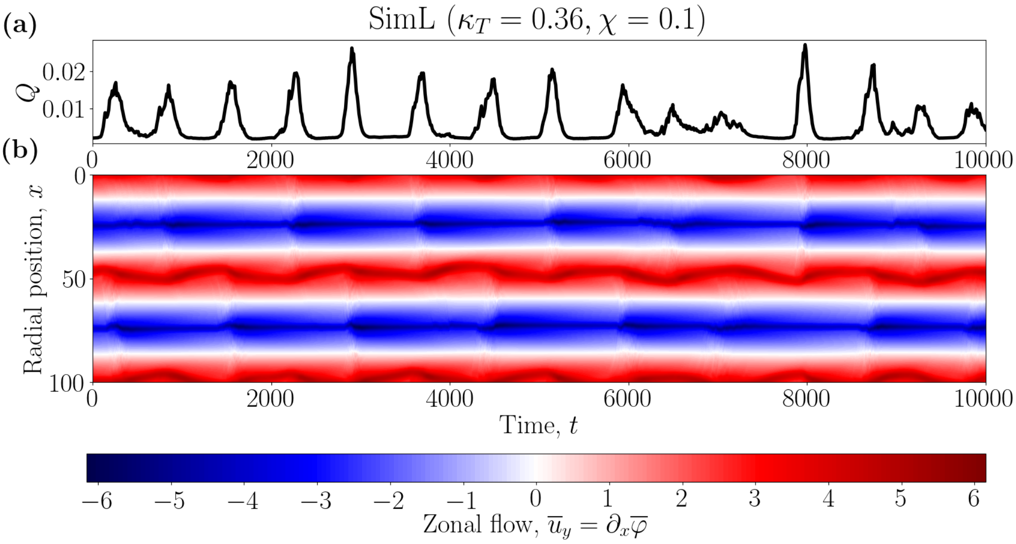


Figure 3.9: **(a)** Box-integrated radial heat flux Q and **(b)** ZF velocity ($\partial_x\overline{\varphi}$) vs. time for SimL ($\kappa_T = 0.36$, $\chi = 0.1$). Each turbulent burst is accompanied by an order-of-magnitude increase in Q and a radial oscillation of the locations of the ZF maxima.

turbulence enhances the ZFs that are doing the shearing through a negative turbulent viscosity. Beyond the Dimits threshold ($\kappa_T > \kappa_T^{\mathrm{c,2D}}$), the turbulent viscosity is positive and strong, ITG-suppressing ZFs cannot be maintained. These results are presented in §3.3, but first, in this section, we shall describe the saturated state near the Dimits threshold.

Figure 3.4 shows the heat flux Q vs. κ_T and χ. We have checked that all simulations have converged by inspection of their heat flux and ZF profiles, and by ensuring that they run for several box-scale diffusion times $t_{\mathrm{box}} \equiv (L_x/2\pi)^2/a\chi$. The turbulent heat flux Q depends strongly on the temperature gradient and increases monotonically with increasing κ_T. In contrast, its dependence on the collisionality is much weaker and non-monotonic (see figure 3.4b). Close to the Dimits threshold, Q decreases with increasing χ (which takes it away from the threshold), whereas farther away from the threshold, it increases and then plateaus with increasing χ. An increase of flux with collisionality for Z-pinch turbulence was noted by Ricci *et al.* (2006).

In what follows, a significant fraction of the detailed analysis is done using two simulations of the low-collisionality near-marginal state with parameters $\kappa_T = 0.36$, $\chi = 0.1$, $L_x = 100$, $L_y = 150$, one with higher (507×337) and one with lower (167×167) number of Fourier modes (the lower-resolution simulation is used for longer runs due to its lower computational cost). They have the same initial condition, taken from an already saturated simulation. Both the low- and high-resolution simulations show good convergence of their spectra (see figure 3.5). We shall refer to these two simulations as "SimL" and "SimH", respectively.

In the near-marginal[3] Dimits state, turbulence is suppressed by a quasi-static "zonal staircase" arrangement of the ZFs and zonal temperature perturbations. This structure is reminiscent of the "$\boldsymbol{E}\times\boldsymbol{B}$ staircase" observed in global GK simulations (Dif-Pradalier *et al.*, 2010, 2015, 2017; Villard *et al.*, 2013, 2014). The zonal staircase consists of interleaved regions of strong zonal shear that suppresses the ITG turbulence in those regions, and localised turbulent patches at the turning points of the ZF velocity. We

---

[3] The dependence of heat flux on the temperature gradient is very steep for fully developed ITG turbulence (see Barnes *et al.*, 2011) and so it is expected that the temperature gradient in a real device will be determined largely by the marginal gradient, i.e., the one above which strong turbulence develops.

shall refer to the former as the "shear zones" (see §3.2.1) and to the latter as the "convection zones" (see §3.2.2). A typical near-marginal ZF configuration can be seen in figure 3.6 and corresponding snapshots of the perturbed temperature in figure 3.7. Turbulence is always present, in a highly localised form, in the convection zones, but not in the shear zones.

The ZF in the staircase is not steady, but subject to viscous decay. In the low-collisionality ($\chi \lesssim 1$), near-marginal regime, this decay is slow and turbulent bursts are triggered periodically when the zonal shear in the shear zones has decayed to a level that is insufficient for the suppression of turbulence. These bursts lead to a significant (order-of-magnitude) increase in the radial heat flux. Similar bursts were reported by Kobayashi & Rogers (2012) in entropy-mode-driven Z-pinch turbulence. In our system, they are seeded by highly localised, coherent, turbulent structures, reminiscent of those reported by van Wyk *et al.* (2016) in gyrokinetic turbulence with an imposed equilibrium flow shear (see §3.2.3). A typical turbulent burst is illustrated in figures 3.6, 3.7 and 3.8 where we see the evolution of the quiescent state into a turbulent one and then back. Figure 3.9 shows a longer time evolution for the same parameters, illustrating the (quasi)periodic nature of the bursts. At higher collisionality, the ZFs decay faster, the turbulent bursts start to overlap, and it becomes difficult to isolate quiescent periods from turbulent ones. This state is more homogeneous in time and does not have well-defined oscillations, unlike the bursty state at low collisionality. However, we find that the mechanism that governs the stability of the Dimits state and the transition to strong turbulence is very similar for all values of collisionality that we have explored (see §3.3.2).

We now proceed to describe the features of the zonal staircase in more detail.

3.2.1 Shear zones

3.2.1.1 Suppression of turbulence

The zonal staircase is arranged in such a way that it efficiently suppresses turbulence in the shear zones via strong ZF shear. We find that this shear, $S \equiv \partial_x^2 \overline{\varphi}$, satisfies $S \gtrsim \gamma_{\max}$, where $\gamma_{\max} \equiv \max \operatorname{Im}(\omega_{\boldsymbol{k}})$ is the largest linear ITG growth rate determined from the dispersion relation (2.35). The notion that ITG turbulence requires comparable

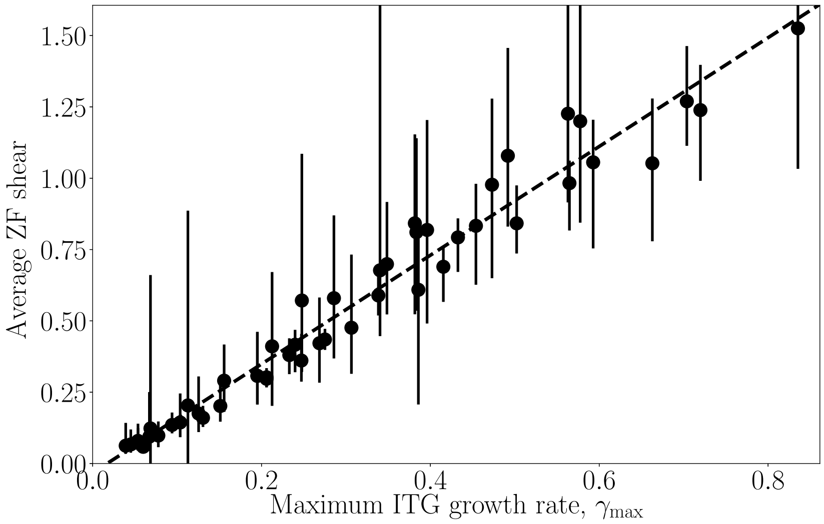


Figure 3.10: ZF shear S (time- and space-averaged over the shear zones) vs. maximum linear ITG growth rate $\gamma_{\max}$ of (2.35). The data is taken from a number of simulations over a range of parameters: $\kappa_T \in [0.16, 7.29]$ and $\chi \in [0.1, 10]$. The error bars represent the smallest and the largest values of the spatially averaged ZF shear for each simulation. The best-fit line (dashed) is $S \approx 2\gamma_{\max}$.

S and $\gamma_{\max}$ to be suppressed by shear is known as the "quench rule" (Waltz *et al.*, 1994, 1998; Kinsey *et al.*, 2005; Kobayashi & Rogers, 2012). Quantitatively, this is supported by figure 3.10, which shows that the time- and space-averaged (over the shear zones) shear satisfies $S \approx 2\gamma_{\max}$ over a range of simulation parameters.[4]Note that the *particular snapshots* of zonal profiles seen in figure 3.6 suggest $S \approx \gamma_{\max}$. However, the *time-averaged* S is larger due to the variation of shear over time (see also figure 3.11).

3.2.1.2 Decay of zonal flows

Let us describe the viscous decay of the ZFs. The equation for the evolution of the ZFs is given by the zonal part of (2.22):

$$\partial_t \partial_x^2 \overline{\varphi} = \partial_x^2 \overline{\partial_x \varphi \partial_y (\varphi + T)} + \chi \partial_x^4 \big(a\overline{\varphi} - b\overline{T} \big). \tag{3.19}$$

---

[4] The averaging is performed numerically over regions of near-uniform zonal shear, where, at every time step, we identify the radial locations of the uniform shear zones by applying the following conditions: $\partial_x^3 \overline{\varphi} < 0.1 \max\{\partial_x^3 \overline{\varphi}\}$ to isolate regions of near-uniform shear, and $\partial_x^2 \overline{\varphi} < 0.5 \max\{\partial_x^2 \overline{\varphi}\}$ to exclude the large variations of shear around the ZF extrema (see figure 3.6).

Integrating (3.19) once yields

$$\partial_t \overline{u}_y = \partial_x \Big[\overline{(\partial_x \varphi)\partial_y(\varphi + T)} + \chi \partial_x^2 \big(a\overline{\varphi} - b\overline{T} \big) \Big] = -\partial_x (\Pi_t + \Pi_d), \tag{3.20}$$

where the zonal flow velocity is $\overline{u}_y \equiv \partial_x \overline{\varphi}$ and we have identified the turbulent, $\Pi_t \equiv -\overline{(\partial_x \varphi)\partial_y(\varphi + T)}$, and diffusive, $\Pi_d \equiv -\chi \partial_x^2 \big(a\overline{\varphi} - b\overline{T} \big)$, radial fluxes of poloidal momentum. The integration constant in (3.20) is zero because both sides of the equation are exact derivatives with respect to x and our domain is periodic. Integrating (3.20) once more yields a term that is not necessarily an exact derivative — the turbulent momentum flux Π_t:

$$\partial_t \overline{\varphi} + \Pi_t + \Pi_d = \Pi, \tag{3.21}$$

where the integration constant $\Pi = (1/L_x) \int_0^{L_x} dx \, \Pi_t$ is the total box-averaged poloidal momentum flux. However, (2.22) and (2.23) are invariant under the symmetry

$$x \mapsto -x, \; y \mapsto y, \; \varphi \mapsto -\varphi, \; T \mapsto -T. \tag{3.22}$$

Under this symmetry, $\Pi \mapsto -\Pi$, a property of our model inherited from gyrokinetics (Parra *et al.*, 2011). Therefore, assuming that the time- and volume-averaged solutions to (2.22) and (2.23) respect (3.22), we conclude that, at least statistically, $\Pi = 0$. Indeed, our numerical simulations show that the right-hand side of (3.21) vanishes (up to the accuracy of the integration scheme).

During the quiescent periods of the Dimits-state evolution (i.e., between turbulent bursts), the turbulent momentum flux in the shear zones is negligible compared to the diffusive momentum flux, $\Pi_t \ll \Pi_d$. This is a consequence of the suppression of the $k_y \neq 0$ ITG modes by the zonal shear.[5] We also find that the zonal temperature gradient $\partial_x \overline{T}$ is approximately constant in the quiescent shear zones (see figure 3.6), so $\partial_x^2 \overline{T} = 0$. Therefore, (3.21) becomes

$$\partial_t \overline{\varphi} = a\chi \partial_x^2 \overline{\varphi}. \tag{3.23}$$

This is a diffusion equation governing the viscous decay of the ZFs with a collisional viscosity $a\chi$. As figure 3.11 shows, quiescent periods of low heat flux and, thus, low

---

[5] Note that Π_t is not small if there is turbulence present in the shear zones (which happens in the run up to and during turbulent bursts) — we shall investigate Π_t in §3.3.

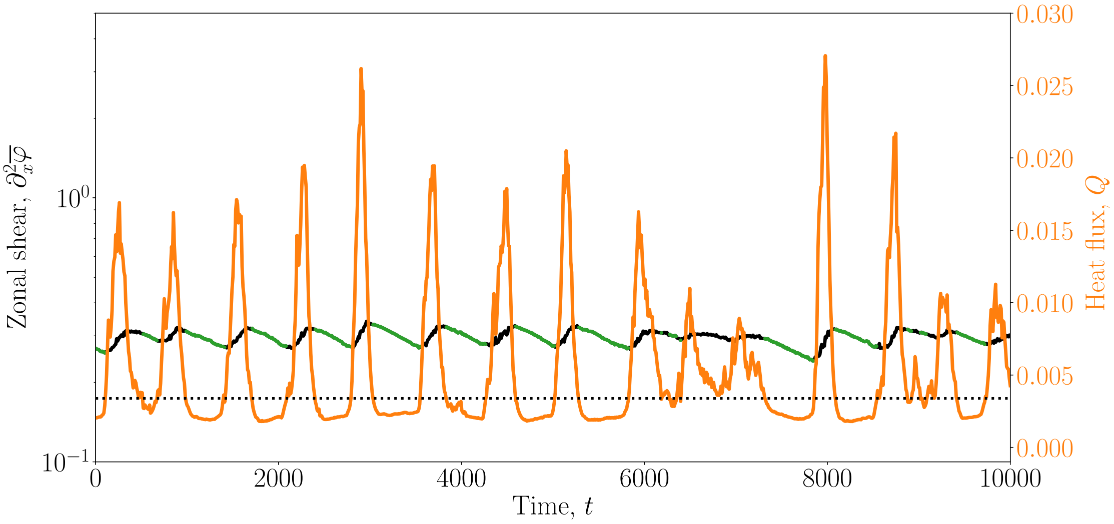


Figure 3.11: Heat flux Q (orange) and zonal shear $\partial_x^2\overline{\varphi}$ (black) vs. time for SimL. The highlighted (in green) sections of the zonal shear correspond to the quiescent periods. They are identified as those in which Q is smaller than a threshold value (dashed black line), defined as 60% of the time-averaged Q. The average decay rate of the zonal shear in the thus-identified quiescent periods is $\gamma \approx -3.5 \times 10^{-4}$ and the decay rate given by (3.24) is $\gamma_\mathrm{s} \approx -3.6 \times 10^{-4}$.

levels of nonzonal perturbations, are correlated with the periods of decay of the zonal shear. We find that, despite the ever-present turbulence in the convection zones, where (3.23) does not hold, the decay rate of the zonal shear is closely approximated by the viscous decay rate of the longest-wavelength ZF that comprises the zonal staircase, viz.,

$$\gamma_\mathrm{s} = -a\chi\left(\frac{2\pi n}{L_x}\right)^2, \tag{3.24}$$

where n is the number of periods of the zonal staircase in the domain of radial size L_x.

Let us now discuss what the ZF periodicity is.

3.2.1.3 Scale of zonal flows

In general, increasing/decreasing the radial extent of the integration domain by a factor increases/decreases the number n of shear zones by the same factor. This suggests that the characteristic length scale of the staircase, viz., the time-averaged radial separation of ZF extrema, is determined by a box-size-independent mechanism. Ascertaining

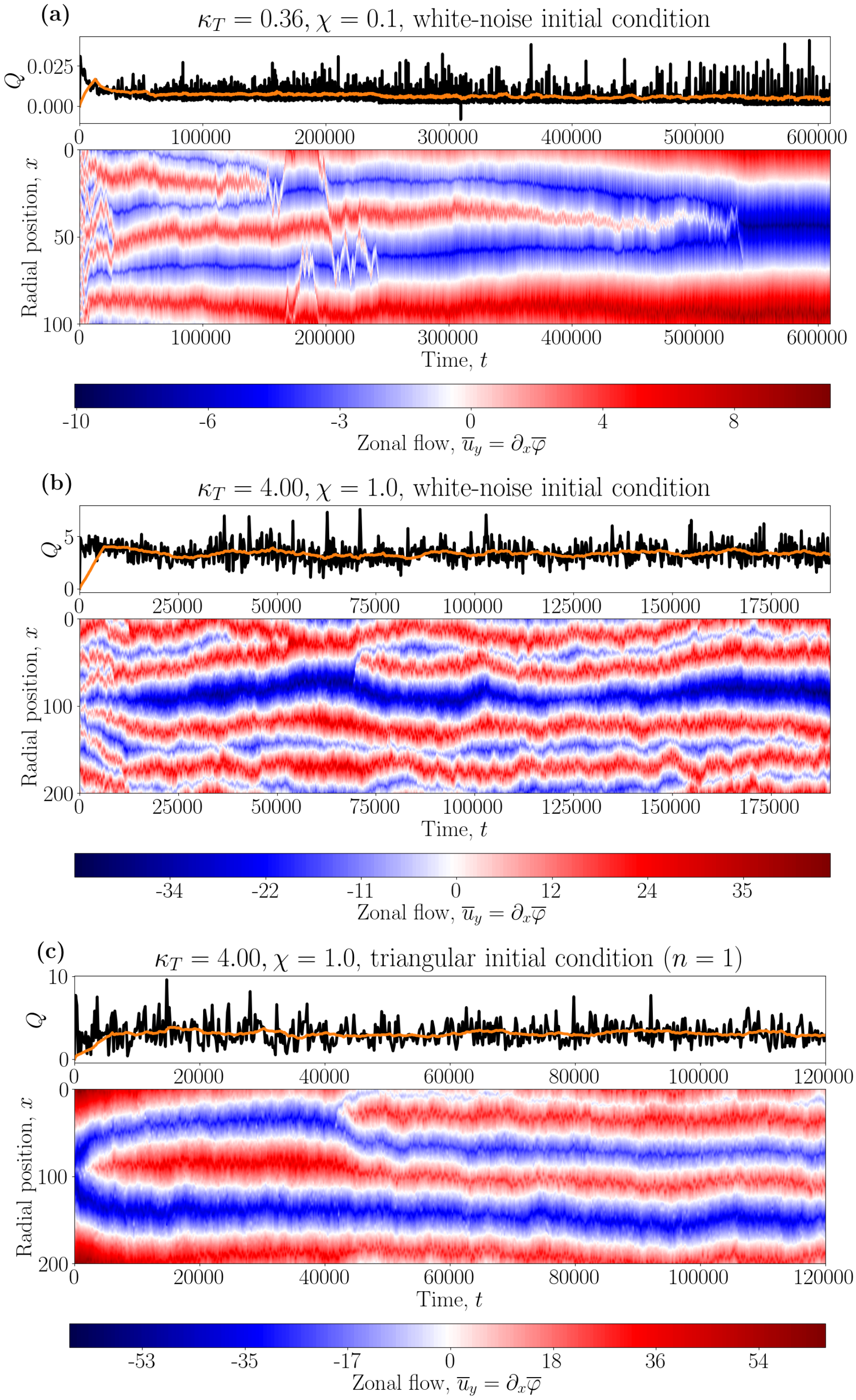


Figure 3.12: **(a)** Heat flux, raw (black) and rolling-averaged with a window of 6000 time units (orange) for $\kappa_T = 0.36$, $\chi = 0.1$, $L_x = 100$, $L_y = 100$. This is the same as SimL, but we have used small-amplitude white noise as an initial condition. The diffusion time for the box-scale ZF is $t_{\text{box}} \approx 11200$. Convergence to a box-sized ZF occurs on a very long timescale ($> 50 t_{\text{box}}$). **(b)** Same as (a), but for $\kappa_T = 4$, $\chi = 1$, $L_x = 200$, $L_y = 200$. Here $t_{\text{box}} \approx 4500$. **(c)** Same as (b), but with a single-peak triangular ZF as an initial condition.

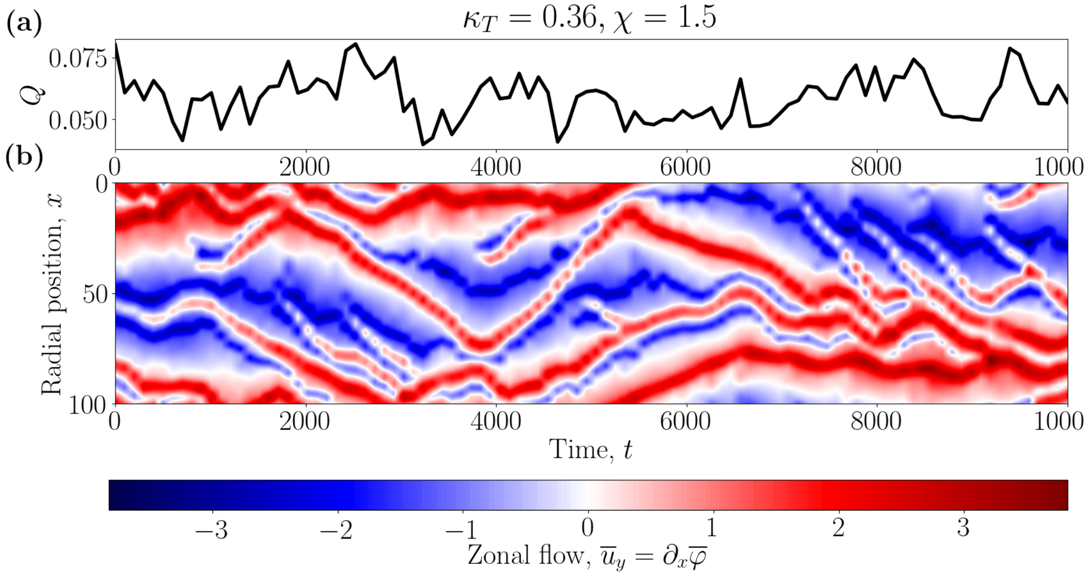


Figure 3.13: Same as figure 3.9, but for $\chi = 1.5$. The diffusion time for the box-scale ZF is $t_{\text{box}} = (L_x/2\pi)^2/a\chi \approx 750$. The locations of the ZF extrema of the staircase drift significantly over times comparable to t_{box}.

definitively whether this is the case is made difficult by the numerically observed time scales of convergence, which are at least an order of magnitude larger than the longest linear time scales, i.e., than the box-scale diffusion time $t_{\text{box}} = (L_x/2\pi)^2/a\chi$ (see figure 3.12). Note that the long-time evolution of the zonal profile and its length scale is not accompanied by a significant change in the average heat flux. In fact, the latter appears to reach saturation on time scales comparable to the box-scale diffusion time. Therefore, it is reasonable to trust the numerical values of the box-averaged heat flux (e.g., those shown in figure 3.4), even though we could not be certain that the zonal profiles have reached ultimate saturation.

As we increase collisionality and thus move away from the near-marginal regime and into the collisionality-independent regime (the plateau seen in figure 3.4b), the ZFs become more dynamic — they can merge, split and drift, as shown in figure 3.13b. Here we focus on the near-marginal regime and the transition to strong turbulence, so this higher-collisionality regime will not be studied.

Even though the zonal staircase arises naturally from white-noise initial conditions for both the zonal and the nonzonal fields, its shape suggests initialising the ZFs with

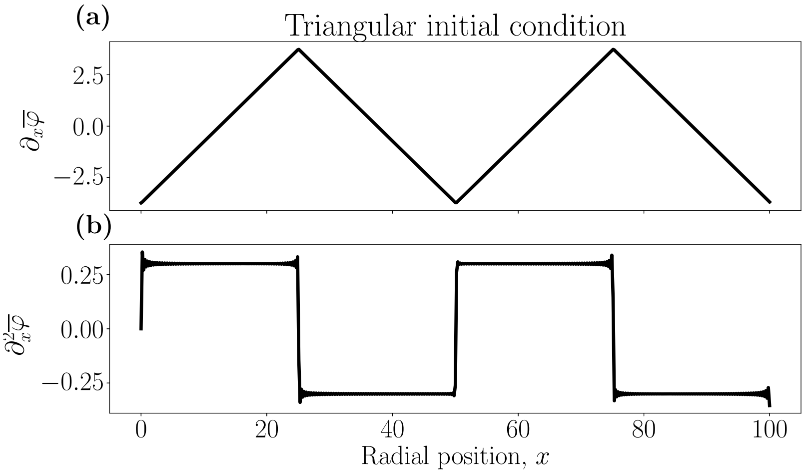


Figure 3.14: An example of the "triangular" initial condition used to initialise $\overline{\varphi}$ (see §3.2.1.3). Plotted are **(a)** the ZF ($\partial_x\overline{\varphi}$) and **(b)** the zonal shear ($S = \partial_x^2\overline{\varphi}$) for SimH parameters $L_x = 100$ and 507 radial Fourier modes. The value of the zonal shear $S = 0.3$ is chosen to be approximately equal to the time-averaged zonal shear in SimH and SimL (see also figures 3.6 and 3.11). The ringing artefacts (due to the finite number of Fourier modes) are viscously ironed out by the time the linearly unstable modes start growing and are not present in the saturated state. The exact form of $\overline{\varphi}$ is given in (3.25).

a "triangular" pattern. We find that this helps achieve more quickly a "less noisy" and more symmetric final state, which is easier to handle both numerically and analytically. Of course, we do not know in advance how many "steps" the staircase will "choose" to have in the saturated state, so their number for the "triangular" initial condition is just an informed guess. Most results in this chapter are from simulations that used such a triangular ZF initial condition, including SimL and SimH. Notable exceptions are figures 3.4, 3.10, and 3.22, where we used data from many simulations, some with white-noise initial conditions and others with "triangular" ones. Mathematically, the triangular initial condition is

$$\overline{\varphi} = \sum_{m\ \mathrm{odd}} \frac{4Si}{l_x}\left(\frac{l_x}{2\pi m}\right)^3 e^{2\pi i m x/l_x}, \tag{3.25}$$

where l_x is the desired ZF period (constrained by $L_x/l_x \in \mathbb{N}$) and S is the desired zonal shear in the shear zones. An example of the ZF profile (3.25) is given in figure 3.14.

3.2.2 Convection zones

The convection zones located at the extrema of the ZFs contain localised patches of ITG turbulence and have a high radial turbulent heat conductivity (see figures 3.6 and 3.7). The imposed equilibrium temperature gradient is flattened in the convection zones and slightly steepened in the shear zones. This results in a staircase-like radial temperature profile, shown in figure 3.15. The turbulence in the convection zones is driven by a tertiary instability, localised by the zonal shear. In the low-collisionality, near-marginal regime, which we consider to be the most important (see the footnote on page 28), there is a qualitative difference between the way in which the tertiary instability operates at the ZF maxima and minima. A similar difference exists in both the Hasegawa-Mima equation (Zhu *et al.*, 2018*b*) and gyrokinetics (McMillan *et al.*, 2011).

We shall now describe the features of the convection zones. The details of the tertiary instability are worked out in appendix B.

3.2.2.1 Turbulence at ZF minima

At the ZF minima, we find both ITG and Kelvin-Helmholtz tertiary instabilities. The former is dominant (faster) and saturates by producing a zonal temperature gradient that cancels the background temperature gradient. This effectively decouples the evolution of the temperature perturbations from that of the electrostatic potential and leaves a KH mode that seems to determine the poloidal wavenumber at the ZF minima (the peak at $k_y \approx 0.26$ in figure 3.5 is precisely the wavenumber of the fastest-growing KH mode at the ZF minima). Further details on the tertiary instability at the ZF minima and its saturation can be found in appendix B.3.1.

3.2.2.2 Turbulence at ZF maxima

In contrast to the ZF minima, the regions around the ZF maxima cannot support a Kelvin-Helmholtz instability because the Rayleigh-Kuo criterion for instability is not satisfied there (Kuo, 1949; Zhu *et al.*, 2018*a*): see appendix B.2. The ITG instability in these regions is significantly weaker than that at the ZF minima and does not appear to saturate in a similar fashion (by cancelling the equilibrium temperature gradient). The profile of the zonal temperature gradient shown in figure 3.6 suggests that the

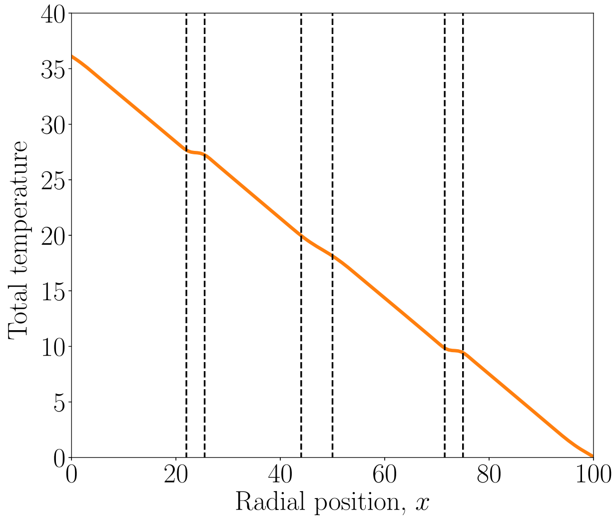


Figure 3.15: Time-averaged normalised total temperature (SimL), relative to the absolute temperature T_R at the right edge of the domain: $(\tau L_B/2\rho_s)\big(T_i - T_R + \overline{\delta T}\big) = (L_x - x)\kappa_T + \overline{T}$. A strong flattening of the gradient is visible around the ZF minima at $x \approx 25$ and 75, a weaker one around the ZF maximum at $x \approx 50$.

instability might not even be localised to the ZF maximum itself: there are two peaks of the zonal temperature gradient visible on either side of the ZF maximum at $x \approx 47$ at $t = 10$. The poloidal scale of the modes at the ZF maxima is significantly longer than that at the ZF minima, see appendix B.3.2.

Additionally, an asymmetric flattening of the zonal shear develops on one side of the ZF maximum, accompanied by a drift of the location of this maximum in the opposite direction (such a flattening is seen to the right of the central ZF maximum at $x \approx 47$ in figure 3.6). Eventually, ferdinons are launched in the direction of the flattening (see also §3.2.3). This is likely due to the inability of the diminished zonal shear there to suppress the nonzonal perturbations. The burst of ferdinons causes the ZF maxima to change the direction of its drift and a flattening of the zonal shear develops on the opposite side. This causes an oscillation of the position of the ZF maximum, as seen in figure 3.9b.

Thus, while turbulence is suppressed by zonal shear in the shear zones and by the cancellation of the equilibrum temperature gradient by the zonal temperature around the ZF minima, the regions around the ZF maxima remain locally unstable. As long

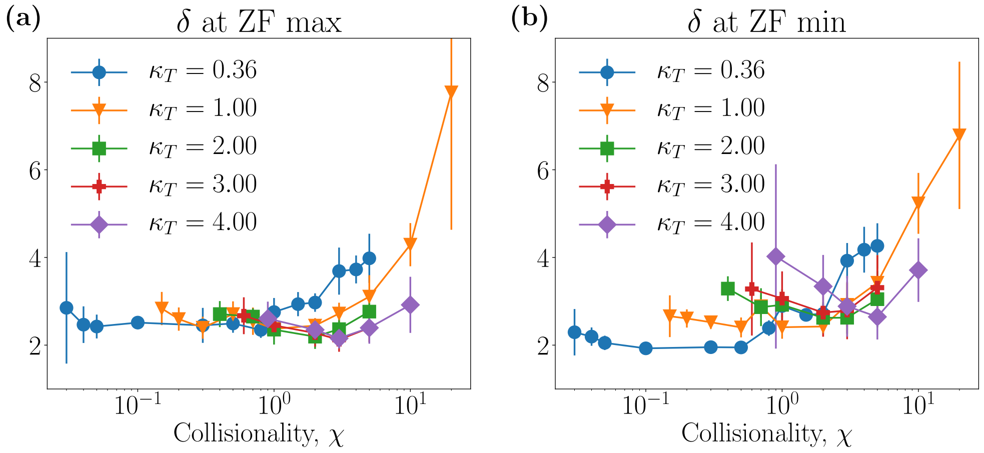


Figure 3.16: Numerically determined time-averaged values of δ, see (3.26), at **(a)** the global maximum and **(b)** global minimum (right) of the ZFs in the simulations whose heat flux is shown in figure 3.4. We find no significant variation of δ with χ or κ_T, except at large χ.

as the zonal shear in the shear zones is strong enough to suppress turbulence, this instability is tamed, with any perturbations launched from the unstable regions into the shear zones unable to survive. Once the zonal shear decays below a certain level, it is no longer able to suppress these perturbations (”ferdinons”, see §3.2.3) and a turbulent burst is initiated. Thus, the quasi-stationary zonal staircase contains the seeds of its own destruction: the perilous combination of decaying ZFs and unstable convections zones around the ZF maxima.

3.2.2.3 Scale of convection zones

The width of the convection zones can be characterised by the quantity

$$\delta \equiv \sqrt{\frac{\partial_x \overline{\varphi}}{\partial_x^3 \overline{\varphi}}}. \tag{3.26}$$

Figure 3.16 shows that δ does not depend very strongly on either κ_T or χ, except far from the marginal state, where collisionality appears to smooth out the gradients in the convection zones and thus increase δ. This suggests that near the Dimits threshold, δ is an $\mathcal{O}(1)$ quantity in the normalised units of (2.22) and (2.23), i.e., it is equal to a few times the sound radius ρ_s.

3.2.3 Ferdinons

After the ZF has decayed sufficiently to weaken its ability to suppress perturbations, vortex-like propagating structures are spawned from the ZF maxima and drift radially through the shear zones. Strikingly similar structures — "ferdinons" — have been observed in GK simulations with imposed background flow shear (van Wyk *et al.*, 2016, 2017). Figures 3.6b and 3.7b show a particular instance of the launching of ferdinons (see also figure 3.8).

As the ferdinons smash into the turbulent modes in the convection zones at the ZF minima, more structures are produced and a burst of turbulence ensues (see figures 3.6b and 3.7b). These travelling structures, as well as the resulting turbulence, cause a significant spike in the box-averaged heat flux (see figure 3.8). As figure 3.6b shows, they do not carry a significant ZF perturbation. They are created and propagate even if the ZF is held artificially constant in the numerical simulations, but the zonal temperature is left to evolve according to (2.23). In other words, a localised ZF perturbation is not an essential part of these structures.

Figure 3.17 shows that ferdinons consist of a vortex dipole and a strong temperature perturbation trapped in one of the vortices of this dipole. There are ferdinons carrying both positive ("hot") and negative ("cold") temperature perturbations. Hot ferdinons drift towards the cooler (right) side of the domain, while the cold ones drift in the opposite direction, towards the hotter (left) side (see also figure 3.8). Figures 3.17a and 3.17b demonstrate that the direction of the drift does not depend on the sign of the zonal shear. Net flow circulation around the ferdinons is also independent of the sign of the shear — it is always anticlockwise for hot and clockwise for cold ones (see figure 3.17c).

Note that the ferdinons that emerge in our simple ITG model bear a striking qualitative resemblance to the avalanches reported by Villard *et al.* (2013) in global GK simulations, namely, they propagate both inwards and outwards, but always with a positive heat flux, and originate from the local maxima of the ZF. Simple soliton-like solutions have already been proposed as a model for GK avalanches (McMillan *et al.*, 2009, 2018). Vortex-dipole structures called "modons" have been investigated in Hasegawa-Mima-like models of turbulence (Horton & Hasegawa, 1994). We do not yet know how and whether any of these are related to the ferdinons that we observe.

Let us discuss what we expect the ferdinon solution to be. Numerically, we find that the existence and propagation of these structures depend crucially on the two ITG-drive terms in (2.22) and (2.23), as well as on the nonlinear terms. In particular, the poloidal localisation of these structures is due to the nonzonal-nonzonal interactions. Indeed, we have found that (2.22) and (2.23) with the nonzonal-nonzonal nonlinear terms taken out (in what is sometimes referred to as the "quasilinear approximation"; see Srinivasan & Young, 2012) do not have ferdinon solutions. However, the quasilinear system does contain propagating nonlinear structures that are *not* localised poloidally, but rather appear to have a definite poloidal wavenumber k_y. These structures might be related to those described by McMillan *et al.* (2009) and Zhou *et al.* (2020). Models have been proposed for structure formation in a sheared flow that rely on the tilting of turbulence by shear and a nonzero group velocity to produce moving structures (McMillan *et al.*, 2018; Zhou *et al.*, 2020). The radial group velocity is (at least in the Hasegawa-Mima-related models) proportional to the product $k_x k_y$,[6] which acquires a definite sign in the presence of flow shear (§3.3.3). However, we observe ferdinons moving in either radial direction, regardless of the sign of the group velocity. Therefore, at the moment, we consider it unlikely that the propagation of ferdinons can be explained using such group-velocity arguments. We leave the detailed investigation of ferdinon generation and propagation for future work.

Understanding ferdinons and their properties can also put an upper bound on the radial scale of the ZF. Indeed, our numerical simulations show that the ZFs can have a well-defined radial scale smaller than the box size (§3.2.1.3). This scale could perhaps be estimated via a causality argument. The ZFs are generated locally (in the radial direction) by the nonlinear interactions of turbulent eddies. Therefore, assuming that ferdinons, and, thus, turbulence, can only propagate a finite radial distance in a region of self-consistently evolving zonal shear, an infinitely wide shear zone cannot be sustained for long because the ZF-driving turbulence will not be able to spread across the

---

[6] The radial group velocity $\partial\omega_{\boldsymbol{k}}/\partial k_x$ is proportional to $k_x k_y$ because $\omega_{\boldsymbol{k}} \propto k_y$ and $\omega_{\boldsymbol{k}}$ depends on k_x only through $k_\perp^2$.

entire shear zone. Note that finite-lifetime ferdinons over a dynamic ZF background, with which they can interact and gain or lose energy, are not in contradiction with the infinite-lifetime ferdinons seen by van Wyk *et al.* (2016), where a constant flow shear was imposed, and thus the shear profile was unable to react to the presence of ferdinons.

Once ferdinons are generated and turbulence develops in the shear zones, our analysis of the viscous decay of the zonal staircase in §3.2.1.2, which ignored the turbulent momentum flux, is no longer valid. Instead, we must focus on the effect of the turbulence on the ZFs. We find that the turbulence in the shear zones has a restoring effect on the zonal staircase in the Dimits regime, whereas beyond the Dimits threshold, it prevents staircase formation.

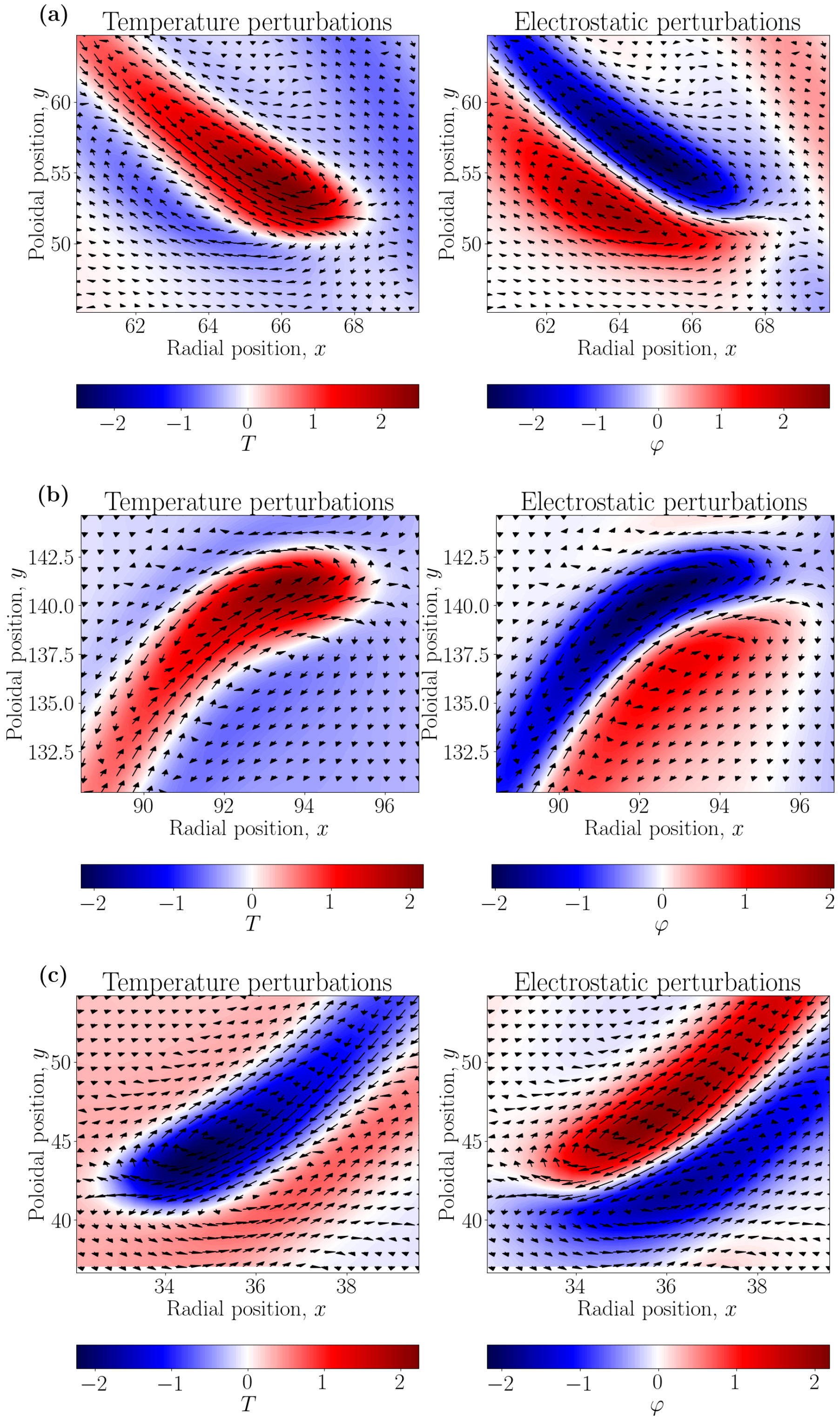


Figure 3.17: **(a)** A close-up view (from figure 3.7b) of the temperature and electrostatic-potential perturbations of a hot ferdinon in a region of negative zonal shear, $S < 0$. The arrows represent the local nonzonal $\boldsymbol{E} \times \boldsymbol{B}$ velocity (in arbitrary units). **(b)** A hot ferdinon in a region of positive zonal shear $S > 0$. **(c)** A cold ferdinon in a region of positive zonal shear $S > 0$.

3.3 Resilience of the zonal state and the Dimits threshold

3.3.1 Turbulent momentum flux

In order to investigate the way in which ZF profiles are formed and maintained during turbulent periods, let us ask the following question: does ITG turbulence in the shear zones (produced in bursts) have a definite effect on the ZFs, and is it to oppose or to feed them? We shall find that this sheared turbulence enhances the ZFs in the Dimits regime and destroys them beyond the Dimits transition.

The ZF evolution equation (3.21) is

$$\partial_t \overline{\varphi} + \Pi_t + \Pi_d = 0. \tag{3.27}$$

In §3.2.1.2, we discussed the effect of the diffusive momentum flux Π_d, viz., the viscous decay of the ZFs. For the rest of this section, we focus on the effects of turbulence by examining the turbulent momentum flux Π_t.

It is evident from (3.27) that ZF saturation requires

$$\langle \Pi_t(t,x) + \Pi_d(t,x) \rangle_{\Delta t} \approx 0 \tag{3.28}$$

to be satisfied at every radial location x, where $\langle f(t) \rangle_{\Delta t} \equiv (1/\Delta t) \int_{\Delta t} dt f(t)$ is a time average in the saturated state over a time Δt longer than the typical evolution time of the ZF (e.g., longer than the duration of turbulent bursts if the saturated state is bursty). Recall that $\Pi_d \approx -a\chi \partial_x^2 \overline{\varphi} = -a\chi S$ in the shear zones (see §3.2.1.2). Therefore, within a shear zone with nearly constant (in time *and* in space) zonal shear, (3.28) tells us that the time-averaged turbulent momentum flux in that shear zone must have a definite value,[7]determined by the local zonal shear. Thus, in the saturated state, Π_t will be correlated with S.

---

[7] Also, the spatial average over a shear zone of the turbulent momentum flux must be nonzero. This is not in contradiction with the argument in §3.2.1.2 that the spatial average over the entire box is zero, viz., $\Pi = 0$, because a definite uniform zonal shear breaks the symmetry (3.22) locally within each shear zone.

To quantify this correlation, let us multiply both sides of (3.27) by $S = \partial_x^2\overline{\varphi}$ and integrate across the radial extent of the domain. We find

$$\int_0^{L_x} dx\ \big(\partial_x^2\overline{\varphi}\big)\partial_t\overline{\varphi} + \int_0^{L_x} dx\ \Pi_t S + \int_0^{L_x} dx\ \Pi_d S = 0. \tag{3.29}$$

Since $\Pi_d = -\chi\partial_x^2(a\overline{\varphi} - b\overline{T}) = -a\chi S + b\chi\partial_x^2\overline{T}$, we have

$$\int_0^{L_x} dx\ \Pi_d S = -a\chi\int_0^{L_x} dx\ S^2 + b\chi\int_0^{L_x} dx\ S\partial_x^2\overline{T} \approx -a\chi\int_0^{L_x} dx\ S^2, \tag{3.30}$$

where we have assumed that the second term is negligible because the main contribution to S comes from the shear zones, where $\partial_x^2\overline{T} \approx 0$ (see §3.2.1.2). Therefore, after integrating by parts the first term in (3.29) and time averaging the resulting equation, we find

$$\left\langle\int_0^{L_x} dx\ \big(\Pi_t S - a\chi S^2\big)\right\rangle_{\Delta t} = -\left\langle\frac{1}{2}\partial_t\int_0^{L_x} dx\ \big(\partial_x\overline{\varphi}\big)^2\right\rangle_{\Delta t} \approx 0. \tag{3.31}$$

This gives a prediction for the effective "turbulent viscosity" in the shear zones:

$$\nu_t \equiv -\frac{\left\langle\int_0^{L_x} dx\ \Pi_t S\right\rangle_{\Delta t}}{\left\langle\int_0^{L_x} dx\ S^2\right\rangle_{\Delta t}} \approx -a\chi. \tag{3.32}$$

Relation (3.32) is, of course, corroborated by numerical simulations: see figure 3.18.

3.3.2 Sign reversal of the turbulent momentum flux at the Dimits threshold

An important consequence of (3.32) is that, in a shear zone, the sign of the turbulent momentum flux must coincide with the sign of the zonal shear. Therefore, if, for certain parameters, sheared turbulence has a momentum flux with a sign *opposing* that of the local shear, saturation mediated by a quasi-static zonal staircase cannot be achieved. We shall see that this is exactly what happens beyond the Dimits threshold.

Let us investigate how turbulence responds to an imposed static zonal profile. We solve (2.22) and (2.23) numerically with an imposed static triangular ZF pattern (i.e., we do not evolve the ZFs at all), in a box of size $L_x = L_y = 100$ and 169×169 Fourier modes, for a range of parameters around the Dimits transition. The chosen ZF pattern is shown in figure 3.19a and is adjusted for every simulation so that the value of the

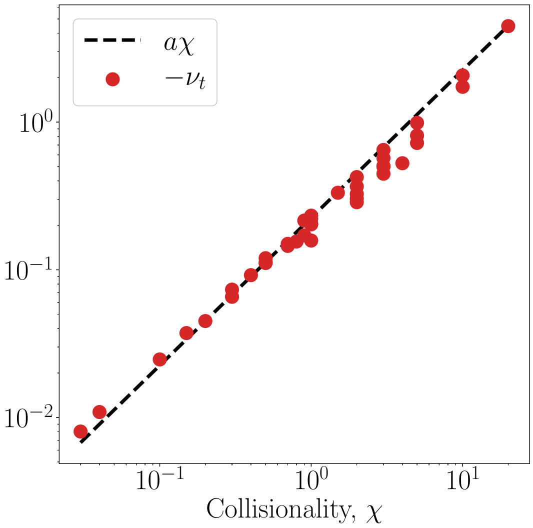


Figure 3.18: Comparison of estimated turbulent viscosity ν_t and the collisional viscosity $a\chi$ for the simulations from figure 3.4. We see that $\nu_t \approx -a\chi$.

zonal shear in the shear zones matches the largest ITG growth rate for that simulation. The chosen radial scale of the ZF (= 100) is inspired by the typical ZF scale that we observe in the low-collisionality regime, and is held fixed as we vary χ and κ_T. Then we calculate the effective turbulent viscosity ν_t associated with the turbulent momentum flux.

As figure 3.19b shows, we find a negative turbulent viscosity ν_t (and, thus, a positive correlation between local zonal shear and turbulent momentum flux) in the Dimits regime and a positive ν_t beyond it (thus, a negative correlation). Let us denote by κ_T^{static} the temperature gradient at which ν_t reverses its sign. The designation "*static*" reflects the fact that this is a numerical result for ITG turbulence with an artificially imposed *static* ZF profile. We find that the value of κ_T^{static} is insensitive to the exact shape of the ZF profile, and, most importantly, it nearly perfectly coincides with the Dimits threshold, i.e., $\kappa_T^{\text{static}} \approx \kappa_T^{\text{c,2D}}$.

Thus, in the Dimits regime, shear zones are resilient because, when the zonal shear there decays due to viscosity and turbulence is thus unleashed, this turbulence acts to reinforce the ZFs and the zonal shear in the shear zones is restored to its turbulence-suppressing level. Beyond the Dimits regime, the zonal staircase cannot be sustained because both turbulence and collisional viscosity act to flatten out the ZFs.

3.3.3 Reynolds stress and diamagnetic stress

Let us analyse what causes the turbulent momentum flux Π_t to reverse its sign at the Dimits transition. We split $\Pi_t = \Pi_\varphi + \Pi_T$ and define

$$\Pi_\varphi \equiv -\overline{(\partial_x\varphi)(\partial_y\varphi)} = \overline{u_y u_x}, \qquad \Pi_T \equiv -\overline{(\partial_x\varphi)(\partial_y T)} = \overline{u_y w_x}, \tag{3.33}$$

where $\boldsymbol{u} = (-\partial_y\varphi, \partial_x\varphi)$ is the $\boldsymbol{E}\times\boldsymbol{B}$ flow and $\boldsymbol{w} = (-\partial_y T, \partial_x T)$ is the diamagnetic flow. Here Π_φ is the radial flux of the poloidal momentum due to the Reynolds stress of the $\boldsymbol{E}\times\boldsymbol{B}$ flow and the "diamagnetic stress" Π_T is a contribution to the momentum flux that physically arises due to the advection of the poloidal diamagnetic flow $w_y = \partial_x T$ by the radial $\boldsymbol{E}\times\boldsymbol{B}$ flow $u_x = -\partial_y\varphi$.[8]To see this, let us take the zonal average of (2.23) and differentiate once with respect to x, to obtain an equation for the zonal diamagnetic flow:

$$\partial_t\overline{w}_y + \partial_x\overline{u_x w_y} - \partial_x\Pi_T = \chi\partial_x^2\overline{w}_y. \tag{3.34}$$

The evolution of the zonal poloidal $\boldsymbol{E}\times\boldsymbol{B}$ flow is described by (3.20), which can be recast as

$$\partial_t\overline{u}_y + \partial_x\overline{u_x u_y} + \partial_x\Pi_T = -\partial_x\Pi_d. \tag{3.35}$$

Added together, (3.34) and (3.35) describe the advection of the total poloidal flow ($\boldsymbol{E}\times\boldsymbol{B}$ + diamagnetic), $v_y \equiv u_y + w_y$, by the radial $\boldsymbol{E}\times\boldsymbol{B}$ flow:

$$\partial_t\overline{v}_y + \partial_x\overline{u_x v_y} = \text{dissipative terms.} \tag{3.36}$$

This makes physical sense because the diamagnetic flow is not a real flow and thus cannot advect anything.

The numerical solutions of (2.22) and (2.23) reveal that Π_φ and Π_T are in competition: on average, Π_φ has the same sign as the zonal shear $S = \partial_x^2\overline{\varphi}$, while Π_T has the opposite sign. This is evident in figure 3.20. Equation (3.27) then tells us that Π_φ feeds

---

[8] Similar terms in the momentum flux play an important role in the GK theory of momentum transport (Parra & Catto, 2009, 2010; Abiteboul, 2012; Calvo & Parra, 2015).

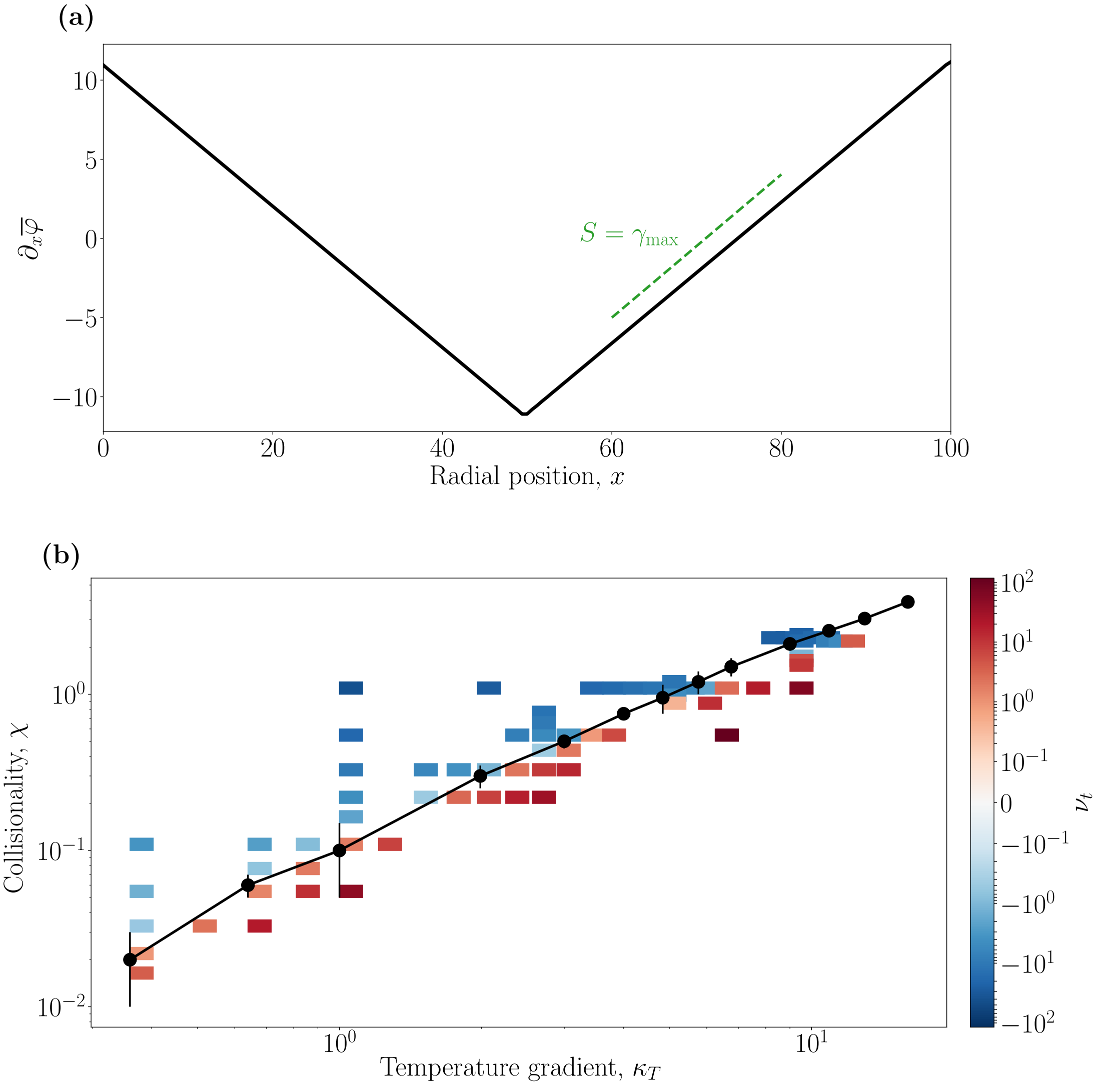


Figure 3.19: **(a)** The artificial triangular zonal pattern used to generate (b). The zonal shear in the shear zones is chosen to be equal in absolute value to the largest ITG growth rate (represented by the dashed green line). **(b)** The effective turbulent viscosity ν_t, as defined by (3.32), for the static triangular ZF profile given in (a) (coloured data points). The black line represents the numerically established Dimits threshold.

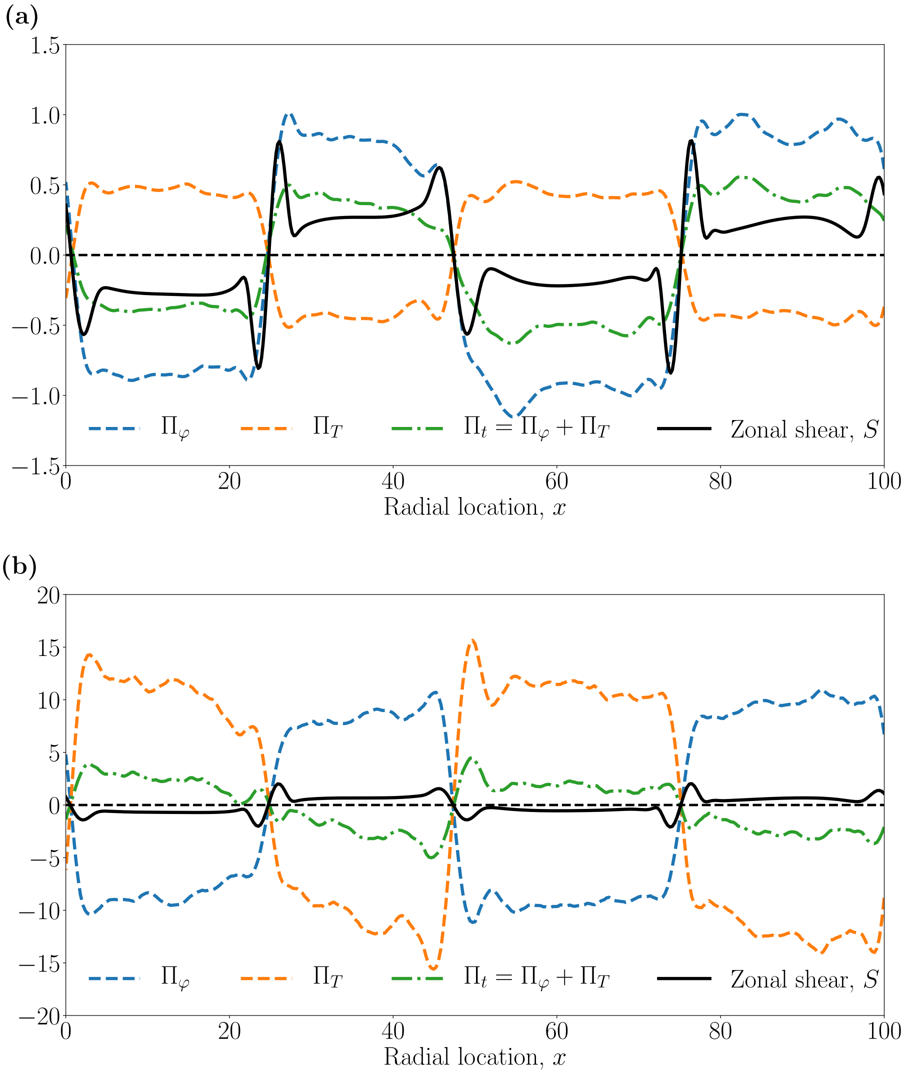


Figure 3.20: **(a)** Time-averaged momentum fluxes Π_φ, Π_T and $\Pi_t = \Pi_\varphi + \Pi_T$ for saturated ITG turbulence over a fixed zonal background. Note the correlation between the signs of the various fluxes and the zonal shear S: the sign of Π_t coincides with that of Π_φ and S and opposes the sign of Π_T. This reflects that the temperature gradient is lower than the Dimits threshold, $\kappa_T = 0.36 < \kappa_T^{c,\mathrm{2D}} \approx 1$. The ZF profile used here was extracted from SimH at $t = 10$, but reduced by a factor of 0.8 in order to allow ITG turbulence to develop in the shear zones. **(b)** Same as (a), but with $\kappa_T = 1.21$. The sign of Π_t now opposes the sign of Π_φ and S and coincides with the sign of Π_T. This reflects that $\kappa_T = 1.21 > \kappa_T^{c,\mathrm{2D}} \approx 1$. The extracted ZF is augmented by a factor of 2 to account for the increased ITG growth rate (due to the larger κ_T). This is necessary for the turbulence to saturate at numerically feasible amplitudes. Note that saturation is possible only because we have fixed the ZF profile. If the ZF is left to evolve according to (3.20), the poloidal momentum generated by the nonzonal perturbations flattens it and the system fails to reach a finite-amplitude saturated state (see §3.3.5).

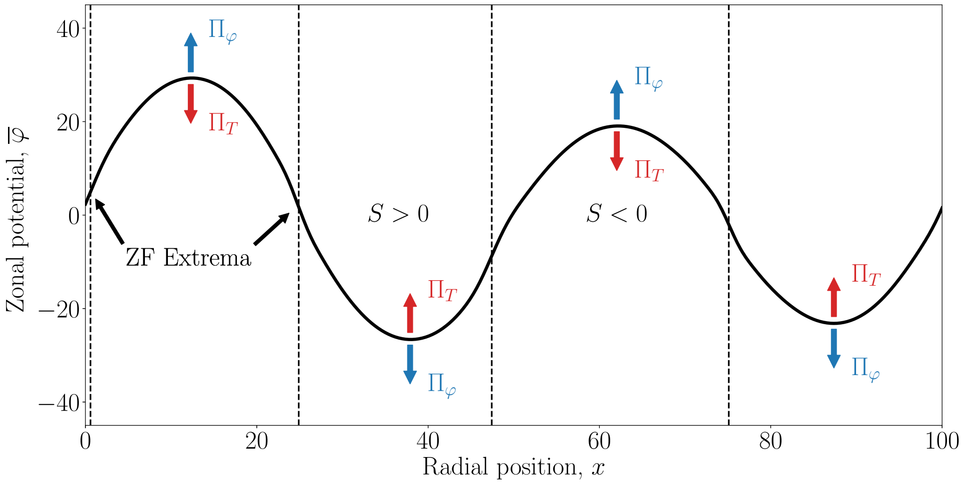


Figure 3.21: An illustration of the enhancing and suppressing effects of Π_φ and Π_T on the ZF. The black curve shows the $\overline{\varphi}(x)$ profile taken from SimH at $t = 10$. The ZF extrema are the locations where $S = \partial_x^2\overline{\varphi} = 0$ (marked by dashed lines).

the ZFs by increasing the zonal potential $\overline{\varphi}$ in the shear zones of negative zonal shear (where $\overline{\varphi}$ is concave) and decreasing it in the shear zones of positive zonal shear (where $\overline{\varphi}$ is convex), whereas Π_T relaxes the ZFs by opposing Π_φ. Their combined effect either steepens or relaxes the ZF velocity $\overline{u}_y = \partial_x\overline{\varphi}$ at the turning points $\partial_x^2\overline{\varphi} = 0$, depending on which stress is larger. Figure 3.21 is an illustration of this. This competition is crucial for the ability of the ZF to reconstitute itself after a turbulent burst and thus sets the threshold for the Dimits regime.

In order to assess what decides the outcome of this competition (i.e., the relative size of Π_φ and Π_T), let us consider how zonal shear affects ITG turbulence. For this purpose, consider a shear zone of radial extent d with a constant zonal shear $S = \partial_x^2\overline{\varphi}$ throughout it. We can then perform the usual shearing-box change of variables $(t, x, y) \mapsto (\tilde{t}, \tilde{x}, \tilde{y})$, where

$$\tilde{t} = t,\ \ \tilde{x} = x,\ \ \tilde{y} = y - Stx. \tag{3.37}$$

This coordinate transformation eliminates the spatially inhomogeneous zonal-advection terms $(\partial_x\overline{\varphi})\partial_y = Sx\partial_y$ in (2.22) and (2.23). Consider a Fourier mode $\varphi, T \propto \exp(i\tilde{k}_x\tilde{x} + i\tilde{k}_y\tilde{y})$ in this shearing frame. In the laboratory frame (t, x, y), this mode has the form

$\varphi, T \propto \exp(ik_x x + ik_y y)$, where

$$k_x = \tilde{k}_x - St\tilde{k}_y,\ k_y = \tilde{k}_y. \tag{3.38}$$

Thus, the ZF shear introduces an effective drift of the laboratory-frame radial wavenumber.[9] The direction of this drift is given by the sign of S, viz., $S > 0$ gives rise to an anticorrelation of k_x and k_y, i.e., $k_x k_y < 0$, whereas for $S < 0$, $k_x k_y > 0$. Integrating the effect of Π_φ over the sheared region, we obtain

$$\frac{1}{d}\int dx\ \Pi_\varphi = -\frac{1}{dL_y}\int dxdy\ (\partial_x\varphi)(\partial_y\varphi) = -\sum_{\boldsymbol{k}} k_x k_y |\varphi_{\boldsymbol{k}}|^2. \tag{3.39}$$

Therefore, on average, Π_φ has the same sign as S, and, thus, feeds the ZFs that generate the shear zones.[10]

We can write a similar expression for the diamagnetic stress:

$$\frac{1}{d}\int dx\ \Pi_T = -\sum_{\boldsymbol{k}} k_x k_y |\varphi_{\boldsymbol{k}}|^2 \mathrm{Re}\left(\frac{T_{\boldsymbol{k}}}{\varphi_{\boldsymbol{k}}}\right). \tag{3.40}$$

Then the total turbulent momentum flux integrated over a shear region is

$$\frac{1}{d}\int dx\ \Pi_t = -\sum_{\boldsymbol{k}} k_x k_y |\varphi_{\boldsymbol{k}}|^2 \left[1 + \mathrm{Re}\left(\frac{T_{\boldsymbol{k}}}{\varphi_{\boldsymbol{k}}}\right)\right]. \tag{3.41}$$

Recall that we already encountered the quantity $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})$ when dealing with the secondary instability in §3.1. There we found that $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) < 0$ for all linearly unstable modes. Thus, linear theory predicts that Π_T and Π_φ are anti-correlated due to the negative sign of $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})$. Now let us perform a more detailed analysis of the linear modes and attempt to construct a model for the Dimits threshold based on it.

3.3.4 Dimits threshold from linear physics

Using our knowledge of ITG perturbations in a region of uniform ZF shear, and of the turbulent momentum flux produced by them, we can make a heuristic linear-physics-based estimate for the Dimits threshold $\kappa_T^{c,\mathrm{2D}}$. In view of (3.41), it is given by the

---

[9] It is certainly true that an equilibrium shear would have such an effect on the turbulence. However, this is not guaranteed for ZFs. Their influence on the turbulence depends crucially on the modified electron response (2.4). This is a distinguishing feature of ion-scale physics that would not exist in, e.g., the electron-scale version of the model presented here.

[10] This is a well-known result in the context of Rossby-wave turbulence (see Vallis, 2017, chapter 15.1.2)

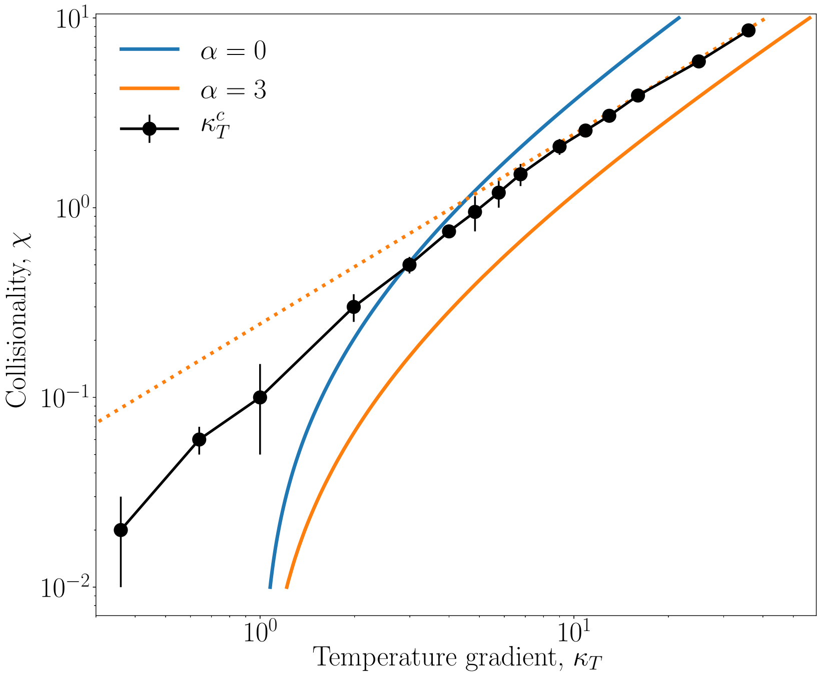


Figure 3.22: Comparison of numerical data to the analytical estimate for the threshold of the Dimits regime (§3.3.4). The black points represent the numerically observed $\kappa_T^{c,2\mathrm{D}}$. The other two curves correspond to the parameters for which the fastest-growing mode with $k_x = \alpha k_y$ satisfies $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) = -1$. The case $\alpha = 0$ corresponds to the threshold for the suppression of the secondary instability of the fastest streamer ($k_x = 0$) mode (see §3.1), and the curve with $\alpha = 3$ asymptotes to the numerically determined slope of the Dimits threshold as $\kappa_T \to \infty$. Its asymptote is represented here by the dotted line and is given by (3.49).

temperature gradient at which the *relevant* ITG modes have $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) = -1$. By *relevant* we mean those ITG modes that dominate the turbulence in the shear zones. It is tempting to assume that these modes would be the most unstable modes in the system. This, however, cannot be the case because the most unstable modes are the radial streamers with $k_x = 0$, but the zonal shear that we find is comparable in magnitude to the largest ITG growth rate ($S \sim \gamma_{\max}$), and, therefore, is bound to break these streamers. Following the discussion in §3.3.3, we may assume that the typical ITG modes in sheared turbulence satisfy $k_x \sim \alpha k_y$, where $\alpha \sim S\tau_{\mathrm{nl}}$ characterises how tilted the mode is, τ_{nl} being the nonlinear correlation time of the turbulence in the shear zones. If $\tau_{\mathrm{nl}}^{-1} \sim \gamma_{\max} \sim S$, then $\alpha \sim 1$.

Thus, we assume that the relevant modes are tilted with $k_x = \alpha k_y$, where $\alpha \sim 1$ is an unknown tilt parameter that depends on the structure of the turbulence. We then look for the temperature gradient κ_T at which the fastest-growing ITG mode with $k_x = \alpha k_y$ satisfies $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) = -1$. This yields a prediction for the Dimits threshold in the (κ_T, χ) plane that we refer to as the “fastest-mode approximation”. Note that there is no *a priori* reason to assume that α is itself not a function of κ_T.

3.3.4.1 High-collisionality limit

We can take the $\chi \to \infty$ limit of the fastest-mode approximation analytically. We formally order $\chi \sim \kappa_T$, as suggested by figure 3.22. We then use the dispersion relation (2.35) to find the growth rate $\mathrm{Im}(\omega_{\boldsymbol{k}})$ and real frequency $\mathrm{Re}(\omega_{\boldsymbol{k}})$ of the fastest mode with $k_x = \alpha k_y$. In §2.3.1.2, we showed that a mode of wavenumber $\boldsymbol{k}$ is unstable if and only if

$$\kappa_T k_y^2 > a\chi^2 k_\perp^6 = a\chi^2(1+\alpha^2)^3 k_y^6, \tag{3.42}$$

so all unstable modes with $k_x = \alpha k_y$, where $\alpha \sim \mathcal{O}(1)$, satisfy

$$k_y^4 < \frac{\kappa_T}{a\chi^2(1+\alpha^2)^3} \sim \mathcal{O}\big(\kappa_T^{-1}\big). \tag{3.43}$$

Similarly, using the results in §2.3.1.1 for the FLR bounds on the region of unstable wavenumbers, we find $k_y^4 \sim \mathcal{O}\big(\kappa_T^{-1}\big)$ in the limit $\kappa_T \to \infty$. Thus, both mechanisms that bound the region of instability (and hence restrict the largest ITG growth rate), lead to the same scaling for the unstable wavenumbers. Therefore, the wavenumber of the most unstable mode must also satisfy $k_y \sim \kappa_T^{-1/4}$. Applying the ordering $\chi \sim \kappa_T \gg 1$ and $k_x = \alpha k_y \sim \kappa_T^{-1/4}$ to (2.35), we find

$$A \sim \mathcal{O}\Big(\kappa_T^{1/2}\Big), \qquad B \sim \mathcal{O}(1), \qquad C \sim \mathcal{O}\Big(\kappa_T^{1/4}\Big), \qquad f \sim \mathcal{O}(1), \qquad g \sim \mathcal{O}(1). \tag{3.44}$$

The unstable solution of (2.35) is

$$-i\omega_{\boldsymbol{k}} = \frac{-A - B + iC + \sqrt{(A - B + iC)^2 + 4fAB - 4igAC}}{2}. \tag{3.45}$$

After expanding it using (3.44), we find

$$\mathrm{Im}(\omega_{\boldsymbol{k}}) \sim \mathcal{O}(1) \ll \mathrm{Re}(\omega_{\boldsymbol{k}}) \sim \mathcal{O}\Big(\kappa_T^{1/4}\Big) \ll \chi k_\perp^2 \sim \mathcal{O}(\sqrt{\kappa_T}), \tag{3.46}$$

$$\mathrm{Re}(\omega_{\boldsymbol{k}}) = -(1-g)C + \mathcal{O}\Big(\kappa_T^{-1/4}\Big) = -\kappa_T(1-b)(1+\alpha^2)k_y^3 + \mathcal{O}\Big(\kappa_T^{-1/4}\Big). \tag{3.47}$$

Therefore, (3.18) gives

$$\mathrm{Re}\left(\frac{T_{\boldsymbol{k}}}{\varphi_{\boldsymbol{k}}}\right) = \frac{k_y \kappa_T \mathrm{Re}(\omega_{\boldsymbol{k}})}{|-i\omega_{\boldsymbol{k}} + \chi k_\perp^2|^2} \approx \frac{k_y \kappa_T \mathrm{Re}(\omega_{\boldsymbol{k}})}{\chi^2 k_\perp^4} \approx -\frac{1-b}{1+\alpha^2}\frac{\kappa_T^2}{\chi^2}. \tag{3.48}$$

Thus, the large-temperature-gradient fastest-mode approximation of the Dimits threshold is a straight line in the (κ_T, χ) plane, given by

$$\kappa_T^{c,2\mathrm{D}} \approx \chi\sqrt{\frac{1+\alpha^2}{1-b}}, \tag{3.49}$$

a posteriori confirming the ordering $\chi \sim \kappa_T$. The numerically determined Dimits threshold is indeed close to a straight line. Fitting the slope of that line to (3.49) yields $\alpha \approx 3$. Comparison of the prediction for the Dimits threshold for this value of α, as well as $\alpha = 0$, which corresponds to the threshold for the secondary instability of a primary streamer (as discussed in §3.1), can be found in figure 3.22. The convergence is slow ($\propto \kappa_T^{-1/4}$), hence the sizeable discrepancy for the values of κ_T shown there, but we consider the asymptotic result to be sound. Note that a value $\alpha \sim \mathcal{O}(1)$ is consistent with the observed tilt of the nonzonal spectra in numerical simulations.

3.3.4.2 Low-collisionality limit

Using a calculation that is nearly identical to the one in §3.1.2, we can analytically take the limit $\chi \to 0$ of the fastest-mode approximation using the collisionless dispersion relation (2.37) and inserting it into (3.14). We obtain that $\kappa_T^{c,2\mathrm{D}} \to 1$ as $\chi \to 0$ for the fastest mode with $k_x = \alpha k_y$, regardless of the value of α. This is a weakness of our "fastest-mode approximation" because the numerical data suggests instead that $\kappa_T^{c,2\mathrm{D}} \to 0$ as $\chi \to 0$. Thus, the assumptions that we made above about the relevance of the fastest-growing modes appear to be inadequate at low collisionality.

To summarise, the assumption that the momentum flux and $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})$ are dominated by the most unstable mode with some tilt given by $\alpha = k_x/k_y$ allows us to predict the Dimits threshold at high collisionality, but fails at low collisionality. This partial success is likely due to the fact that $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})$ for the most unstable mode is independent of $\boldsymbol{k}$ for $\kappa_T \sim \chi \gg 1$, see (3.48). So, not only the most unstable, but in fact all modes in its vicinity will have the same asymptotic value of $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})$. On the other hand, the failure of these assumptions at low collisionality suggests that we cannot use linear theory to predict the threshold there, but must rather focus on the nonlinear structure of the ITG turbulence seeded by ferdinons during bursts.

3.3.5 Beyond the Dimits regime

Beyond the Dimits threshold ($\kappa_T > \kappa_T^{c,2\mathrm{D}}$), our 2D system fails to reach saturation on a scale smaller than the domain size — perturbations grow exponentially and the box-sized streamer ($k_x = 0, k_y = 2\pi/L_y$) eventually dominates the spectrum. Figures 3.23 and 3.24 show that the large-scale, coherent ZFs that comprise the zonal staircase are quickly destroyed and never reappear. This is consistent with the illustration in figure 3.21 and the discussion in §3.3.3. For $\kappa_T > \kappa_T^{c,2\mathrm{D}}$, if a shear zone of coherent zonal shear (like the ones we observe in the Dimits regime) were formed, the turbulent stress Π_t would flatten out the ZF profile. Any coherent zonal shear is thus the harbinger of its own demise due to the momentum flux of the tilted turbulent eddies. The nonzonal perturbations grow exponentially, and so do the ZFs, but the

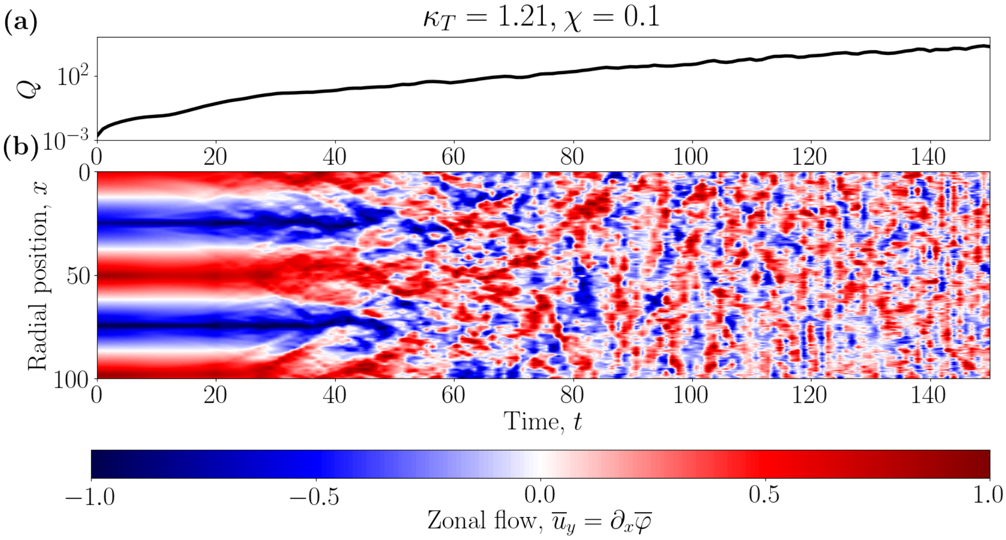


Figure 3.23: Time evolution of **(a)** the heat flux and **(b)** ZF velocity $\overline{u}_y$ beyond the Dimits threshold. The ZF amplitude in (b) is normalised to a maximum of 1 at each time. The initial conditions are the same as those for SimH, but with an augmented temperature gradient $\kappa_T = 1.21 > \kappa_T^{c,2\text{D}} \approx 1$ and a lower resolution of 337×167 Fourier modes due to the numerical cost of simulating the blow-up regime. The ZFs are quickly destroyed and large-scale ZFs never reappear, while the box-averaged heat flux Q grows exponentially. This state is eventually dominated by a streamer with a poloidal scale equal to that of the integration domain (see figure 3.24).

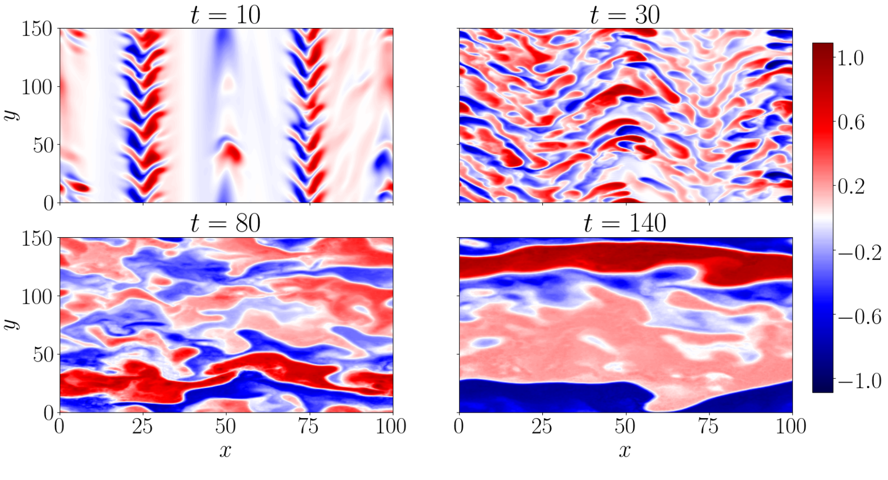


Figure 3.24: Snapshots of temperature perturbations in the blow-up state beyond the Dimits threshold. The amplitudes are normalised to a maximum of 1 at each time. The initial conditions are the same as those for SimH, but with an augmented temperature gradient $\kappa_T = 1.21 > \kappa_T^{c,2\text{D}} \approx 1$ and a lower resolution of 337×167 Fourier modes due to the numerical cost of simulating the blow-up regime.

latter are now dominated by small-scale time-incoherent zonal modes that are unable to quench the instability.

The lack of saturation beyond the Dimits regime in 2D is not surprising. GK simulations have shown that, beyond the Dimits regime in saturated 3D ITG turbulence, the ITG frequency at the injection (”outer”) scale of the perpendicular plane is balanced by the parallel propagation time — the turbulence is in “critical balance” (Barnes *et al.*, 2011):

$$\omega^{\mathrm{o}} \sim v_{\parallel} k_{\parallel}, \tag{3.50}$$

where $v_{\parallel}$ is some appropriate speed of parallel propagation (e.g., the ion thermal speed v_{ti}), and $\omega^{\mathrm{o}} \propto k_y$ is the ITG frequency at the outer scale, proportional to k_y by (2.35). In a tokamak, the smallest allowed value of $k_{\parallel}$ is $k_{\parallel} \sim L_{\parallel}^{-1}$, where $L_{\parallel}$ is the parallel connection length of the device. Thus, a parallel length scale is enforced by the magnetic geometry. The poloidal outer scale (k_y) then follows by (3.50) and the radial outer scale is enforced by zonal shearing ($k_x \sim k_y$). The 2D approximation can be obtained as the $k_{\parallel} \to 0$ limit of the 3D system. In this case, (3.50) implies that $k_y \to 0$, in agreement with the blow up dominated by the box-sized streamer that we observe beyond the Dimits threshold. Thus, the 2D approximation is fundamentally inadequate as a description of fully developed ITG turbulence.[11] In fact, in chapter 4, we shall see that the addition of 3D physics removes the blow up from the system, corroborating the hypothesis that it is caused by the restriction to 2D. However, we have shown that ZF-mediated saturation and the Dimits transition are captured by a 2D model. In fact, we shall see that the physics of turbulent momentum flux that generates the 2D Dimits state is also responsible for the 3D Dimits regime.

---

[11] Also because of the presence of 2D invariants (see §2.2), which can lead to an inverse cascade and energy pile-up at the largest available (box) scale, as they do in 2D hydrodynamic turbulence (Frisch, 1995).

Chapter 4

Going 3D: parasitic instabilities and strong turbulence

We now proceed to study the nonlinear saturated state of (2.18)–(2.20). We solve these equations using an enhanced version of the code we used in chapter 3, whereby (2.18)–(2.20) are solved using a pseudo-spectral algorithm in a triply periodic box of dimensions L_x, L_y, and $L_\parallel$. The linear terms (including the parallel streaming) are integrated implicitly in time, while the nonlinear terms are integrated explicitly using the Adams–Bashforth three-step method. As the 3D model has no dissipation terms that depend on $k_\parallel$,[1] we usually include small (compared to the collisional dissipation) parallel hyperviscosity of the form $\nu k_\parallel^4$. It is incorporated in the equations by replacing $\partial_t \mapsto \partial_t + \nu k_\parallel^4$ for all three fields in the model. The value of ν is typically chosen to give a maximum parallel hyperviscosity of 10% of $\chi k_{\perp,\text{largest}}^2$, where $k_{\perp,\text{largest}}$ is the largest $k_\perp$ included the simulation. This form of hyperviscosity effectively subtracts $\nu k_\parallel^4$ from the growth rate of every mode, but does not alter the linear mode structure, i.e., it does not influence the ratio of Reynolds to diamagnetic stresses given by $\text{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})$. Thus, it dissipates energy without perturbing the saturated state either towards or away from the Dimits regime.

Recall that the 2D model has two distinct nonlinear states — a Dimits regime, where saturation is achieved with the aid of strong ZFs that quench the curvature-ITG instability by shearing the perturbations it produces, and a blow-up regime, where no finite-amplitude saturation is achieved, but amplitudes continue to grow exponentially

---

[1] The collisional parallel diffusion is small in the limit (2.8).

indefinitely (or at least until numerical efforts become futile). This unphysical blow-up is arguably the main limitation of the 2D model, and there are good reasons to believe that it is a consequence of the $k_\parallel = 0$ restriction (see §3.3.5). This will indeed be corroborated below as we find that the 3D model is able to saturate for all values of κ_T and χ that we have investigated numerically.

The low-collisionality Dimits regime (which can be argued to be the most relevant: see the footnote on page 28) of the 3D model is strikingly similar to its 2D counterpart. The saturated state is dominated by quasi-static triangular ZFs that break up the radial domain into regions (shear zones) of constant zonal shear, where turbulence is sheared and thus suppressed (see figures 4.1a–c). Just as in 2D (see figure 3.7), localised patches of turbulence remain present at the turning points of the ZFs, where the zonal shear vanishes.

Periodically, when viscosity has eroded the ZFs and their ability to suppress turbulence has diminished, turbulent bursts are triggered. Just as in 2D, these bursts are foreshadowed by an instability located at the ZF maxima and by the appearance of localised travelling structures (”ferdinons”) produced by this instability. An example of a turbulent burst in the 3D model can be seen in figure 4.2. It is visually indistinguishable from a burst in 2D when viewed as a cross section in the (x, y) plane. We discuss the 3D structure of the Dimits regime in detail in §4.1.

The crucial qualitative change in physics that allowing 3D perturbations brings about is the slab-ITG instability. It is the peculiar nature of the cold-ion 3D slab-ITG instability that allows it to exist at arbitrarily small perpendicular scales $k_\perp \gg 1$ (for zero collisionality), unlike the 2D curvature-ITG instability which is FLR-limited to $k_\perp \lesssim 1$. Recall that in §2.3.2.2, we showed that the collisionless slab-ITG instability extends to infinitely large $k_\parallel$ and $k_\perp$, provided $k_\parallel \approx \kappa_T k_\perp k_y$. This is clearly unphysical as collisional dissipation increases with wavenumber and is bound to suppress the finite slab-ITG growth rate $\mathrm{Im}(\omega_{\boldsymbol{k}}) \sim \kappa_T$. The collisional terms in (2.19) and (2.20) scale as $\sim \chi k_\perp^2$, while that in (2.18) scales as $\chi k_\perp^4/(1 + k_\perp^2)$. Thus, for small-scale modes with $k_\perp \gg 1$, the dissipation rates in all three equations have the same scaling, viz., $\chi k_\perp^2$, and so we can make a simple estimate for the collisional cut-off of the instability $k^2_{\perp,\text{cut-off}}$ by balancing the maximal growth rate with the dissipation rate:

$\mathrm{Im}(\omega_{\boldsymbol{k}}) \sim \kappa_T \sim \chi k^2_{\perp,\text{cut-off}}$. Therefore, increasing κ_T (or decreasing χ) pushes the instability to smaller perpendicular and parallel scales, viz.,

$$k_{\perp,\text{cut-off}} \sim \sqrt{\frac{\kappa_T}{\chi}}. \tag{4.1}$$

This is in stark contrast with the behaviour of the 2D curvature-ITG modes whose cut-off wavenumber scales as $k_{\perp,\text{2D cut-off}} \sim \kappa_T^{-1/4}$ (see §2.3.1). Moreover, the maximal growth rate of the slab-ITG modes scales as $\mathrm{Im}(\omega_{\boldsymbol{k}}) \sim \kappa_T$, while that of the curvature-ITG modes satisfies $\mathrm{Im}(\omega_{\boldsymbol{k},\text{2D}}) \sim \kappa_T^{1/4}$. This implies that there is a natural scale separation between slow, large-scale curvature-driven modes and fast, small-scale slab-ITG modes. Crucially, this scale separation allows small-scale turbulence to be driven both by the equilibrium gradients and by the gradients associated with the large-scale 2D modes (which are themselves generated by the curvature-ITG instability). In fact, as we shall see in §4.2, the latter type of driving dominates in the saturated state to such an extent that the equilibrium temperature gradient can be turned off for the $k_\parallel \neq 0$ mode and the saturated state remains largely unchanged. In other words, the slab-ITG modes are parasitic modes, a type of 3D secondary instability of the 2D curvature-driven ITG modes.

Crucially for the Dimits state, the small-scale instability can be shown always to favour strong, coherent ZFs. It does so in two ways: by providing an effective positive thermal diffusion of the large-scale modes that would otherwise destabilise the ZFs in 2D, and by generating momentum transport that is beneficial for the ZFs (i.e., a negative turbulent viscosity for the zonal flow). This makes the 3D Dimits state much more resilient than the 2D one. In fact, we find that the 3D system enters a Dimits state regardless of the values of the parameters κ_T and χ, provided the domain is "sufficiently" 3D, i.e., provided $L_\parallel$ is large enough and that our numerical simulations have sufficient parallel resolution to resolve the slab-ITG modes.

We now recap the physical mechanism that gives rise to the Dimits regime and also discuss any qualitative and quantitative changes that 3D physics brings about. Then, in §4.2, we turn to the small-scale slab-ITG instability and its consequences for the saturated state. Finally, we examine the circumstances that can prevent the system from establishing a Dimits state and force it into the strongly turbulent regime.

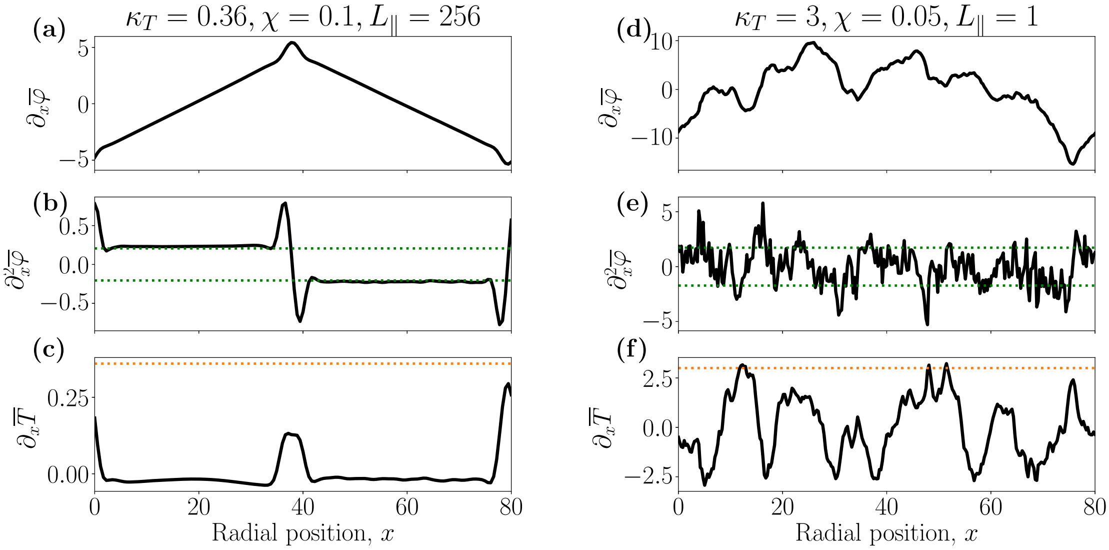


Figure 4.1: Example of instantaneous radial profiles in the Dimits state: **(a)** ZF, **(b)** zonal shear, **(c)** zonal temperature gradient. Example of instantaneous radial profiles in strong turbulence: **(d)** ZF, **(e)** zonal shear, **(f)** zonal temperature gradient. The dotted green lines in (b) and (e) are the largest linear growth rates for the respective simulations. The dotted orange lines in (c) and (f) are the values of the equilibrium temperature gradient κ_T. Just as in 2D, the zonal shear in the Dimits state is determined by the largest linear growth rate. Strongly turbulent ZFs do not produce regions of coherent shear.

4.1 Dimits regime

Recall that the 2D Dimits transition is a sharp transition from a finite-amplitude saturated state with strong ZFs to a "blow-up" state dominated by ever-growing streamers. The key to understanding this is the equation for the zonal electrostatic potential

$$\partial_t \overline{\varphi} + \Pi_\varphi + \Pi_T + \Pi_d = 0, \tag{4.2}$$

where

$$\Pi_\varphi \equiv -\overline{(\partial_x \varphi)(\partial_y \varphi)}, \qquad \Pi_T \equiv -\overline{(\partial_x \varphi)(\partial_y T)}, \qquad \Pi_d \equiv -\chi \partial_x^2 \big(a\overline{\varphi} - b\overline{T}\big) \tag{4.3}$$

are the Reynolds, diamagnetic, and diffusive stresses, respectively. Equation (4.2) describes how the ZFs are generated or eroded by turbulence (via the Reynolds and diamagnetic stresses, depending on their sign) and damped by collisional viscosity. The total turbulent stress $\Pi_t = \Pi_\varphi + \Pi_T$ in (4.2) can also be written as

$$\Pi_t = \Pi_\varphi + \Pi_T = -\sum_{\boldsymbol{k}} k_x k_y |\varphi_{\boldsymbol{k}}|^2 \left[1 + \mathrm{Re}\left(\frac{T_{\boldsymbol{k}}}{\varphi_{\boldsymbol{k}}}\right)\right]. \tag{4.4}$$

Thus, the effect of the mode with wavenumber $\boldsymbol{k}$ on the ZFs depends on the ratio $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})$. Namely, $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) < -1$ implies that the mode will destabilise the ZFs, while $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) > -1$ means that the mode will reinforce the ZFs. This observation is based on the fact that sheared (by the ZFs) turbulence is "tilted" and the sign of $k_x k_y$ is correlated with the sign of the zonal shear in each shear zone. In §3.3.4, we derived a simple estimate for the Dimits threshold at large κ_T that was based on applying these ideas to the linear modes of the 2D system. More generally, we argued that the Dimits transition occurred at the threshold of a nonlinear version of the secondary instability — when sheared by ZFs, turbulence either reinforced these flows and thus a Dimits state was established (the Reynolds stress won), or it failed to do so (the diamagnetic stress won) and saturation had to be reached via a different route that did not rely on zonal shear. In the 2D case, no such alternative route for finite-amplitude saturation exists.

4.1.1 The influence of $L_\parallel$ on the Dimits state

Taking the limit $L_\parallel \to 0$ effectively restricts our model equations (2.18)–(2.20) to 2D, and thus their saturated Dimits state converges to that of the 2D model (2.22) and (2.23). As we increase $L_\parallel$ away from 0, we see elongated nonzonal modes that eventually lose the ability to stay coherent along the field lines if $L_\parallel$ is large enough. Figure 4.3 shows that the typical Dimits-state ferdinons are not true 2D structures and develop a finite parallel extent if the parallel size of the box allows it. However, ferdinons can exist both in 2D and 3D, so we deem their parallel structure nonessential. This is consistent with the work of van Wyk *et al.* (2016, 2017), where the ferdinons were found to be very elongated along the field lines, but still had some structure along the field line. In figures 4.4a and 4.4b, we show the dependence of the turbulent heat flux Q on $L_\parallel$ far below the 2D Dimits threshold ($\kappa_T = 0.36$, $\chi = 0.1$), where turbulent bursts dominate the 2D state, and closer to it ($\kappa_T = 0.8$, $\chi = 0.1$), where the bursts start to overlap in time. As expected, if $L_\parallel$ is small enough, we recover the 2D results. As $L_\parallel$ increases, Q converges in a monotonic way to a definite 3D value that is smaller that the 2D heat flux. Figures 4.4c–f show that, for larger values of $L_\parallel$, the turbulent bursts become more frequent, but shorter in duration and lower in amplitude. There are two effects responsible for this — parallel localisation of turbulence and the development of the parasitic small-scale slab-ITG modes.

Parallel localisation is inevitable because the turbulent nonzonal modes cannot propagate information infinitely quickly along the field lines. This is in contrast with the ZFs which do stay perfectly coherent along the entire domain regardless of $L_\parallel$. The ability of the ZFs to communicate infinitely fast along the field lines is a consequence of the asymptotic limit of small mass ratio. In this limit, the electrons are able to stream along the magnetic field on a time scale $k_\parallel v_{te}$ that is much shorter than the typical ITG time scale. Here $v_{te} = \sqrt{2T_e/m_e}$ is the electron thermal speed, which satisfies $v_{te}/v_{ti} \sim \sqrt{m_i/m_e} \gg 1$ in the small-mass-ratio limit. This fast streaming leads to the modified adiabatic electron response (2.4). This puts the ZFs at an unfair advantage because the turbulent stresses in (4.2) are parallel averages and so a turbulent burst that is localised to a fraction $\Delta L_\parallel / L_\parallel$ of the parallel extent of the box has its turbulent

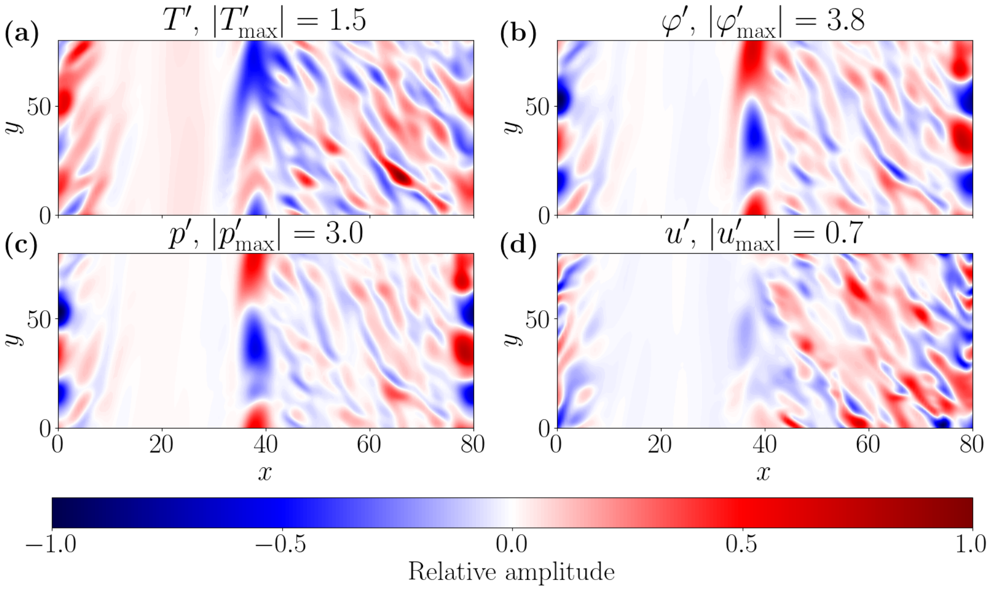
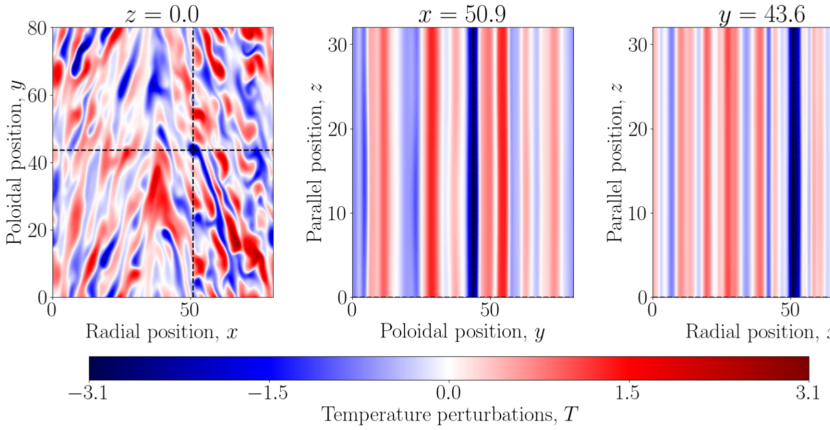


Figure 4.2: Snapshots of the nonzonal **(a)** temperature T', **(b)** potential φ', **(c)** pressure $p' = \varphi' + T'$, and **(d)** parallel velocity u'. The colour scale is relative to the maximum absolute amplitude in each panel (given in the panels' titles). We see that Dimits-state turbulent eddies have u perturbations, as well as T and φ perturbations.

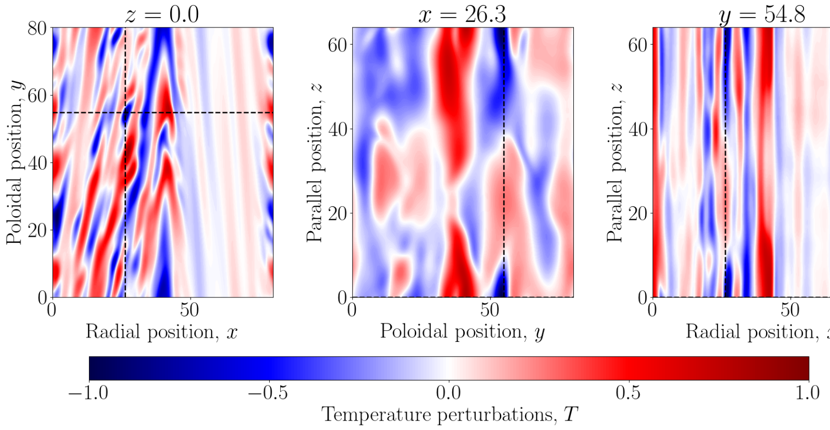
stress diminished by a factor of $\Delta L_\parallel / L_\parallel$. As we increase $L_\parallel$, every such localised burst provides a smaller "kick" to the ZFs and so it takes less time for the ZFs to decay to a level that permits the development of a new burst. Note that the turbulent heat flux Q is also a spatial average of the turbulent fields and it too is diminished for a localised burst. Thus, we expect smaller, more frequent bursts, and this is precisely what is observed.

Additionally, we find small-scale slab-ITG modes that feed off the perpendicular temperature gradients associated with the ferdinons. The presence of this three-dimensional small-scale parasitic instability can be detected via the parallel velocity u, because the latter is only involved in the 3D slab-ITG modes and not in the 2D curvature-ITG modes. Figure 4.5 shows an example of a ferdinon that is "infected" with such small-scale slab-ITG instabilities. As we shall discuss in §4.2, the small-scale instability leads to an effective increase in thermal diffusion, and thus an increase in the effective damping at large scales that reduces the large-scale temperature perturbations. This additional damping likely contributes to the reduced Q of the 3D Dimits state.

(a) $L_\parallel = 32$



(b) $L_\parallel = 64$



(c) $L_\parallel = 256$

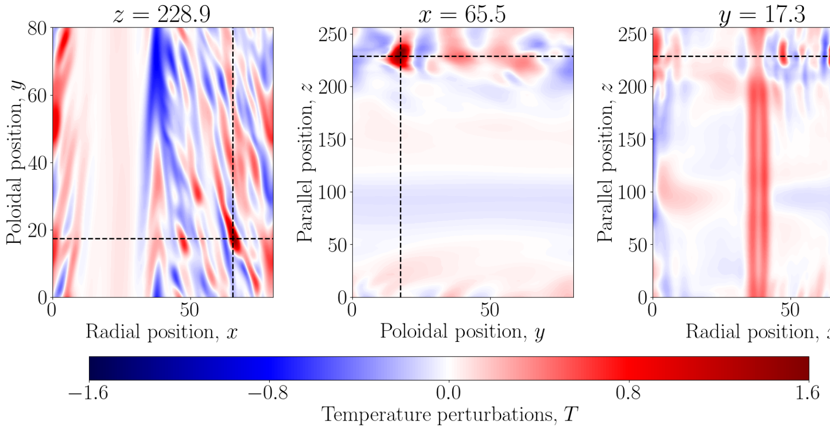


Figure 4.3: Snapshots of the 3D temperature perturbations associated with a ferdinon for $\kappa_T = 0.36$, $\chi = 0.1$. The plots on each row are cross sections in different planes from the same simulation at the same t for **(a)** $L_\parallel = 32$, **(b)** $L_\parallel = 64$, and **(c)** $L_\parallel = 256$. The black dashed lines visualise the intersections of the cross-sectional planes. As we increase $L_\parallel$, turbulence loses the ability to stay coherent along the parallel extent of the box and the bursts become localised in z.

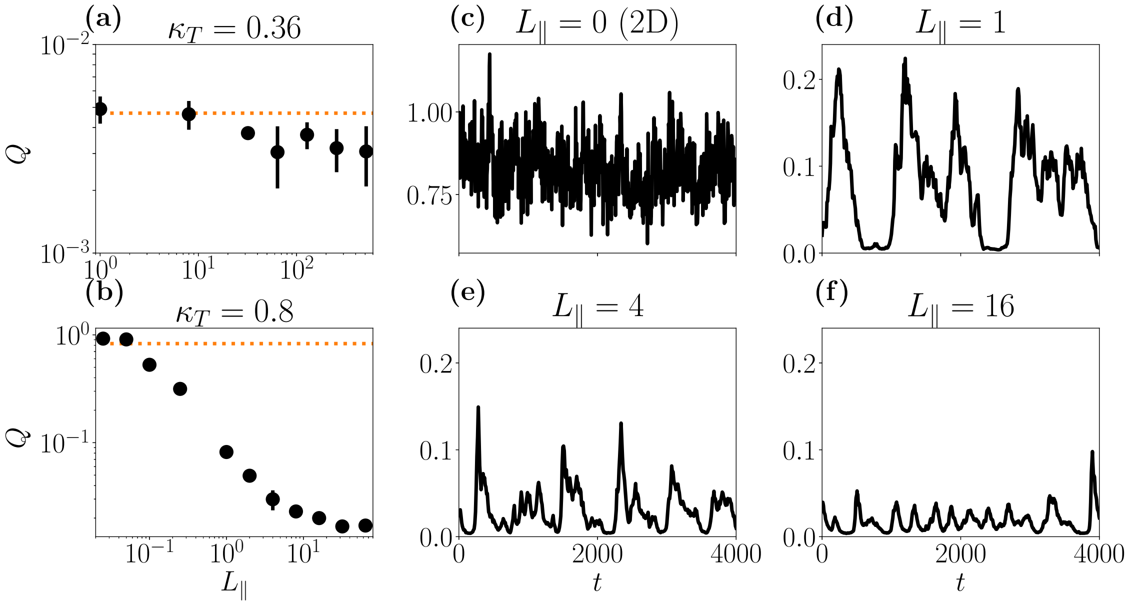


Figure 4.4: **(a)** Dependence of the time-averaged heat flux Q on the parallel size of the box $L_\parallel$ for $\chi = 0.1$, $L_x = L_y = 80$, and $\kappa_T = 0.36$. The orange dotted line denotes the time-averaged heat flux for the 2D state. **(b)** Same as a), but with $\kappa_T = 0.8$. **(c)**–**(f)** Time evolution of the heat flux Q for $\kappa_T = 0.8$, $\chi = 0.1$, $L_x = 80$, $L_y = 80$ and four different values of $L_\parallel$ (notated on each panel). As $L_\parallel$ increases, the turbulent bursts become more frequent and less violent, and the time-averaged Q drops.

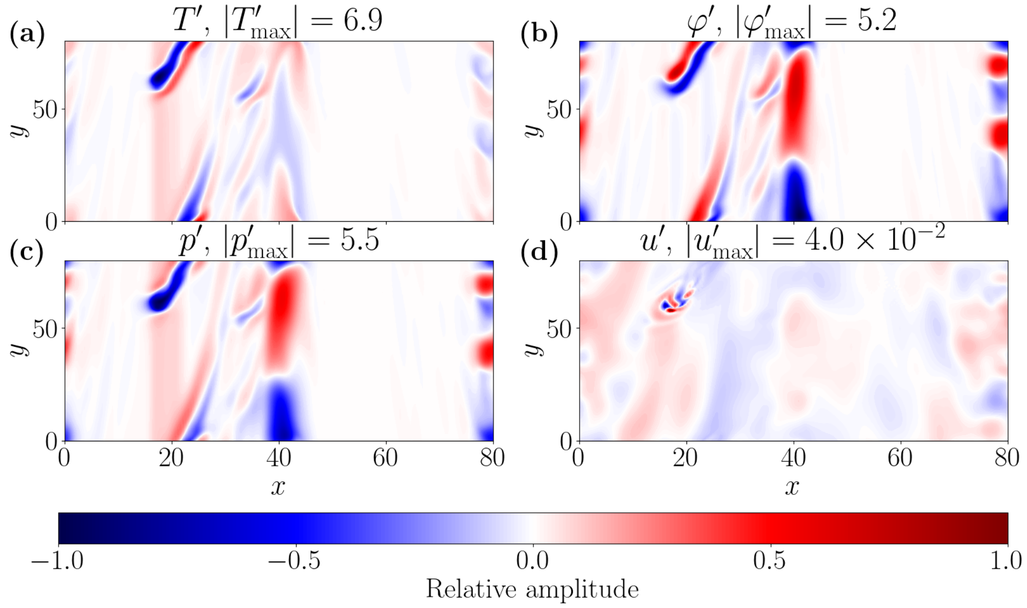


Figure 4.5: Snapshots of nonzonal **(a)** temperature, **(b)** potential, **(c)** pressure, and **(d)** parallel velocity at fixed z in the Dimits state with parameters $\kappa_T = 0.8$, $\chi = 0.1$, $L_x = 80$, $L_y = 80$, and $L_\parallel = 1$. The colour scale is relative to the maximum absolute amplitude in each panel (given in the panels' titles). Small-scale slab-ITG modes driven by the gradients of the ferdinon are evident in (d); however, the small u' amplitude indicates that they are negligible. Indeed, the conservation law (2.31) suggests that the amplitudes of u, φ, and T should be comparable if all of the three fields are dynamically important. This is confirmed by our numerical simulations of strong turbulence where we find that they are indeed comparable (see figure 4.6).

4.2 The slab-ITG parasitic instability

4.2.1 Numerical evidence

Let us now address the small-scale slab-ITG instability. This instability exists only in the 3D model and is the most important distinction between the 2D and 3D models. It is the presence of this instability that enables, in 3D, the existence of a strongly turbulent saturated state, i.e., one in which there are no strong, coherent ZFs (the zonal profiles of such a state are shown in figures 4.1d–f). We find that its most distinctive feature is the concentration of pressure perturbations at perpendicular scales that are much larger than the typical scales for the perturbations in φ and T (or, to be more precise, the absence in the pressure perturbations of the small-scale structure present in φ and T). This is manifest in figure 4.6. Recall that in §2.3.2.2, we showed that small pressure perturbations (as compared to the perturbations in electrostatic potential and temperature) are characteristic of the small-scale ($k_\perp \gg 1$) slab-ITG instability, see (2.63). However, the small-scale structure that we see in figure 4.6 is not a consequence of the equilibrium-driven instability. In fact, the equilibrium-driven slab-ITG instability is inconsequential for the strongly turbulent saturated state. To show this, we ran "modified" simulations where κ_T was set to 0 for all modes with $k_\parallel \neq 0$ (this is straightforward to do in our spectral code). This removed the equilibrium-driven linear instability from all 3D ($k_\parallel \neq 0$) modes. Comparing the spectra of the two conserved quantities $W_{\boldsymbol{k}}$ and $I_{\boldsymbol{k}}$ (see §2.2), we see that turning off the equilibrium temperature gradient for the 3D modes has no appreciable effect on the structure of turbulence (see figure 4.7). As figure 4.8 shows, these modified simulations are also visually indistinguishable from "regular" ones. We see no appreciable change in the saturated turbulent heat flux Q either. Note that the nonlinear interactions of the 2D ($k_\parallel = 0$) modes cannot produce the 3D modes that we see in these "modified" simulations. Therefore, these 3D modes must be produced by a parasitic slab-ITG instability of the 2D fields into which energy is injected.

Additionally, the spectra of $W_{\boldsymbol{k}}$ and $I_{\boldsymbol{k}}$ are inconsistent with the region of linear instability of the dispersion relation (2.34). Namely, figures 4.7a and 4.7b show that $W_{\boldsymbol{k}}$ and $I_{\boldsymbol{k}}$ of the linearly unstable modes of (2.34) are orders of magnitude smaller than

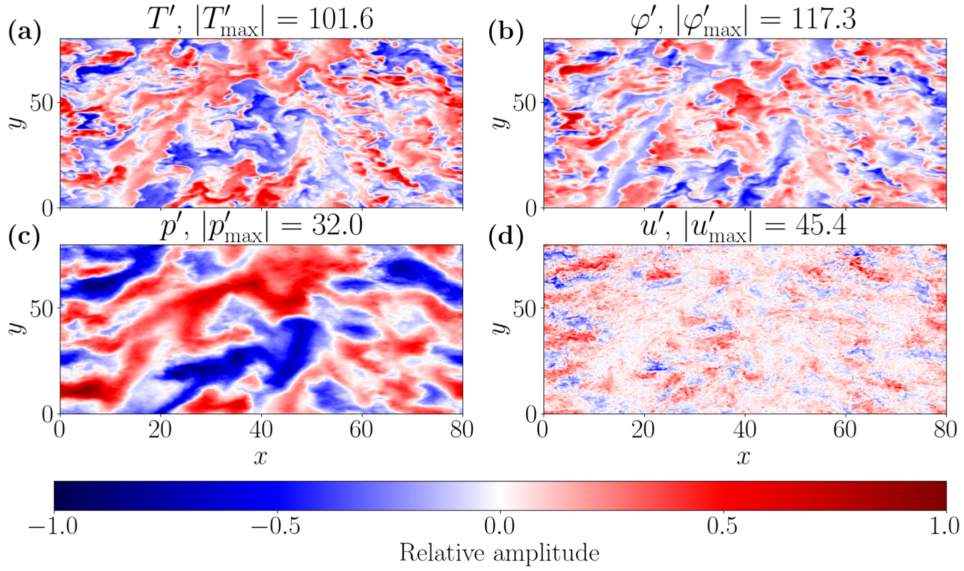


Figure 4.6: Snapshots of nonzonal temperature **(a)**, potential **(b)**, pressure **(c)**, and parallel velocity **(d)** at fixed z in the strongly turbulent state with parameters $\kappa_T = 3$, $\chi = 0.05$, $L_x = 80$, $L_y = 80$, $L_\parallel = 1$, solved at the resolution of $287 \times 287 \times 85$ Fourier modes. The colour scale is relative to the maximum absolute amplitude in each panel (given in the panels' titles). Time-averaged spectra from the same simulation are shown in figure 4.7.

the corresponding spectral peaks of the two conserved quantities. This too suggests that the linearly unstable slab-ITG modes are unimportant in the saturated state. In fact, we can use the turbulent spectra to determine the "dominant" perpendicular scale as a function of the parallel scale $k_\parallel$. As figures 4.7a and 4.7b show, this is the scale at which $I_{\boldsymbol{k}}$ peaks and the dependence of $W_{\boldsymbol{k}}$ on $k_\perp$ changes from flat to steeply declining. To extract this scale, we define $k_{\perp,I}(k_\parallel)$ as the $k_\perp$ that maximises $I_{\boldsymbol{k}}$ at a fixed $k_\parallel$. Figure 4.9a shows that $k_{\perp,I}(k_\parallel)$ lies outside the region of linear instability of (2.34). Thus, the 3D structure of the saturated state is not produced by the linear slab-ITG instability.

Recall that in §2.3.2.2, we showed that the equilibrium-driven slab-ITG instability is localised at $k_\parallel \sim \kappa_T k_\perp^2$. A similar relationship holds for $k_{\perp,I}(k_\parallel)$, viz., $k_\parallel \approx \kappa_T^{\mathrm{eff}} k_{\perp,I}^2$, where κ_T^{eff} can be thought of as an effective temperature gradient. Figure 4.9b shows that this κ_T^{eff} is several times larger than the equilibrium temperature gradient, again confirming that the 3D structure of the strongly turbulent saturated state cannot be

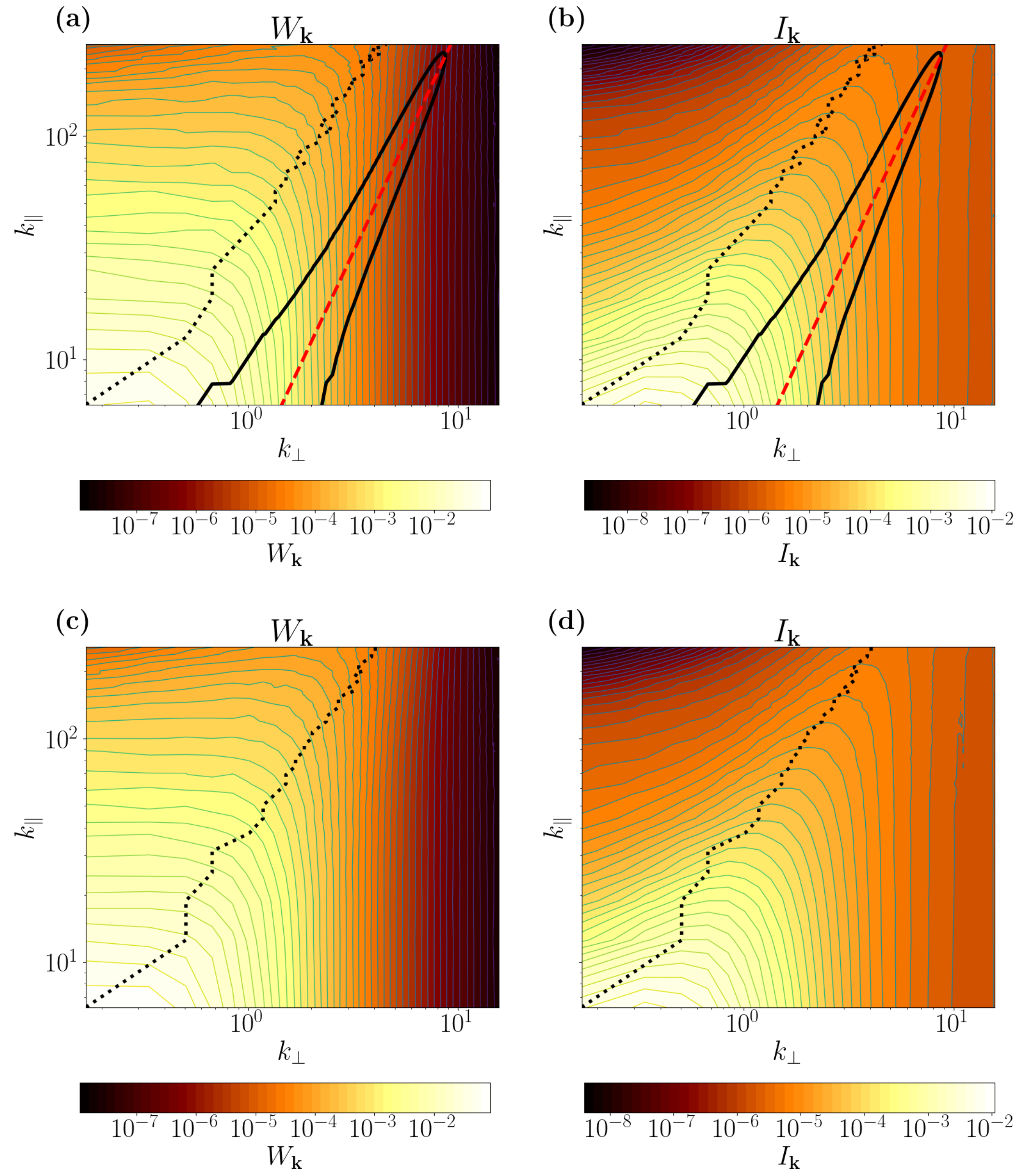


Figure 4.7: Time-averaged spectra **(a)** $W_{\boldsymbol{k}}$ and **(b)** $I_{\boldsymbol{k}}$, defined by (2.32) and (2.33), respectively, in the strongly turbulent state with parameters $\kappa_T = 3$, $\chi = 0.05$, $L_x = 80$, $L_y = 80$, and $L_\parallel = 1$. The solid black lines demarcate the region of linear instability for $k_x = 0$, and the red dashed line is $k_\parallel = \kappa_T k_\perp^2$, where the collisionless modes with largest growth rate reside (see §2.3.2.2). We can see that the largest contributions to the two conserved quantities are offset from the region of linear instability. The dotted black line denotes the peak $k_{\perp,I}(k_\parallel)$ of $I_{\boldsymbol{k}}$ at fixed $k_\parallel$. Zonal profiles and cross-sectional snapshots from the same simulation are shown in figure 4.1 and figure 4.6, respectively. The spectra of the saturated state with the same parameters, but with κ_T set to 0 for all $k_\parallel \neq 0$ modes, are given in **(c)** and **(d)**. We can see that turning off the equilibrium gradient for the 3D modes does not alter the spectra significantly. Snapshots from this "modified" simulation are shown in figure 4.8.

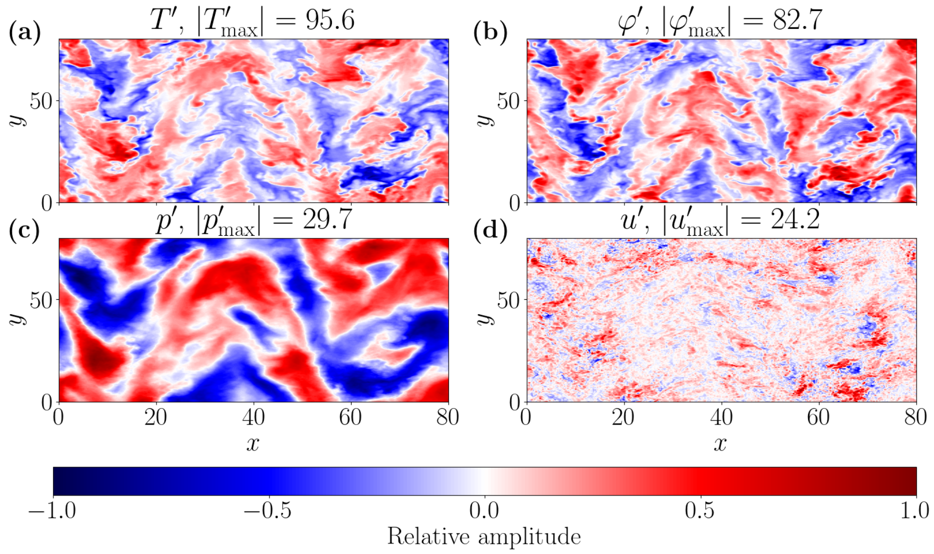


Figure 4.8: Snapshots of nonzonal temperature **(a)**, potential **(b)**, pressure **(c)**, and parallel velocity **(d)** at fixed z in the strongly turbulent state with parameters $\kappa_T = 3$, $\chi = 0.05$, $L_x = 80$, $L_y = 80$, $L_\parallel = 1$, but from a "modified" simulation, i.e., with κ_T set to 0 for the $k_\parallel \neq 0$ modes. The colour scale is relative to the maximum absolute amplitude in each panel (given in the panels' titles). Visually, the saturated state is identical to that shown in figure 4.6.

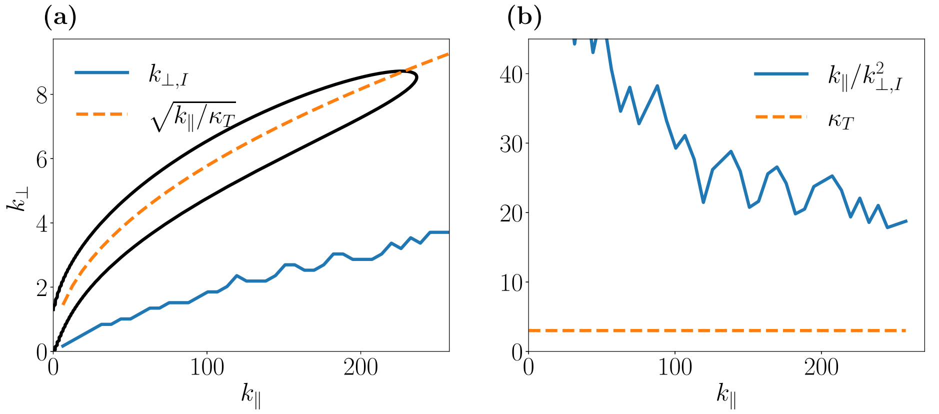


Figure 4.9: **(a)** Comparison of the location of the spectral peak $k_{\perp,I}(k_\parallel)$ (blue) and the location of the peak of the collisionless linear instability driven by the equilibrium gradient (orange dashed), given by $k_\parallel = \kappa_T k_\perp^2$. The black curve circumscribes the region of linear instability. **(b)** Comparison of the equilibrium temperature gradient κ_T and the "effective" temperature gradient κ_T^{eff} defined as $k_\parallel = \kappa_T^{\mathrm{eff}}(k_\parallel)k^2_{\perp,I}$. The data is from a simulation with parameters $\kappa_T = 3$, $\chi = 0.05$, $L_x = 80$, $L_y = 80$, and $L_\parallel = 1$. The spectra of this simulation are given in figure 4.7. This rough estimate for κ_T^{eff} of about 5–10 times larger than κ_T matches the calculated growth rate of the small-scale instability, see figure 4.10.

described by the analysis of §2.3.2.2. As we will see shortly, the effective κ_T^{eff} in question is actually the gradient of the large-scale temperature perturbations.

4.2.2 Scale-separated equations for curvature-ITG and slab-ITG modes

The numerical analysis above leads us to believe that the 3D structure of the saturated state is a consequence of an instability driven not by the equilibrium gradient κ_T, but rather by the gradients of the 2D perturbations. Let us attack on the analytical front. As we discussed at the start of chapter 4, the 3D slab-ITG modes are naturally scale-separated from the 2D curvature-ITG modes both in wavenumber and in frequency. Let us then introduce the parallel average

$$\langle f\rangle_{\parallel} \equiv \int \frac{dz'}{L_z} f(z'). \tag{4.5}$$

This average allows us to split (2.18)–(2.20) into the slow 2D modes governed by the curvature-ITG instability at large perpendicular ($k_\perp \ll 1$) scales, and the fast 3D slab-ITG modes, which live at small perpendicular ($k_\perp \gg 1$) scales.[2] We define the small-scale perturbations as $\widetilde{f} \equiv f - \langle f\rangle_{\parallel}$. The large-scale equations are then

$$\begin{aligned}\partial_t\Big(\langle\varphi'\rangle_{\parallel} - \nabla_\perp^2\langle\varphi\rangle_{\parallel}\Big) - \partial_y\Big(\langle\varphi\rangle_{\parallel} + \langle T\rangle_{\parallel}\Big) + \kappa_T\partial_y\nabla_\perp^2\langle\varphi\rangle_{\parallel}\\ +\Big\{\langle\varphi\rangle_{\parallel}, \langle\varphi'\rangle_{\parallel} - \nabla_\perp^2\langle\varphi\rangle_{\parallel}\Big\} + \boldsymbol{\nabla}_\perp\cdot\Big\{\boldsymbol{\nabla}_\perp\langle\varphi\rangle_{\parallel}, \langle T\rangle_{\parallel}\Big\} + \chi\nabla_\perp^4\Big(a\langle\varphi\rangle_{\parallel} - b\langle T\rangle_{\parallel}\Big) =\\ -\Big\langle\{\widetilde{\varphi}, \widetilde{\varphi}' - \nabla_\perp^2\widetilde{\varphi}\} - \boldsymbol{\nabla}_\perp\cdot\Big\{\boldsymbol{\nabla}_\perp\widetilde{\varphi}, \widetilde{T}\Big\}\Big\rangle_{\parallel},\end{aligned} \tag{4.6}$$

$$\partial_t\langle T\rangle_{\parallel} + \kappa_T\partial_y\langle\varphi\rangle_{\parallel} + \Big\{\langle\varphi\rangle_{\parallel}, \langle T\rangle_{\parallel}\Big\} - \chi\nabla_\perp^2\langle T\rangle_{\parallel} = -\Big\langle\Big\{\widetilde{\varphi}, \widetilde{T}\Big\}\Big\rangle_{\parallel}, \tag{4.7}$$

$$\partial_t\langle u\rangle_{\parallel} + \Big\{\langle\varphi\rangle_{\parallel}, \langle u\rangle_{\parallel}\Big\} - d\chi\nabla_\perp^2\langle u\rangle_{\parallel} = -\langle\{\widetilde{\varphi}, \widetilde{u}\}\rangle_{\parallel}. \tag{4.8}$$

The right-hand sides of (4.6)–(4.8) represent the influence of the 3D slab-ITG modes on the large-scale fields. Subtracting (4.6)–(4.8) from (2.18)–(2.20), we find the small-scale

---

[2] A more accurate analysis should not average over the entire parallel extent of the box, but only over l_z defined to be larger that the scale of the slab-ITG modes and smaller than the parallel scale of the curvature-ITG-like modes. As discussed in §2.3.2.1, modes with $k_{\parallel} \ll 1$ behave like curvature-ITG modes with finite-$k_{\parallel}$ modifications. Here we have taken a cruder approach for the sake of simplifying the analysis. However, note that modes with $k_{\parallel} \lesssim 1$ are usually not included in our simulations of strong turbulence for numerical reasons as we need a large maximum $k_{\parallel}$ in order to resolve the slab-ITG instability (see §4.3.2). Thus, this cruder approach is sufficient for the analysis of the simulations that we report in §4.2.6.

equations:

$$\begin{aligned}
\partial_t\big(\widetilde{\varphi}' - \nabla_\perp^2\widetilde{\varphi}\big) + \partial_\parallel\widetilde{u} - \partial_y(\widetilde{\varphi} + \widetilde{T}) + \kappa_T\partial_y\nabla_\perp^2\widetilde{\varphi} + \widetilde{\{\widetilde{\varphi}, \widetilde{\varphi}' - \nabla_\perp^2\widetilde{\varphi}\}} + \boldsymbol{\nabla}_\perp\cdot\widetilde{\Big\{\boldsymbol{\nabla}_\perp\widetilde{\varphi}, \widetilde{T}\Big\}} \\
+\chi\nabla_\perp^4\Big(a\widetilde{\varphi} - b\widetilde{T}\Big) = -\Big[\Big\{\langle\varphi\rangle_\parallel, \widetilde{\varphi}' - \nabla_\perp^2\widetilde{\varphi}\Big\} + \Big\{\widetilde{\varphi}, \langle\varphi'\rangle_\parallel - \nabla_\perp^2\langle\varphi\rangle_\parallel\Big\} \\
+\boldsymbol{\nabla}_\perp\cdot\Big\{\boldsymbol{\nabla}_\perp\langle\varphi\rangle_\parallel, \widetilde{T}\Big\} + \boldsymbol{\nabla}_\perp\cdot\Big\{\boldsymbol{\nabla}_\perp\widetilde{\varphi}, \langle T\rangle_\parallel\Big\}\Big], \qquad (4.9)
\end{aligned}$$

$$\partial_t\widetilde{T} + \kappa_T\partial_y\widetilde{\varphi} + \widetilde{\Big\{\widetilde{\varphi}, \widetilde{T}\Big\}} - \chi\nabla_\perp^2\widetilde{T} = -\Big[\Big\{\langle\varphi\rangle_\parallel, \widetilde{T}\Big\} + \Big\{\widetilde{\varphi}, \langle T\rangle_\parallel\Big\}\Big], \qquad (4.10)$$

$$\partial_t\widetilde{u} + \partial_\parallel\Big(\widetilde{\varphi} + \widetilde{T}\Big) + \widetilde{\{\widetilde{\varphi}, \hat{u}\}} - d\chi\nabla_\perp^2\hat{u} = -\Big[\Big\{\langle\varphi\rangle_\parallel, \widetilde{u}\Big\} + \Big\{\widetilde{\varphi}, \langle u\rangle_\parallel\Big\}\Big]. \qquad (4.11)$$

In order to simplify the following analysis, we shall assume that the spatial and temporal scales of (4.9)–(4.11) are infinitesimal compared to the scales of (4.6)–(4.8) (i.e., we assume scale separation).

4.2.3 Parasitic slab-ITG instability

First, we investigate the slab-ITG instability in the presence of large-scale modes. Let us simplify (4.9)–(4.11) by linearising them and assuming that the 2D modes have perpendicular scales $k_{\perp,\mathrm{2D}} \ll k_{\perp,\mathrm{3D}} \sim 1$. Equations (4.9)–(4.11) then become

$$\begin{aligned}
\Big(\partial_t + \langle\boldsymbol{V}_E\rangle_\parallel\cdot\boldsymbol{\nabla}_\perp\Big)\big(\widetilde{\varphi}' - \nabla_\perp^2\widetilde{\varphi}\big) + \partial_\parallel\widetilde{u} - \partial_y(\widetilde{\varphi} + \widetilde{T}) \\
+ \boldsymbol{\kappa}_T\cdot\boldsymbol{\nabla}_\perp\nabla_\perp^2\widetilde{\varphi} + \boldsymbol{\kappa}_n\cdot\boldsymbol{\nabla}_\perp\widetilde{\varphi} = -\chi\nabla_\perp^4(a\widetilde{\varphi} - b\widetilde{T}), \qquad (4.12)
\end{aligned}$$

$$\Big(\partial_t + \langle\boldsymbol{V}_E\rangle_\parallel\cdot\boldsymbol{\nabla}_\perp\Big)\widetilde{T} + \boldsymbol{\kappa}_T\cdot\boldsymbol{\nabla}_\perp\widetilde{\varphi} = \chi\nabla_\perp^2\widetilde{T}, \qquad (4.13)$$

$$\Big(\partial_t + \langle\boldsymbol{V}_E\rangle_\parallel\cdot\boldsymbol{\nabla}_\perp\Big)\widetilde{u} + \partial_\parallel\Big(\widetilde{\varphi} + \widetilde{T}\Big) = d\nabla_\perp^2\hat{u}, \qquad (4.14)$$

where

$$\langle\boldsymbol{V}_E\rangle_\parallel \equiv \hat{\boldsymbol{z}}\times\boldsymbol{\nabla}_\perp\langle\varphi\rangle_\parallel, \quad \boldsymbol{\kappa}_n \equiv -\hat{\boldsymbol{z}}\times\boldsymbol{\nabla}_\perp\langle\varphi'\rangle_\parallel, \quad \boldsymbol{\kappa}_T \equiv \kappa_T\hat{\boldsymbol{y}} - \hat{\boldsymbol{z}}\times\boldsymbol{\nabla}_\perp\langle T\rangle_\parallel \qquad (4.15)$$

represent the $\boldsymbol{E}\times\boldsymbol{B}$ advection, the local density gradient, and the total local temperature gradient (large-scale perturbations plus equilibrium),[3] respectively. Note that only the nonzonal electrostatic potential φ' gives rise to a density perturbation — this is a consequence of the modified adiabatic electron response (2.4). Note also that we

---

[3] We have kept the equilibrium gradient for completeness. As we showed in §4.2.1, its effect is negligible compared to that of the gradients associated with the large-scale perturbations.

have ignored the large-scale $\langle u\rangle_\parallel$ because it is not involved in any linear instability. In our numerical simulations, we find that the 2D $\langle u\rangle_\parallel$ perturbations are many orders of magnitude smaller than the other perturbations and are thus irrelevant for the saturated state. This is consistent with the discussion leading to (2.22) and (2.23).

Ignoring collisions (i.e., setting $\chi = 0$) and taking the gradients of the large-scale fields to be constant over the small scales at which (4.12)–(4.14) hold, we can investigate the small-scale linear instability in a way analogous to the discussion in §2.3. In particular, we shall focus on the $k_\parallel \sim k_\perp^2 \gg 1$ regime analysed in §2.3.2.2. We look for Doppler-shifted solutions to (4.12)–(4.14) of the form $\propto \exp\left[-i\left(\omega_{\boldsymbol{k}} + \langle \boldsymbol{V}_E\rangle_\parallel \cdot \boldsymbol{k}\right)t + i\boldsymbol{k}\cdot\boldsymbol{r}\right]$. We ignore the magnetic-drift term $-\partial_y\left(\widetilde{\varphi} + \widetilde{T}\right)$ in (4.12) because it is subdominant for the slab-ITG modes with $k_\perp \gg 1$. The dispersion relation for these modes is

$$\left(\omega_{\boldsymbol{k}}^2 - \frac{k_\parallel^2}{1+k_\perp^2}\right)(\omega_{\boldsymbol{k}} + \boldsymbol{\kappa}_T\cdot\boldsymbol{k}) = \frac{\omega_{\boldsymbol{k}}^2}{1+k_\perp^2}[(\boldsymbol{\kappa}_n + \boldsymbol{\kappa}_T)\cdot\boldsymbol{k}]. \tag{4.16}$$

Since (4.12)–(4.14) describe real fields, (4.16) must be invariant under $\boldsymbol{k} \mapsto -\boldsymbol{k}$, $\omega_{\boldsymbol{k}} \mapsto -\omega_{\boldsymbol{k}}^*$. We may, therefore, assume that $\boldsymbol{\kappa}_T\cdot\boldsymbol{k} > 0$ without loss of generality. Repeating the arguments of §2.3.2.2, we define $\omega_{\boldsymbol{k}} \equiv \hat{\omega}_{\boldsymbol{k}}\boldsymbol{\kappa}_T\cdot\boldsymbol{k}$, $k_\parallel \equiv \hat{k}_\parallel\boldsymbol{\kappa}_T\cdot\boldsymbol{k}$. Then (4.16) turns out to be equivalent to our old dispersion relation (2.48), but now with

$$\hat{\gamma}_{\boldsymbol{k}}^2 = \frac{(\boldsymbol{\kappa}_n + \boldsymbol{\kappa}_T)\cdot\boldsymbol{k}}{2k_\perp^2\boldsymbol{\kappa}_T\cdot\boldsymbol{k}}. \tag{4.17}$$

Thus, the results of §2.3.2.2 carry over to the instability described by (4.16). In particular, the slab-ITG instability exists if $\hat{\gamma}_{\boldsymbol{k}}^2 > 0$, i.e., if $(\boldsymbol{\kappa}_n + \boldsymbol{\kappa}_T)\cdot\boldsymbol{k}$ and $\boldsymbol{\kappa}_T\cdot\boldsymbol{k}$ have the same sign, and is localised at $k_\parallel \approx \pm\boldsymbol{\kappa}_T\cdot\boldsymbol{k}k_\perp$. Its growth rate is given by

$$\operatorname{Im}(\omega_{\boldsymbol{k}}) \approx \operatorname{Re}\sqrt{\frac{\boldsymbol{\kappa}_T\cdot\boldsymbol{k}(\boldsymbol{\kappa}_n + \boldsymbol{\kappa}_T)\cdot\boldsymbol{k}}{2k_\perp^2}}. \tag{4.18}$$

As expected, this is the same as (2.62) if $\boldsymbol{\kappa}_n = 0$ and $\boldsymbol{\kappa}_T = \kappa_T\hat{\boldsymbol{y}}$. In figure 4.10b, we show the maximum growth rate obtained from the numerical solution of the full (with collisionality and magnetic curvature turned back on) dispersion relation with the addition of the local temperature and density gradients of the large-scale fields. As expected from the numerical analysis in §4.2.1, the small-scale instability driven by the large-scale gradients is significantly ($\approx$ 5 times in this case) stronger than

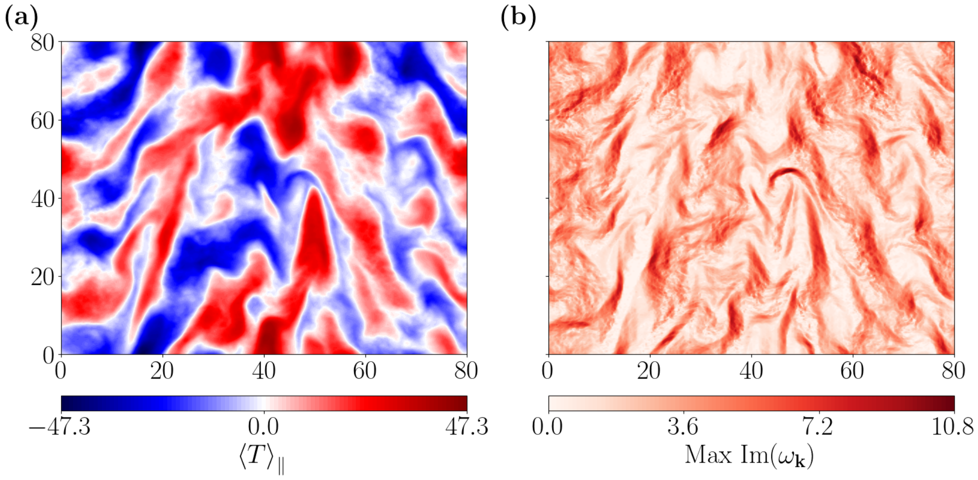


Figure 4.10: **(a)** Snapshot of the 2D temperature perturbation $\langle T\rangle_\parallel$ in the (x, y) plane. The data is taken from the same $\kappa_T = 3$, $\chi = 0.05$ simulation that we showed in figure 4.6. The 2D temperature perturbations lack the small-scale structure that could be seen in figure 4.6a. **(b)** Small-scale growth rate in the (x, y) plane. This plot is obtained by finding the maximum growth rate of the full (with collisionality and magnetic curvature) dispersion relation (2.34) with the addition of the local temperature and density gradients of the large-scale fields at every point. For this simulation, $\kappa_T = 3$, and so the largest collisionless growth rate, given by (2.62), is $\kappa_T/\sqrt{2} \approx 2.1$. It is thus evident that the influence of the gradients of the large-scale fields dominates over that of the equilibrium gradient κ_T by a factor of 5. In fact, the "effective" κ_T^{eff} that we estimated for the same simulation in figure 4.9b is a factor of 5–10 larger that the equilibrium gradient κ_T.

the equilibrium-driven instability. This is consistent with the estimate of the effective temperature gradient κ_T^{eff} for the slab-ITG instability that we did in §4.2.1 and is shown in figure 4.9b.

Note that if $\boldsymbol{\kappa}_n \neq 0$, (4.18) implies that modes with $\boldsymbol{\kappa}_T \cdot \boldsymbol{k}(\boldsymbol{\kappa}_n + \boldsymbol{\kappa}_T) \cdot \boldsymbol{k} < 0$ are linearly stable. But is there a $\boldsymbol{\kappa}_n$ that quenches the slab-ITG instability for all $\boldsymbol{k}$? Let $\boldsymbol{\kappa}_T \equiv |\boldsymbol{\kappa}_T|\boldsymbol{a}$, $\boldsymbol{\kappa}_T + \boldsymbol{\kappa}_n \equiv |\boldsymbol{\kappa}_T|(\alpha\boldsymbol{a} + \beta\boldsymbol{b})$, where $|\boldsymbol{a}| = |\boldsymbol{b}| = 1$. Then, $\boldsymbol{\kappa}_T \cdot \boldsymbol{k}(\boldsymbol{\kappa}_n + \boldsymbol{\kappa}_T) \cdot \boldsymbol{k} = |\boldsymbol{\kappa}_T|^2 k_a(\alpha k_a + \beta k_b)$, where $k_a \equiv \boldsymbol{k} \cdot \boldsymbol{a}$ and $k_b \equiv \boldsymbol{k} \cdot \boldsymbol{b}$. Choosing $k_b = 0$ implies that $\alpha \leq 0$. Choosing $k_a > 0$ and k_b such that $\beta k_b > -\alpha k_a$ gives $\boldsymbol{\kappa}_T \cdot \boldsymbol{k}(\boldsymbol{\kappa}_n + \boldsymbol{\kappa}_T) > 0$, unless $\beta = 0$. Therefore, in order to quench the slab-ITG instability for all $\boldsymbol{k}$, we need

$$\boldsymbol{\kappa}_n \parallel \boldsymbol{\kappa}_T, \quad \frac{\boldsymbol{\kappa}_n \cdot \boldsymbol{\kappa}_T}{|\boldsymbol{\kappa}_T|^2} \leq -1. \tag{4.19}$$

Later, in §4.2.6, we shall show that the effect that the growing small-scale modes have on the large-scale fields is to push the latter toward the stability condition (4.19). Now, we shall investigate the role of the small-scale slab-ITG modes in stabilising the strong ZFs of the Dimits regime.

4.2.4 Turbulent stress of the slab-ITG modes

In §3.3.4, we obtained a prediction for the critical gradient $\kappa_T^{c,\mathrm{2D}}(\chi)$ above which a Dimits state with strong ZFs cannot be sustained. This prediction was based on considerations of the ratio $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})$ for the linear modes with largest growth rate. As explained in §3.3 and §4.1, this ratio determines the balance of Reynolds and diamagnetic stresses for an individual Fourier mode: if $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) > -1$, then the Reynolds stress is larger and the mode favours a Dimits state, otherwise its diamagnetic stress is larger and the mode helps suppress the coherent ZFs needed for the Dimits state. In 2D, this ratio is sensitive to both the temperature gradient κ_T and the collisionality χ, and thus an appropriate balance between these two parameters is required in order to have $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) > -1$ for the dominant modes and thus to keep the system in the Dimits state. In particular, for $\kappa_T \gg 1$, the Dimits threshold is given by $\kappa_T/\chi = \mathrm{const}$.

Let us adopt a similar approach for the fastest-growing small-scale slab-ITG modes located at $k_\parallel \approx \boldsymbol{\kappa}_T \boldsymbol{\cdot} \boldsymbol{k} k_\perp$. Equation (2.63) tells us that these modes satisfy $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) = -1 + \mathcal{O}\big(k_\perp^{-2}\big)$ for $k_\perp \gg 1$. Therefore, to lowest order, the slab-ITG modes are Dimits-marginal, i.e., their Reynolds and diamagnetic stresses balance out. This means that the lowest-order collisionless calculations of §2.3.2.2 are insufficient for our needs. While we can extend the calculations to $\mathcal{O}\big(k_\perp^{-2}\big)$, ignoring χ is an oversimplification. Collisionality is necessary for saturation because it provides the only sink of energy in the system. And, as we shall shortly see, even far away from the collisional cut-off, nonzero χ can provide a $\mathcal{O}\big(k_\perp^{-1}\big)$ correction to $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})$ and hence render any collisionless higher-order correction irrelevant.

4.2.5 Effects of collisionality

As discussed in the beginning of chapter 4, the collisional cut-off for the slab-ITG instability scales as $k_{\perp,\text{cut-off}}^2 \sim \kappa_T/\chi$. In other words, in order to damp the instability at a given $k_\perp$, we need a collisionality $\chi \sim \kappa_T/k_\perp^2$. It turns out that even an asymptotically

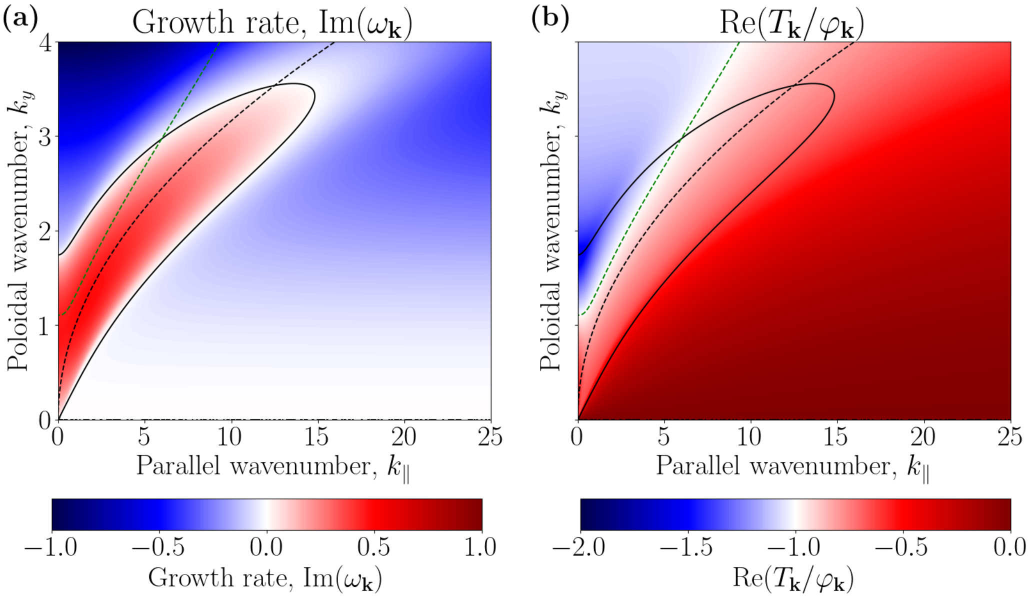


Figure 4.11: **(a)** Linear growth rate and **(b)** the ratio $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})$ of the most unstable mode versus $k_\parallel$ and k_y for the most unstable $k_x = 0$ modes for $\kappa_T = 1$, $\chi = 0.1$. The green dashed line is the $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) = -1$ line. The black dashed line is the location of the largest collisionless growth rate $k_\parallel = \kappa_T k_y^2$. While the collisionless modes would have the green line on top of the black one, we see that the addition of collisions has shifted the linearly unstable modes towards the Dimits-favourable $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) > -1$ ratio.

weaker collisionality $\chi \sim \kappa_T/k_\perp^3$ produces a $\mathcal{O}(1/k_\perp)$ correction to $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})$ for the fastest-growing slab-ITG modes. The details of the relevant calculation can be found in appendix C.2. We find that for χ ordered as $\kappa_T \sim \chi k_\perp^3$, the most unstable small-scale slab-ITG modes are still located at $k_\parallel = \boldsymbol{\kappa}_T \cdot \boldsymbol{k} k_\perp$ and satisfy

$$\frac{T_{\boldsymbol{k}}}{\varphi_{\boldsymbol{k}}} = -1 - \sqrt{-\hat{\gamma}_{\boldsymbol{k}}^2 + \frac{i(a+b-1)\chi k_\perp^2}{2\boldsymbol{\kappa}_T \cdot \boldsymbol{k}}} + \mathcal{O}\big(k_\perp^{-2}\big), \tag{4.20}$$

where we take the branch of the square root with a positive imaginary part. Since $\hat{\gamma}_{\boldsymbol{k}}^2 > 0$ and $\boldsymbol{\kappa}_T \cdot \boldsymbol{k} > 0$ (see §4.2.3), we find that the sign of the real part of the square root in (4.20) is given by the sign of $a+b-1$. Specifically, we have $a = 9/40$, $b = 67/160$, so $a + b - 1 < 0$, hence the square root has a negative real part and $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) > -1$. Thus, collisionality always pushes the otherwise Dimits-marginal small-scale slab-ITG modes to side with the Reynolds stress and reinforce the ZFs.

The sensitivity of $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})$ to the numerical factors a and b allows us to carry out a simple test of the above theory. We restart a 3D Dimits-state simulation above the 2D Dimits threshold, i.e., with $\kappa_T > \kappa_T^{c,\mathrm{2D}}$, but setting $a = 1$ for all nonzonal modes. Linearly, this increases viscosity and reduces growth rates, without affecting zonal physics. Intuitively, one might expect that with an increased damping of the turbulence, the Dimits state should become “stronger”. However, such reasoning does not take into account the structure of the 3D modes and the change to the balance of Reynolds and diamagnetic stress, which is a consequence of the change of the sign of $a + b - 1$. Indeed, in this numerical experiment, we discover that the Dimits regime is destroyed and strong turbulence sets in, just as the analysis above predicts. This is clear evidence that the most consequential role of collisionality for the Dimits regime of (2.18)–(2.20) is not to dissipate turbulent energy, but rather to regulate the turbulent stress via the ratio $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})$.

4.2.6 Large-scale response: effective thermal diffusion

We expect that the growth of small-scale slab-ITG modes that are driven by large-scale gradients, will diminish the amplitudes of the driving large-scale fields. This is an intuitive consequence of the conservation laws (2.28) and (2.31) in (2.18)–(2.20): as the

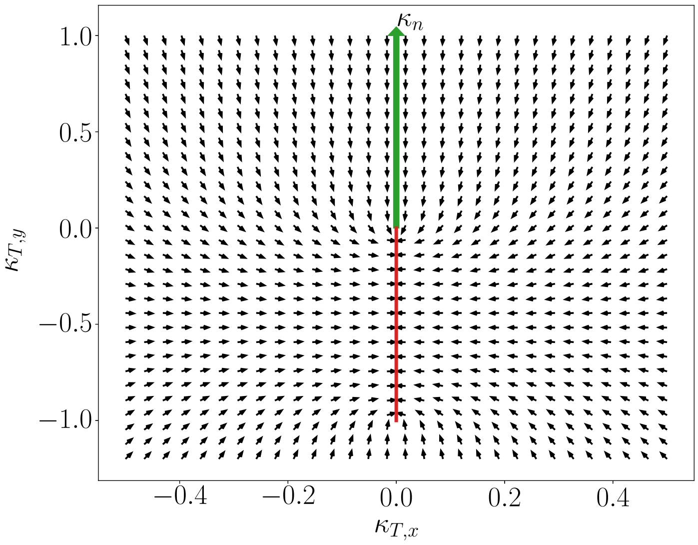


Figure 4.12: This plot shows the direction of the small-scale heat flux $-\hat{\boldsymbol{z}} \times \langle \widetilde{\boldsymbol{Q}} \rangle_{\parallel}$ of the most unstable slab-ITG modes as a function of $\boldsymbol{\kappa}_T = -\hat{\boldsymbol{z}} \times \boldsymbol{\nabla}_{\perp} \langle T \rangle_{\parallel}$. We have chosen a coordinate system in which the large-scale density gradient is $\boldsymbol{\kappa}_n = (0, 1)$, denoted by the green arrow. The red line shows the values of $\boldsymbol{\kappa}_T$ for which the slab-ITG instability has zero growth rate, according to (4.19). The black arrows represent the direction of $-\hat{\boldsymbol{z}} \times \langle \widetilde{\boldsymbol{Q}} \rangle_{\parallel}$. We see that $\langle \widetilde{\boldsymbol{Q}} \rangle_{\parallel}$ pushes the large-scale temperature gradient $\boldsymbol{\kappa}_T$ towards the linearly stable region.

parasitic instability is driven by the nonlinear terms that conserve $W = \sum_{\boldsymbol{k}} W_{\boldsymbol{k}}$ and $I = \sum_{\boldsymbol{k}} I_{\boldsymbol{k}}$, an increase of, say, $W_{\boldsymbol{k}}$ for some $\boldsymbol{k}$ must be accompanied by a decrease at another wavenumber in order to keep $W = \sum_{\boldsymbol{k}} W_{\boldsymbol{k}}$ constant. In particular, an increase in small-scale amplitudes is accompanied by a decrease in large-scale ones.

Let us now calculate explicitly the influence of small-scale slab-ITG modes on the large-scale modes and show that this is indeed true. This influence is represented by the terms of the form $\langle\{.,.\}\rangle_{\parallel}$ on the right-hand sides of (4.6)–(4.8).

First, consider (4.7). The relevant term is

$$-\left\langle\left\{\widetilde{\varphi}, \widetilde{T}\right\}\right\rangle_{\parallel} = -\boldsymbol{\nabla}_{\perp} \cdot \left\langle\left[(\hat{\boldsymbol{z}} \times \boldsymbol{\nabla}_{\perp}\widetilde{\varphi})\widetilde{T}\right]\right\rangle_{\parallel} \equiv -\boldsymbol{\nabla}_{\perp} \cdot \left\langle\widetilde{\boldsymbol{Q}}\right\rangle_{\parallel}, \tag{4.21}$$

where $\left\langle\widetilde{\boldsymbol{Q}}\right\rangle_{\parallel}$ is the turbulent heat flux associated with the small-scale modes. Let us compute it assuming scale separation. As stated in §4.2.2, we assume that the small-scale equations are solved in an infinitesimal (compared to the large scales) box, and thus the parallel average is equivalent to an average over the small-scale box:

$$-\left\langle\widetilde{\boldsymbol{Q}}\right\rangle_{\parallel} = \sum_{\boldsymbol{q}} i\hat{\boldsymbol{z}} \times \boldsymbol{q}\widetilde{\varphi}_{-\boldsymbol{q}}\widetilde{T}_{\boldsymbol{q}} \approx -\sum_{\boldsymbol{q}} \frac{\hat{\boldsymbol{z}} \times \boldsymbol{q}}{q_{\perp}\sqrt{2}} \sqrt{\frac{(\boldsymbol{\kappa}_T + \boldsymbol{\kappa}_n) \cdot \boldsymbol{q}}{\boldsymbol{\kappa}_T \cdot \boldsymbol{q}}} |\widetilde{\varphi}_{\boldsymbol{q}}|^2, \tag{4.22}$$

where we have assumed that the sum is dominated by the wavenumbers $\boldsymbol{q}$ with the largest linear growth rate, and so replaced $\widetilde{T}_{\boldsymbol{q}}/\widetilde{\varphi}_{\boldsymbol{q}}$ with the collisionless expression (2.63) for the modes with $k_{\parallel} = \boldsymbol{\kappa}_T \cdot \boldsymbol{k} k_{\perp}$ that maximise this growth rate. Note that the small-scale fields $\widetilde{T}_{\boldsymbol{q}}$ and $\widetilde{\varphi}_{\boldsymbol{q}}$, and thus $\left\langle\widetilde{\boldsymbol{Q}}\right\rangle_{\parallel}$ itself, depend implicitly on the position variable of the large-scale equations (4.6)–(4.8).

In order to verify that $\left\langle\widetilde{\boldsymbol{Q}}\right\rangle_{\parallel}$ does indeed damp the large-scale temperature perturbations $\langle T\rangle_{\parallel}$, we multiply (4.7) by $\langle T\rangle_{\parallel}$ and integrate over space to find

$$\begin{aligned}
&\partial_t \int d^3\boldsymbol{r} \frac{1}{2}\langle T\rangle_{\parallel}^2 + \text{linear terms} = \int d^3\boldsymbol{r} \left\langle\widetilde{\boldsymbol{Q}}\right\rangle_{\parallel} \cdot \boldsymbol{\nabla}_{\perp}\langle T\rangle_{\parallel} \\
&\approx -\int d^3\boldsymbol{r} \sum_{\boldsymbol{q}} \frac{(\hat{\boldsymbol{z}} \times \boldsymbol{q}) \cdot \boldsymbol{\nabla}_{\perp}\langle T\rangle_{\parallel}}{q_{\perp}\sqrt{2}} \sqrt{\frac{(\boldsymbol{\kappa}_T + \boldsymbol{\kappa}_n) \cdot \boldsymbol{q}}{\boldsymbol{\kappa}_T \cdot \boldsymbol{q}}} |\widetilde{\varphi}_{\boldsymbol{q}}|^2 \\
&= -\int d^3\boldsymbol{r} \sum_{\boldsymbol{q}} \sqrt{\frac{\boldsymbol{\kappa}_T \cdot \boldsymbol{q}(\boldsymbol{\kappa}_T + \boldsymbol{\kappa}_n) \cdot \boldsymbol{q}}{2q_{\perp}^2}} |\widetilde{\varphi}_{\boldsymbol{q}}|^2 = -\int d^3\boldsymbol{r} \sum_{\boldsymbol{q}} \mathrm{Im}(\omega_{\boldsymbol{k}})|\widetilde{\varphi}_{\boldsymbol{q}}|^2,
\end{aligned} \tag{4.23}$$

where $\mathrm{Im}(\omega_{\boldsymbol{k}})$ is the slab-ITG growth rate (4.18). Thus, the linearly unstable small-scale modes have a sign-definite effect on $\langle T\rangle_\parallel$: they provide additional dissipation.

$\left\langle\widetilde{\boldsymbol{Q}}\right\rangle_\parallel$ depends on $\langle T\rangle_\parallel$ in a nontrivial way. Let us quantify its influence on $\langle T\rangle_\parallel$ by working out its direction as a function of $\boldsymbol{\kappa}_T$. We can write (4.22) as

$$-\left\langle\widetilde{\boldsymbol{Q}}\right\rangle_\parallel \approx -\sum_{\boldsymbol{q}} \frac{\hat{\boldsymbol{z}}\times\hat{\boldsymbol{q}}}{\sqrt{2}}\sqrt{\frac{(\boldsymbol{\kappa}_T+\boldsymbol{\kappa}_n)\cdot\hat{\boldsymbol{q}}}{\boldsymbol{\kappa}_T\cdot\hat{\boldsymbol{q}}}}|\widetilde{\varphi}_{\boldsymbol{q}}|^2, \tag{4.24}$$

where $\hat{\boldsymbol{q}} = \boldsymbol{q}/q_\perp$ is the direction of the wavevector $\boldsymbol{q}$. Note that the slab-ITG growth rate (4.18) also depends only on $\hat{\boldsymbol{q}}$ and not $|\boldsymbol{k}|$. Let us assume that $\left\langle\widetilde{\boldsymbol{Q}}\right\rangle_\parallel$ is dominated by the fastest-growing slab-ITG modes, and let their wavevector direction be $\hat{\boldsymbol{q}}_{\max}$, so $\left\langle\widetilde{\boldsymbol{Q}}\right\rangle_\parallel$ is parallel to $\hat{\boldsymbol{z}}\times\hat{\boldsymbol{q}}_{\max}$.[4] In figure 4.12, we illustrate the influence on $\boldsymbol{\kappa}_T$ of the contribution to $\left\langle\widetilde{\boldsymbol{Q}}\right\rangle_\parallel$ from the most unstable small-scale modes. As expected, we find that the small-scale heat flux pushes the large-scale gradient $\boldsymbol{\kappa}_T$ towards the linearly stable configuration (4.19).

Now consider (4.6), the evolution equation of $\langle\varphi\rangle_\parallel$. The relevant nonlinear terms are

$$\left\langle\{\widetilde{\varphi},\widetilde{\varphi}'-\nabla_\perp^2\widetilde{\varphi}\}+\boldsymbol{\nabla}_\perp\cdot\left\{\boldsymbol{\nabla}_\perp\widetilde{\varphi},\widetilde{T}\right\}\right\rangle_\parallel = \boldsymbol{\nabla}_\perp\cdot\langle\{\boldsymbol{\nabla}_\perp\widetilde{\varphi},\widetilde{p}\}\rangle_\parallel$$

$$= \boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp : \sum_{\boldsymbol{q}}(\hat{\boldsymbol{z}}\times\boldsymbol{q})\boldsymbol{q}\left(1+\mathrm{Re}\frac{\widetilde{T}_{\boldsymbol{q}}}{\widetilde{\varphi}_{\boldsymbol{q}}}\right)|\widetilde{\varphi}_{\boldsymbol{q}}|^2 \equiv \boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp : \left\langle\widetilde{\Pi}\right\rangle_\parallel. \tag{4.25}$$

The collisionless modes satisfy $1+\mathrm{Re}(\widetilde{T}_{\boldsymbol{q}}/\widetilde{\varphi}_{\boldsymbol{q}}) = \mathcal{O}(q_\perp^{-2})$, whence $\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp : \left\langle\widetilde{\Pi}\right\rangle_\parallel \sim k_\perp^2|\hat{\varphi}|^2$. However, (4.22) implies $\boldsymbol{\nabla}_\perp\cdot\left\langle\widetilde{\boldsymbol{Q}}\right\rangle_\parallel \sim k_\perp|\hat{\varphi}|^2 \gg \boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp : \left\langle\widetilde{\Pi}\right\rangle_\parallel$ since the large-scale modes live at $k_\perp \ll 1$. Thus, the main effect of the small-scale modes is to provide a feedback to the large-scale temperature in the form of the additional heat flux $\left\langle\widetilde{\boldsymbol{Q}}\right\rangle_\parallel$.

4.3 Breaking the Dimits state

Recall that the 2D critical gradient $\kappa_T^{c,\mathrm{2D}}$ was found to be an increasing function of χ. Naïvely, this makes sense on the basis of "more dissipation means less turbulence": one expects that one should be able to compensate an increase in the drive κ_T by

---

[4] This is equivalent to the statement that the equilibrium-driven heat flux is in the radial direction.

an appropriate increase in χ and thus keep the system in the Dimits state. However, this simple picture is false. Collisionality and drive are important for maintaining the Dimits state not because they provide dissipation and injection of energy, but rather because they adjust the ratio $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}})$ of the linearly unstable modes. In 2D, this ratio is sensitive to both κ_T and χ; however, this is not the case in 3D as the small-scale slab-ITG modes always favour the Dimits state. First, their turbulent momentum flux was shown to satisfy $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) \approx -1$ (see §4.2.4) with collisions pushing this ever so slightly in the Dimits-stable direction of $\mathrm{Re}(T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}) \gtrsim -1$ (see §4.2.5). Secondly, they provide effective thermal diffusion for the large-scale $\langle T\rangle_{\parallel}$, which in turn reduces the absolute value of $\langle T_{\boldsymbol{k}}\rangle_{\parallel}/\langle\varphi_{\boldsymbol{k}}\rangle_{\parallel}$ and partially suppresses the tendency of large-scale modes to destroy the Dimit state. Our numerical simulations show that the combination of the mode structure of the small-scale instability and its influence on large-scale modes proves to be enough to keep the system in the Dimits regime regardless of κ_T and χ. As κ_T increases beyond the 2D Dimits threshold $\kappa_T^{\mathrm{c,2D}}$, the 2D modes flip the sign of their turbulent momentum flux and start eroding the ZFs, but the small-scale slab-ITG modes are able to provide enough ZF drive in order to maintain the Dimits state (see figure 4.13). Figure 4.14a illustrates the Dimits saturation mechanism.

However, the small-scale slab-ITG modes are able to maintain the Dimits state only if the 3D system is “3D enough”. Namely, if we restrict the system in z by either squeezing it with a small $L_{\parallel}$ or by underresolving large-$k_{\parallel}$ modes, we can break the ZF-dominated Dimits regime and enter a stongly turbulent state. The former of these methods can be deemed “physical” in the sense that real systems can be geometrically limited along the magnetic field. The latter is a numerical artifact, but an important one that deserves our attention nonetheless. Regardless of how the Dimits state is broken, amplitudes remain finite. The small-scale instability is able to extract energy efficiently from the large-scale ($k_{\perp} \lesssim 1$) fields, into which the curvature-ITG inputs energy, and dump it into the small scales $k_{\perp} > 1$ of the slab-ITG instability whence it cascades to smaller scales, where dissipation can take it out of the system. Figure 4.14b show the flow of energy in the strongly turbulent state.

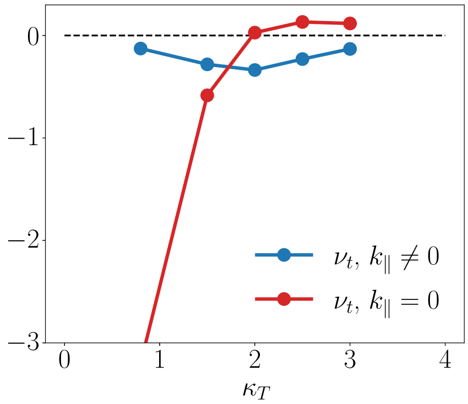


Figure 4.13: Dependence of the turbulent viscosity ν_t, defined by (3.32), on the temperature gradient κ_T for $\chi = 0.1$ and $L_\parallel = 1$. The 2D Dimits regime ends at $\kappa_T^{c,\mathrm{2D}} \approx 1$. In 3D simulations, the 2D modes reverse their turbulent viscosity (red) somewhat after $\kappa_T^{c,\mathrm{2D}}$, which is consistent with a reduction in $\langle T_{\boldsymbol{k}} \rangle_\parallel / \langle \varphi_{\boldsymbol{k}} \rangle_\parallel$ brought about by the additional thermal diffusion on the large-scale temperature. However, the 3D slab-ITG modes continue to feed the ZFs through a negative turbulent viscosity (blue). The data is taken from simulations with fixed ZF profiles.

4.3.1 Varying $L_\parallel$

Figure 4.15 shows a typical example of the dependence of the saturated turbulent heat flux Q on the parallel size of the box $L_\parallel$ for parameters κ_T and χ that lie beyond the 2D Dimits regime. For such parameters, the $L_\parallel = 0$ system does not reach finite-amplitude saturation. For $L_\parallel$ large enough, a Dimits state is established. Here Q is independent of $L_\parallel$, just as it was for parameters that were within the 2D Dimits threshold, see §4.1. As $L_\parallel$ is decreased, the ZFs break up and the system enters a strongly turbulent state. In figure 4.15, this happens for $L_\parallel < 1$. As $L_\parallel$ approaches 0, Q starts to increase rapidly, signifying the approach to the 2D state, where a blow up occurs.

Therefore, for each set of κ_T and χ, there exists a critical $L_\parallel^\mathrm{c}$ such that the system is in the Dimits state for $L_\parallel > L_\parallel^\mathrm{c}$ and in the strongly turbulent regime for $L_\parallel < L_\parallel^\mathrm{c}$. It is clear that $L_\parallel^\mathrm{c} = 0$ if $\kappa_T < \kappa_T^{c,\mathrm{2D}}$, i.e., if κ_T and χ are such that the 2D system is able to reach saturation. However, the dependence of $L_\parallel^\mathrm{c}$ on κ_T and χ for $\kappa_T > \kappa_T^{c,\mathrm{2D}}$ is not known at this point. This is largely due to the numerical cost of trying to resolve simultaneously both the large $k_\parallel$ of the small-scale modes and the box-sized $k_\parallel \sim L_\parallel^{-1}$.

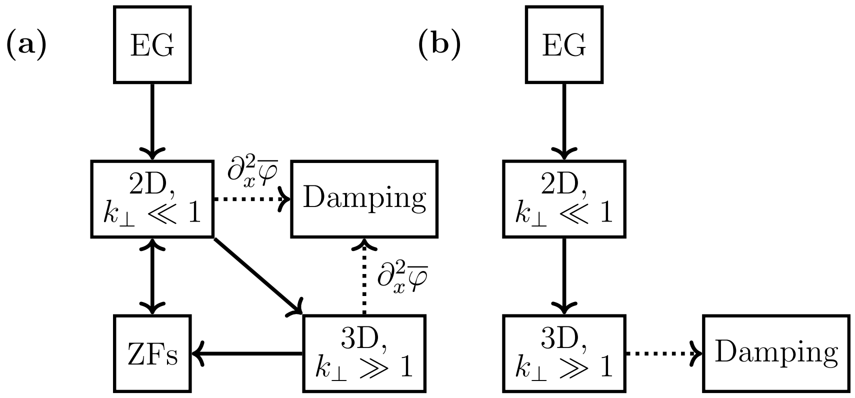


Figure 4.14: An illustration of the flow of energy in **(a)** the Dimits regime and **(b)** strong turbulence. In the Dimits regime, the equilibrium gradients (EG) inject energy into large-scale modes via the curvature-ITG instability. These can then inject energy into the ZFs via the secondary instability (see §3.1) and into small-scale perturbations via the parasitic slab-ITG instability (see §4.2). In the 2D Dimits regime ($\kappa_T < \kappa_T^{c,2D}$), the 2D large-scale modes generate a negative turbulent viscosity on the ZFs and hence reinforce the Dimits state. For ($\kappa_T > \kappa_T^{c,2D}$), the 2D modes erode the ZFs, i.e., energy flows from the ZFs to the 2D modes (hence the double-sided arrow connecting them). In contrast, the small-scale modes always drive the ZFs (see §4.2.4). Saturation is achieved by the action of the zonal shear $\partial_x^2\overline{\varphi}$ on turbulence. On the other hand, if a Dimits state cannot be achieved, the energy injected into the large-scale modes is transferred to small scales via the slab-ITG instability whence it cascades to even smaller, linearly stable scales where it is taken out of the system. The ZFs are not involved in this process.

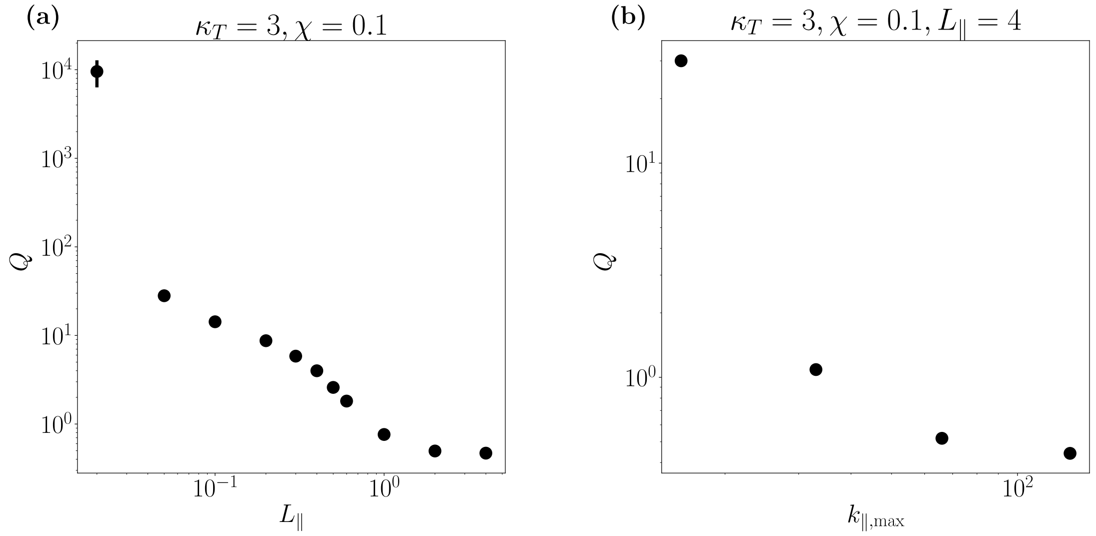


Figure 4.15: Dependence of the saturated turbulent heat flux Q **(a)** on the parallel size of the box $L_\parallel$ and **(b)** on the largest parallel Fourier mode $k_{\parallel,\max}$ that is included in the simulation.

4.3.2 Parallel resolution

The scale separation of the large-scale curvature-ITG modes and the small-scale slab-ITG modes increases the numerical cost of solving (2.18)–(2.20). When the parallel resolution, i.e., the largest $k_\parallel$ in the simulation, is too small, we destroy the Dimits state numerically and push the system into a strong-turbulence regime for parameters for which a Dimits state exists with an appropriate parallel resolution. This is shown in figure 4.15b. Empirically, we have found that a good rule of thumb is "not to chop the leaves" of the instability, i.e., to make sure that the unstable "leaves" at $k_\parallel \sim \kappa_T k_\perp^2$ (see figure 2.3) are all included in the wavenumbers of the simulation.[5] This, however, rapidly increases the numerical cost of the simulations. Recall that the instability cut-off satisfies $k^2_{\perp,\text{cut-off}} \sim \kappa_T/\chi$. Thus, for a fixed χ, the largest unstable $k_\parallel$ is given by $k_{\parallel,\max} \sim \kappa_T^2/\chi$. Therefore, the number of Fourier modes required to properly resolve a simulation scales as κ_T^2, in addition to scaling with $L_\parallel$. This quickly renders simulation efforts futile, even for a fluid code.

---

[5] Of course, this is but a rule of thumb and cannot be entirely accurate. As discussed in §4.2.1, the small-scale instability is driven not by κ_T, but rather by the gradients of the large-scale fields. In other words, the linear 3D modes shown in figure 2.3 are irrelevant for the saturated state. However, we expect that the saturated large-scale temperature perturbations scale with κ_T and so this rule of thumb is a good heuristic guide for setting up simulations.

Chapter 5

Discussion and future directions

We have demonstrated that the low-transport Dimits state of ITG turbulence in a periodic slab, threaded by a constant-curvature magnetic field, is dominated by a quasi-static zonal staircase of strong ZFs with patchwise-constant zonal shear. This zonal staircase suppresses turbulent transport at all radial locations by either shearing the eddies or by a local flattening of the equilibrium temperature gradient that drives turbulence. The mechanism that generates and sustains the ZFs in the Dimits regime and undermines them beyond it is linked to the turbulent momentum flux of ITG modes in the presence of a coherent zonal shear. Namely, in the Dimits regime, the response of ITG turbulence to strong (comparable to the ITG-instability growth rate), coherent zonal shear can be described in terms of a negative turbulent viscosity that reinforces the ZFs. This turbulent viscosity vanishes at the Dimits threshold and becomes positive beyond it, thus impeding any strong zonal shear that could suppress turbulence. Viewed this way, the Dimits transition is caused by a change in the properties of sheared ITG turbulence. In the model considered here, the turbulent momentum flux consists of the usual Reynolds stress, familiar from hydrodynamics, and a diamagnetic contribution. We find that the former acts to reinforce the ZFs, while the latter opposes the ZFs. In general, therefore, determining whether a set of equilibrium parameters lies within the Dimits regime, requires one to make a statement about the combined momentum flux of all turbulent modes. In §3.3.4, we employed the heuristic assumption that the momentum flux is determined predominantly by the most unstable modes with a finite tilt in order to model the Dimits threshold in 2D.

This method is found to work reasonably well in the strongly driven limit, but fails for low values of the equilibrium temperature gradient and of the collisionality.

The mechanism for the Dimits transition described above is not directly tied to the onset of the radially localised tertiary instability found at the extrema of the ZF. It is more appropriate to view it as the quenching of some nonlinear version of the secondary instability. This deviates from the status-quo understanding of the Dimits regime, which was established by Rogers *et al.* (2000), and is also in contrast with the transition seen in the Hasegawa-Wakatani or the Terry-Horton equations, where the tertiary instability was argued to determine the threshold for the strong-turbulence regime (Zhu *et al.*, 2020*a*,*b*). These systems do not contain the diamagnetic stress discussed in §3.3.3 or, indeed, any other form of poloidal momentum flux apart from the Reynolds stress. Thus, by construction, they lack the effects discussed in this paper. This suggests that the term "Dimits transition" is used by the community to denote different things. However, it is worth mentioning that Zhu *et al.* (2020*a*) report turbulent bursts triggered by travelling structures in the Hasegawa-Wakatani system. Due to the lack of diamagnetic stress, these turbulent bursts are bound to restore the ZFs.

As far as we are aware, the recognition that the Dimits state depends on the momentum-transport properties of sheared ITG turbulence is both novel and potentially crucial for understanding these phenomena. Both our simplified fluid model and the close-to-experiment GK simulations of van Wyk *et al.* (2016) report that this sheared turbulence is highly structured and formed of long-lived ferdinons. Thus, such simplified models, which are significantly easier to study both analytically and numerically, could be a viable sandbox model for investigating nonlinear structure formation in ITG turbulence. Additionally, if the averaged properties of sheared turbulence correspond to those of a collection of (independent) ferdinons, we might be able to use the momentum flux of a ferdinon in order to make predictions about sheared turbulence. "Catching" a ferdinon in numerical simulations and attempting an analytical model of it is therefore an important outstanding task for future research.

The details of the 3D Dimits state turned out to be dependent on the peculiar nature of cold-ion physics and the scale separation between large-scale 2D curvature-driven and small-scale 3D slab-ITG modes. As this is only one asymptotic limit of GK, it is difficult to extrapolate any quantitative predictions. However, this does raise the question of whether we truly understand the strongly turbulent saturated state of ITG turbulence. The flow of energy depicted in figure 4.14b is qualitatively distinct from the scenario of critically balance turbulence (Barnes *et al.*, 2011) where energy is injected at a particular outer scale and cascaded down to viscous scales. Unless there is a crucial detail that is missing in our study (e.g., equilibrium flow shear and/or magnetic shear, which are known to have a profound impact on turbulence, see, e.g., Burrell, 1997; Newton *et al.*, 2010), the results presented here suggest that, at least in some parameter regimes, the simple critical-balance picture may be incomplete.

The fact that the Dimits state is governed by essentially the same physical mechanisms in both the 2D and 3D cold-ion Z-pinch systems gives us not only hope that one day we could understand the Dimits regime of full-blown GK, but also a solid foundation of numerical and analytical work upon which to build. To reiterate, we have showed that, within the Dimits regime, the time-averaged poloidal momentum flux of turbulence sheared by a region of constant zonal shear acts to reinforce the ZFs and, thus, zonal shear. On the other hand, beyond the Dimits regime, the overall sign of the momentum flux reverses and the ZFs are destroyed by the turbulence. The validity of this statement is a well-posed question that can be answered by GK numerical simulations. Therefore, armed with the results presented in this work, our natural next step is to investigate the Dimits transition in local GK simulations. Despite the significantly larger number of parameters in realistic GK simulations, we suspect that such an investigation can be carried out in a (qualitatively) similar way to the work presented here. Looking even further into the future, there are some discrepancies in the structure of the local zonal staircase seen in our model and the global $\boldsymbol{E} \times \boldsymbol{B}$ staircase of Dif-Pradalier *et al.* (2010). This is likely due to the fundamental difference between local and global models, viz., that in the former, we impose gradients and solve for fluxes, while in the latter, we impose fluxes and solve for gradients. How our local results translate into global simulations is a difficult question that will be the focus of future work.

Appendix A

The ITG model

A.1 Derivation of the model equations

A.1.1 The gyrokinetic equation

The derivation of our fluid model is similar to the one by Newton *et al.* (2010). We start from the ion gyrokinetic (GK) equation in a Z-pinch-like equilibrium, as discussed in §2.1.1 (for a review of gyrokinetics, see Abel *et al.*, 2013). The ion distribution function is $f_i = F_i + \delta f_i$, where $F_i = n_i/\big(\pi^{3/2} v_{ti}^3\big)\exp(-v^2/v_{ti}^2)$ is the equilibrium ion distribution function and

$$\delta f_i = -\frac{Ze\phi}{T_i}F_i + h\big(t, \boldsymbol{R}, v_\parallel, v_\perp\big). \tag{A.1}$$

The gyrocentre distribution h satisfies

$$\frac{\partial}{\partial t}\left(h - \frac{ZeF_i}{T_i}\langle\phi\rangle_R\right) + v_\parallel\frac{\partial h}{\partial z} + \boldsymbol{V}_D\boldsymbol{\cdot}\frac{\partial h}{\partial \boldsymbol{R}} + \langle\boldsymbol{V}_E\rangle_R\boldsymbol{\cdot}\left[\frac{\partial h}{\partial \boldsymbol{R}} - \hat{\boldsymbol{x}}\left(\frac{v^2}{v_{ti}^2} - \frac{3}{2}\right)\frac{F_i}{L_T}\right] = \langle C_l[h]\rangle_R. \tag{A.2}$$

Here $T_i = m_i v_{ti}^2/2$ and $L_T^{-1} = -\partial_x \ln T_i$ are the ion equilibrium temperature and temperature gradient, respectively. The ion mass is m_i, the ion charge is Ze, and $\Omega_i = ZeB/m_i c$ is the ion gyrofrequency. $\boldsymbol{B}$ is the equilibrium magnetic field and $\langle.\rangle_r$ and $\langle.\rangle_R$ denote, respectively, the gyroaverages at fixed position and fixed ion guiding centre $\boldsymbol{R} = (X, Y, Z) = \boldsymbol{r} - \hat{\boldsymbol{b}}\times\boldsymbol{v}/\Omega_i$, where $\hat{\boldsymbol{b}} = \boldsymbol{B}/B$ is the unit vector parallel to the magnetic field. These gyroaverages are defined as

$$\langle f(\boldsymbol{R})\rangle_r \equiv \langle f(\boldsymbol{r} - \boldsymbol{\rho}(\theta))\rangle = \int_0^{2\pi}\frac{d\theta}{2\pi}\, f(\boldsymbol{r} - \boldsymbol{\rho}(\theta)), \tag{A.3}$$

$$\langle f(\boldsymbol{r})\rangle_R \equiv \langle f(\boldsymbol{R} + \boldsymbol{\rho}(\theta))\rangle = \int_0^{2\pi}\frac{d\theta}{2\pi}\, f(\boldsymbol{R} + \boldsymbol{\rho}(\theta)), \tag{A.4}$$

where $\langle . \rangle$ denotes the average with respect to the gyroangle θ, $\boldsymbol{\rho}(\theta) \equiv \boldsymbol{b} \times \boldsymbol{v}/\Omega_i$, $\boldsymbol{v} = v_\parallel \hat{\boldsymbol{b}} + v_\perp(\cos\theta \hat{\boldsymbol{y}} - \sin\theta \hat{\boldsymbol{x}})$, and the unit vectors $\{\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{b}}\}$ form a right-handed orthonormal basis, as shown in figure 2.1. We will require the following properties of the gyroaverage:

$$\langle \boldsymbol{\rho} \rangle = 0, \qquad \langle \boldsymbol{\rho}\boldsymbol{\rho} \rangle = \frac{1}{2}\frac{v_\perp^2}{v_{ti}^2}\rho_i^2 \mathbb{1}_\perp, \tag{A.5}$$

where $\mathbb{1}_\perp$ is the identity matrix in the (x, y) plane. Using (A.5) and Taylor expanding (A.3) and (A.4), we obtain

$$\begin{aligned} \langle h \rangle_r &= \left[1 + \frac{1}{4}\frac{v_\perp^2}{v_{ti}^2}\rho_i^2 \nabla_\perp^2 + \mathcal{O}\big(k_\perp^4 \rho_i^4\big)\right] h(\boldsymbol{r}), && \text{(A.6)} \\ \langle \varphi \rangle_R &= \left[1 + \frac{1}{4}\frac{v_\perp^2}{v_{ti}^2}\rho_i^2 \nabla_\perp^2 + \mathcal{O}\big(k_\perp^4 \rho_i^4\big)\right] \varphi(\boldsymbol{R}). && \text{(A.7)} \end{aligned}$$

The velocities $\boldsymbol{V}_D$ and $\langle \boldsymbol{V}_E \rangle_R$ in (A.2) are the magnetic and (gyroaveraged) $\boldsymbol{E} \times \boldsymbol{B}$ drifts, respectively, given by

$$\begin{aligned} \boldsymbol{V}_D &= \frac{1}{\Omega}\left[v_\parallel^2 \hat{\boldsymbol{b}} \times (\boldsymbol{b} \cdot \nabla \boldsymbol{b}) + \frac{v_\perp^2}{2}\boldsymbol{b} \times \nabla \ln B\right] = \frac{1}{\Omega_i B}\left(v_\parallel^2 + \frac{v_\perp^2}{2}\right)\hat{\boldsymbol{b}} \times \boldsymbol{\nabla}_\perp B, && \text{(A.8)} \\ \langle \boldsymbol{V}_E \rangle_R &= \frac{c}{B}\hat{\boldsymbol{b}} \times \frac{\partial \langle \phi \rangle_R}{\partial \boldsymbol{R}}. && \text{(A.9)} \end{aligned}$$

The second equality in (A.8) is obtained by assuming that the magnetic field is created by currents external to the spatial domain, so $\boldsymbol{\nabla} \times \boldsymbol{B} = 0$. For the Z-pinch geometry and coordinates discussed in chapter 2 and shown in figure 2.1, we obtain

$$\boldsymbol{V}_D = -\frac{\rho_i v_{ti}}{L_B}\left(\frac{v_\parallel^2}{v_{ti}^2} + \frac{v_\perp^2}{2v_{ti}^2}\right)\hat{\boldsymbol{y}}. \tag{A.10}$$

The term on the right-hand side of (A.2) is the gyroaveraged linearised Landau collision operator. To lowest order in the mass-ratio expansion ($m_e/m_i \ll 1$), only ion-ion collisions contribute and the linearised operator is given by

$$C_l[h] = \frac{\nu_i v_{ti}^3}{n_i}\left.\frac{\partial}{\partial \boldsymbol{v}}\right|_{\boldsymbol{r}} \cdot \left\{F_i(\boldsymbol{v}) \int d^3\, \boldsymbol{v}' F_i(\boldsymbol{v}') \mathsf{U} \cdot \left[\left.\frac{\partial}{\partial \boldsymbol{v}}\right|_{\boldsymbol{r}} \frac{h(\boldsymbol{v})}{F_i(\boldsymbol{v})} - \left.\frac{\partial}{\partial \boldsymbol{v}'}\right|_{\boldsymbol{r}} \frac{h(\boldsymbol{v}')}{F_i(\boldsymbol{v}')}\right]\right\}, \tag{A.11}$$

where $\mathsf{U} = (\boldsymbol{u}^2 \mathsf{I} - \boldsymbol{u}\boldsymbol{u})/u^3$, $\boldsymbol{u} = \boldsymbol{v} - \boldsymbol{v}'$, and

$$\nu_i = \frac{2\pi Z^4 e^4 n_i \ln \Lambda_i}{m_i^2 v_{ti}^3} \tag{A.12}$$

is the ion-ion collision frequency, where $\ln \Lambda_i$ is the Coulomb logarithm (Helander & Sigmar, 2002).

The electrostatic gyrokinetic equation (A.2) is closed by the quasineutrality condition $\delta n_e = Z\delta n_i$ and the modified adiabatic electron response (see §2.1.2 and Abel & Cowley, 2013)

$$\delta f_e = \frac{e\phi'}{T_e} F_e, \tag{A.13}$$

where the nonzonal electrostatic potential φ' is defined in (2.6). Therefore,

$$\frac{1}{n_i} \int d^3\boldsymbol{v}\ \langle h \rangle_r = \frac{Ze}{T_i}\phi + \frac{e}{T_e}\phi'. \tag{A.14}$$

Putting all of this together, we arrive at the 2D electrostatic gyrokinetic system in our Z-pinch equilibrium:

$$\begin{aligned}\frac{\partial}{\partial t}(h - \langle \varphi \rangle_R F_i) + v_\parallel \frac{\partial h}{\partial z} + \frac{\rho_i v_{ti}}{2L_T}\left(\frac{v^2}{v_{ti}^2} - \frac{3}{2}\right) F_i \frac{\partial \langle \varphi \rangle_R}{\partial Y} - \frac{\rho_i v_{ti}}{L_B}\left(\frac{v_\parallel^2}{v_{ti}^2} + \frac{v_\perp^2}{2v_{ti}^2}\right)\frac{\partial h}{\partial Y} \\ + \frac{1}{2}\rho_i v_{ti}\{\langle \varphi \rangle_R, h\} = \langle C_l[h] \rangle_R,\end{aligned} \tag{A.15}$$

$$\frac{1}{n_i} \int d^3\boldsymbol{v}\ \langle h \rangle_r = \varphi + \tau\varphi', \tag{A.16}$$

where the normalised (to ion units) electric potential is $\varphi = Ze\phi/T_i$ and the temperature ratio is $\tau = T_i/ZT_e$.

A.1.2 Lowest-order solution

We now apply the high-collisionality, long-wavelength, cold-ion ordering (2.8). In this expansion, we write $h = h^{(0)} + h^{(1/2)} + h^{(1)}$, where $\mathcal{O}(h^{(1)}) \sim \mathcal{O}(\sqrt{\tau} h^{(1/2)}) \sim \mathcal{O}(\tau h^{(0)})$. Then, to lowest order, equation (A.15) gives $C_l[h^{(0)} + h^{(1/2)}] = 0$, whose solution is a perturbed Maxwellian

$$h^{(0)}(\boldsymbol{R}) + h^{(1/2)}(\boldsymbol{R}) = \left[\frac{\delta N(\boldsymbol{R})}{n_i} + \frac{\delta T(\boldsymbol{R})}{T_i}\left(\frac{v^2}{v_{ti}^2} - \frac{3}{2}\right) + \frac{2v_\parallel U_\parallel(\boldsymbol{R})}{v_{ti}^2}\right] F_i. \tag{A.17}$$

Here the abstract quantities $\delta N(\boldsymbol{R})$, $\delta T(\boldsymbol{R})$, and $U_\parallel(\boldsymbol{R})$ are taken to be functions of the guiding centre $\boldsymbol{R}$. The ordering $U_\parallel \sim \tau c_s \varphi$, which we established using (2.8), implies that

$$\frac{2v_\parallel U_\parallel}{v_{ti}^2} \sim \frac{\tau c_s \varphi}{v_{ti}} \sim \sqrt{\tau}\varphi. \tag{A.18}$$

Therefore, the perturbed parallel flow does not enter into $h^{(0)}$. We define the solution for the distribution function to the two lowest orders as

$$h^{(0)} = \left[\frac{\delta N}{n_i} + \frac{\delta T}{T_i}\left(\frac{v^2}{v_{ti}^2} - \frac{3}{2}\right)\right] F_i, \tag{A.19}$$

$$h^{(1/2)} = \frac{2v_\parallel U_\parallel}{v_{ti}^2} F_i. \tag{A.20}$$

We also impose the following solubility conditions on $h^{(1)}$:

$$\int d^3\boldsymbol{v}\ h^{(1)} = \int d^3\boldsymbol{v}\ v^2 h^{(1)} = \int d^3\boldsymbol{v}\ v_\parallel h^{(1)} = 0. \tag{A.21}$$

These effectively absorb all density, temperature, and parallel-flow perturbations into the lowest-order solutions $h^{(0)}$ and $h^{(1/2)}$.

Substituting (A.17) into the quasineutrality equation (A.16) and using (A.6), we obtain

$$\frac{1}{n_i}\int d^3\boldsymbol{v}\ \left(1 + \frac{1}{4}\frac{v_\perp^2}{v_{ti}^2}\rho_i^2\nabla_\perp^2\right)\left[\frac{\delta N}{n_i} + \frac{\delta T}{T_i}\left(\frac{v^2}{v_{ti}^2} - \frac{3}{2}\right)\right] F_i = \varphi + \tau\varphi', \tag{A.22}$$

where we have absorbed the density and temperature moments of h into $h^{(0)}$ by imposing (A.21).

Formally, in writing down (A.22), we have only assumed that the density moment of $h^{(1)}$ vanishes. We will use the condition that $h^{(1)}$ has a zero temperature moment in appendix A.2.

Under the ordering $\tau \sim k_\perp^2\rho_i^2 \ll 1$, (A.22) yields

$$\frac{\delta N}{n_i} = \varphi + \tau\varphi' - \frac{1}{4}\rho_i^2\nabla_\perp^2\left(\varphi + \frac{\delta T}{T_i}\right) + \mathcal{O}\big(k_\perp^4\rho_i^4\varphi\big). \tag{A.23}$$

We now proceed to take density and temperature moments of the GK equation (A.15) at fixed particle position $\boldsymbol{r}$ and retain only the lowest-order terms in the ordering (2.8). We will find that the density moment of (A.15) vanishes to lowest order. Hence we are required to expand that moment to order $\mathcal{O}(k_\perp^2\rho_i^2 h)$. We shall only require terms up to $\mathcal{O}(h)$ for the temperature moment.

A.1.3 Density moment

Let us consider the density moment at fixed particle position, $(1/n_i)\int d^3\boldsymbol{v}\ \langle\cdot\rangle_r$, of the terms in (A.15). The first term is

$$\frac{1}{n_i}\int d^3\boldsymbol{v}\ \langle h - \langle\varphi\rangle_R F_i\rangle_r = \varphi + \tau\varphi' - \left(1 + \frac{1}{2}\rho_i^2\nabla_\perp^2\right)\varphi = \tau\varphi' - \frac{1}{2}\rho_i^2\nabla_\perp^2\varphi, \tag{A.24}$$

where we have used quasineutrality (A.16) and the lowest-nontrivial-order expressions for the gyroaverages (A.6) and (A.7).

The parallel-velocity moment is, using (A.20),

$$\frac{1}{n_i}\int d^3\boldsymbol{v}\ v_\parallel\partial_\parallel\left\langle h^{(1/2)}\right\rangle_r \approx \partial_\parallel U_\parallel. \tag{A.25}$$

The next term is

$$\begin{aligned}&\frac{1}{n_i}\int d^3\boldsymbol{v}\ \frac{\rho_i v_{ti}}{2L_T}\left(\frac{v^2}{v_{ti}^2} - \frac{3}{2}\right)F_i\left\langle\frac{\partial\langle\varphi\rangle_R}{\partial Y}\right\rangle_r\\ &\approx \frac{1}{n_i}\int d^3\boldsymbol{v}\ \frac{\rho_i v_{ti}}{2L_T}\left(\frac{v^2}{v_{ti}^2} - \frac{3}{2}\right)F_i\left(1 + \frac{1}{2}\frac{v_\perp^2}{v_{ti}^2}\rho_i^2\nabla_\perp^2\right)\frac{\partial\varphi}{\partial y}\\ &= \frac{\rho_i v_{ti}}{2L_T}\frac{1}{2}\rho_i^2\nabla_\perp^2\frac{\partial\varphi}{\partial y}.\end{aligned} \tag{A.26}$$

To lowest order in $k_\perp^2\rho_i^2 \ll 1$, the magnetic-drift term gives

$$\begin{aligned}&\frac{1}{n_i}\int d^3\boldsymbol{v}\ \frac{\rho_i v_{ti}}{L_B}\left(\frac{v_\parallel^2}{v_{ti}^2} + \frac{v_\perp^2}{2v_{ti}^2}\right)\left\langle\frac{\partial h(\boldsymbol{R})}{\partial Y}\right\rangle_r\\ &\approx \frac{1}{n_i}\int d^3\boldsymbol{v}\ \frac{\rho_i v_{ti}}{L_B}\left(\frac{v_\parallel^2}{v_{ti}^2} + \frac{v_\perp^2}{2v_{ti}^2}\right)\frac{\partial h^{(0)}(\boldsymbol{r})}{\partial y}\\ &= \frac{1}{n_i}\int d^3\boldsymbol{v}\ \frac{\rho_i v_{ti}}{L_B}\frac{2}{3}\frac{v^2}{v_{ti}^2}\frac{\partial h^{(0)}(\boldsymbol{r})}{\partial y} = \frac{\rho_i v_{ti}}{L_B}\frac{\partial}{\partial y}\left(\varphi + \frac{\delta T}{T_i}\right),\end{aligned} \tag{A.27}$$

where we have used

$$\int dv\ v_\parallel^2 h^{(0)} = \int dv\ \frac{v_\perp^2}{2}h^{(0)} = \int dv\ \frac{v^2}{3}h^{(0)}, \tag{A.28}$$

which is a consequence of the isotropic form (A.19) of $h^{(0)}$.

The density moment of the nonlinear term in (A.15) is

$$
\begin{aligned}
&\frac{1}{n_i}\int d^3\boldsymbol{v}\ \langle\{\langle\varphi\rangle_R(\boldsymbol{R}), h(\boldsymbol{R})\}\rangle_r \\
&\approx \frac{1}{n_i}\int d^3\boldsymbol{v}\ \left\langle\left\{\varphi(\boldsymbol{R}) + \frac{1}{4}\frac{v_\perp^2}{v_{ti}^2}\rho_i^2\nabla_\perp^2\varphi(\boldsymbol{R}), h(\boldsymbol{R})\right\}\right\rangle_r \\
&\approx \frac{1}{n_i}\int d^3\boldsymbol{v}\ \left\langle\left\{\varphi(\boldsymbol{r}) - \boldsymbol{\rho}\cdot\boldsymbol{\nabla}_\perp\varphi(\boldsymbol{r}) + \frac{1}{2}\boldsymbol{\rho\rho} : \boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\varphi(\boldsymbol{r}) + \frac{1}{4}\frac{v_\perp^2}{v_{ti}^2}\rho_i^2\nabla_\perp^2\varphi(\boldsymbol{r}), h(\boldsymbol{r}-\boldsymbol{\rho})\right\}\right\rangle_r \\
&= \frac{1}{n_i}\int d^3\boldsymbol{v}\ \Big\langle \{\varphi(\boldsymbol{r}), h(\boldsymbol{r}-\boldsymbol{\rho})\} + \boldsymbol{\rho\rho} : \{\boldsymbol{\nabla}_\perp\varphi(\boldsymbol{r}), \boldsymbol{\nabla}_\perp h(\boldsymbol{r})\} \\
&\qquad\qquad + \frac{1}{2}\boldsymbol{\rho\rho} : \{\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\varphi(\boldsymbol{r}), h(\boldsymbol{r})\} + \frac{1}{4}\frac{v_\perp^2}{v_{ti}^2}\big\{\rho_i^2\nabla_\perp^2\varphi(\boldsymbol{r}), h(\boldsymbol{r})\big\}\Big\rangle_r \\
&= \frac{1}{n_i}\int d^3\boldsymbol{v}\ \Big[\{\varphi(\boldsymbol{r}), \langle h\rangle_r(\boldsymbol{r})\} + \langle\boldsymbol{\rho\rho}\rangle_r : \{\boldsymbol{\nabla}_\perp\varphi(\boldsymbol{r}), \boldsymbol{\nabla}_\perp h(\boldsymbol{r})\} \\
&\qquad\qquad + \frac{1}{2}\langle\boldsymbol{\rho\rho}\rangle_r : \{\boldsymbol{\nabla}_\perp\boldsymbol{\nabla}_\perp\varphi(\boldsymbol{r}), h(\boldsymbol{r})\} + \frac{1}{4}\frac{v_\perp^2}{v_{ti}^2}\big\{\rho_i^2\nabla_\perp^2\varphi(\boldsymbol{r}), h(\boldsymbol{r})\big\}\Big] \\
&= \frac{1}{n_i}\int d^3\boldsymbol{v}\ \Big[\{\varphi(\boldsymbol{r}), \langle h\rangle_r(\boldsymbol{r})\} + \frac{1}{2}\frac{v_\perp^2}{v_{ti}^2}\rho_i^2\mathbb{1}_\perp : \{\boldsymbol{\nabla}_\perp\varphi(\boldsymbol{r}), \boldsymbol{\nabla}_\perp h(\boldsymbol{r})\} \\
&\qquad\qquad + \frac{1}{2}\frac{v_\perp^2}{v_{ti}^2}\big\{\rho_i^2\nabla_\perp^2\varphi(\boldsymbol{r}), h(\boldsymbol{r})\big\}\Big],
\end{aligned}
\tag{A.29}
$$

where we have used (A.5). Using the lowest-order contribution (A.19) to h and the fact that $\{g, g\} = 0$ for any g, (A.29) becomes

$$
\begin{aligned}
&\{\varphi, \tau\varphi'\} + \frac{1}{2}\rho_i^2\mathbb{1}_\perp : \left\{\boldsymbol{\nabla}_\perp\varphi, \boldsymbol{\nabla}_\perp\frac{\delta T}{T_i}\right\} + \frac{1}{2}\rho_i^2\left\{\nabla_\perp^2\varphi, \varphi + \frac{\delta T}{T_i}\right\} \\
&= \left\{\varphi, \tau\varphi' - \frac{1}{2}\rho_i^2\nabla_\perp^2\varphi\right\} + \frac{1}{2}\rho_i^2\boldsymbol{\nabla}_\perp\cdot\left\{\boldsymbol{\nabla}_\perp\varphi, \frac{\delta T}{T_i}\right\}.
\end{aligned}
\tag{A.30}
$$

Finally, collecting terms, dividing by τ, and introducing $\rho_s = \rho_i/\sqrt{2\tau}$, we obtain

$$
\begin{aligned}
&\frac{\partial}{\partial t}\big(\varphi' - \rho_s^2\nabla_\perp^2\varphi\big) - \frac{\rho_i v_{ti}}{\tau L_B}\frac{\partial}{\partial y}\left(\varphi + \frac{\delta T}{T_i}\right) + \frac{\rho_i v_{ti}}{2L_T}\frac{\partial}{\partial y}\big(\rho_s^2\nabla_\perp^2\varphi\big) \\
&\quad + \frac{1}{2}\rho_i v_{ti}\left(\big\{\varphi, \varphi' - \rho_s^2\nabla_\perp^2\varphi\big\} + \rho_s^2\boldsymbol{\nabla}_\perp\cdot\left\{\boldsymbol{\nabla}_\perp\varphi, \frac{\delta T}{T_i}\right\}\right) = \frac{1}{\tau n_i}\int d^3\boldsymbol{v}\ \langle\langle C_l[h]\rangle_R\rangle_r.
\end{aligned}
\tag{A.31}
$$

This will become (2.9) after we calculate the collisional term in appendix A.1.6.3.

A.1.4 Temperature moment

In a similar way, let us consider the temperature moments, $(1/n_i)\int d^3\boldsymbol{v} v^2/v_{ti}^2\langle\cdot\rangle_r$, of the terms in (A.15) to lowest order in $k_\perp^2\rho_i^2 \ll 1$. The first term is

$$\frac{1}{n_i}\int d^3\boldsymbol{v}\ \frac{v^2}{v_{ti}^2}\langle h - \langle\varphi\rangle_R F_i\rangle_r \approx \frac{1}{n_i}\int d^3\boldsymbol{v}\ \frac{v^2}{v_{ti}^2}\big(h^{(0)} - \varphi F_i\big) = \frac{3}{2}\frac{\delta T}{T_i}, \tag{A.32}$$

where we have used (A.19).

The parallel-streaming term is

$$\frac{1}{n_i}\int d^3\boldsymbol{v} v_\parallel \partial_\parallel \frac{v^2}{v_{ti}^2}\big\langle h^{(1/2)}\big\rangle_r = \frac{5}{2}\partial_\parallel U_\parallel. \tag{A.33}$$

The temperature-gradient term is

$$\begin{aligned}&\frac{1}{n_i}\int d^3\boldsymbol{v}\ \frac{\rho_i v_{ti}}{2L_T}\frac{v^2}{v_{ti}^2}\left(\frac{v^2}{v_{ti}^2} - \frac{3}{2}\right)F_i\left\langle\frac{\partial\langle\varphi\rangle_R}{\partial Y}\right\rangle_r \\ &\approx \frac{1}{n_i}\int d^3\boldsymbol{v}\ \frac{\rho_i v_{ti}}{2L_T}\frac{v^2}{v_{ti}^2}\left(\frac{v^2}{v_{ti}^2} - \frac{3}{2}\right)F_i\frac{\partial\varphi}{\partial y} = \frac{3}{2}\frac{\rho_i v_{ti}}{2L_T}\frac{\partial\varphi}{\partial y}.\end{aligned} \tag{A.34}$$

Note that under the ordering (2.8), the parallel-streaming term is $\mathcal{O}(\tau)$ smaller than the temperature-gradient term, viz.,

$$\partial_\parallel U_\parallel \sim \frac{\tau c_s}{L_B} \ll \frac{\rho_i v_{ti}}{2L_T}\frac{\partial\varphi}{\partial y} \sim \frac{\tau c_s}{L_T} \sim \frac{\kappa_T c_s}{L_B}. \tag{A.35}$$

The magnetic-drift term is

$$\begin{aligned}&\frac{1}{n_i}\int d^3\boldsymbol{v}\ \frac{\rho_i v_{ti}}{L_B}\frac{v^2}{v_{ti}^2}\left(\frac{v_\parallel^2}{v_{ti}^2} + \frac{v_\perp^2}{2v_{ti}^2}\right)\left\langle\frac{\partial h(\boldsymbol{R})}{\partial Y}\right\rangle_r \\ &\approx \frac{1}{n_i}\int d^3\boldsymbol{v}\ \frac{\rho_i v_{ti}}{L_B}\frac{v^2}{v_{ti}^2}\left(\frac{v_\parallel^2}{v_{ti}^2} + \frac{v_\perp^2}{2v_{ti}^2}\right)\frac{\partial h^{(0)}(\boldsymbol{r})}{\partial y} \\ &= \frac{1}{n_i}\int d^3\boldsymbol{v}\ \frac{\rho_i v_{ti}}{L_B}\frac{2}{3}\frac{v^4}{v_{ti}^4}\frac{\partial h^{(0)}(\boldsymbol{r})}{\partial y} = \frac{5}{2}\frac{\rho_i v_{ti}}{L_B}\frac{\partial}{\partial y}\left(\varphi + 2\frac{\delta T}{T_i}\right),\end{aligned} \tag{A.36}$$

where we have used the isotropy of $h^{(0)}$ again. Again, by the ordering (2.8), this term is $\mathcal{O}(\tau)$ smaller than the temperature-gradient term (A.34). Hence it will not contribute to the final expression for the temperature moment of (A.15).

The nonlinear term is

$$\frac{1}{n_i}\int d^3\boldsymbol{v}\ \frac{v^2}{v_{ti}^2}\langle\{\langle\varphi\rangle_R, h\}\rangle_r \approx \frac{1}{n_i}\int d^3\boldsymbol{v}\ \frac{v^2}{v_{ti}^2}\{\varphi, h^{(0)}\} = \frac{3}{2}\left\{\varphi, \frac{\delta T}{T_i}\right\}. \tag{A.37}$$

Collecting terms, we find that the temperature moment of (A.15) is

$$\frac{\partial}{\partial t}\frac{\delta T}{T_i} + \frac{\rho_i v_{ti}}{2L_T}\frac{\partial \varphi}{\partial y} + \frac{1}{2}\rho_i v_{ti}\left\{\varphi, \frac{\delta T}{T_i}\right\} = \frac{2}{3n_i}\int d^3\boldsymbol{v}\ \frac{v^2}{v_{ti}^2}\langle\langle C_l[h]\rangle_R\rangle_r. \tag{A.38}$$

This will become (2.10) after we calculate the collisional term in appendix A.1.6.2.

A.1.5 Parallel-velocity moment

Finally, we take the parallel-velocity moment, $(1/n_i)\int d^3\boldsymbol{v}\ v_\parallel\langle\cdot\rangle_r$, of (A.15). The first term is the time derivative of

$$\frac{1}{n_i}\int d^3\boldsymbol{v}\ v_\parallel\langle h - \langle\varphi\rangle_R F_i\rangle_r \approx \frac{1}{n_i}\int d^3\boldsymbol{v}\ v_\parallel h^{(1/2)} = U_\parallel. \tag{A.39}$$

The parallel-streaming term is

$$\frac{1}{n_i}\int d^3\boldsymbol{v}\ v_\parallel^2\partial_\parallel\langle h\rangle_r \approx \frac{1}{n_i}\int d^3\boldsymbol{v}\ v_\parallel^2\partial_\parallel h^{(0)} = \frac{1}{2}v_{ti}^2\partial_\parallel(\varphi + T). \tag{A.40}$$

The temperature-gradient term integrates to 0 because the integrand is odd in $v_\parallel$. The magnetic-gradient term is one order of $L_T/L_B \sim \mathcal{O}(\tau) \ll 1$ smaller than the rest (the magnetic curvature is absent from (A.38) for the same reason). The nonlinear term integrates to

$$\frac{1}{n_i}\int d^3\boldsymbol{v}\ v_\parallel\langle\{\langle\varphi\rangle_R, h\}\rangle_r \approx \frac{1}{n_i}\int d^3\boldsymbol{v}\ v_\parallel\{\varphi, h^{(1/2)}\} = \frac{1}{2}\rho_i v_{ti}\{\varphi, U_\parallel\}. \tag{A.41}$$

Finally, the parallel-velocity moment of the collisional operator is

$$\frac{1}{n_i}\int d^3\boldsymbol{v}\ v_\parallel\langle\langle C_l[h]\rangle_R\rangle_r \approx \frac{1}{n_i}\int d^3\boldsymbol{v}\ v_\parallel C_l[h^{(1/2)}] = d\chi\nabla_\perp^2 U_\parallel, \tag{A.42}$$

where $d = 9/10$ is a numerical factor (see Newton *et al.*, 2010). Putting together (A.39)–(A.42), we arrive at (2.11).

A.1.6 Moments of the collision operator

A.1.6.1 Gyroaveraged collision operator

The collision term in (A.2), expanded in a Fourier basis, $h(\boldsymbol{R}) = \sum_{\boldsymbol{k}} h_{\boldsymbol{k}} e^{i\boldsymbol{k}.\boldsymbol{R}}$, is

$$\begin{aligned}\langle C_l[h]\rangle_R &= \sum_{\boldsymbol{k}}\left\langle C_l\left[h_{\boldsymbol{k}}e^{i\boldsymbol{k}\cdot\boldsymbol{R}}\right]\right\rangle_R = \sum_{\boldsymbol{k}}\left\langle C_l\left[h_{\boldsymbol{k}}e^{-i\boldsymbol{k}\cdot\boldsymbol{\rho}}\right]e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\right\rangle_R \\ &= \sum_{\boldsymbol{k}}\left\langle C_l\left[h_{\boldsymbol{k}}e^{-i\boldsymbol{k}\cdot\boldsymbol{\rho}}\right]e^{i\boldsymbol{k}\cdot\boldsymbol{\rho}}\right\rangle e^{i\boldsymbol{k}\cdot\boldsymbol{R}},\end{aligned} \tag{A.43}$$

where all derivatives and integrals with respect to $\boldsymbol{v}$ in the collision operator are taken at fixed $\boldsymbol{r}$, and the $\langle . \rangle$ operation is the gyroangle average, see (A.3) and (A.4). In order to obtain collisional terms in (A.31) and (A.38), we expand the exponential factors $e^{\pm i\boldsymbol{k}\cdot\boldsymbol{\rho}}$ in the small quantity $\boldsymbol{k}\cdot\boldsymbol{\rho}$. Recall that (A.31) and (A.38) are contained at different orders in the GK equation. We need (A.43) to order $\mathcal{O}(k_\perp^2\rho_i^2\nu_i h)$ for the temperature moment and to order $\mathcal{O}(k_\perp^4\rho_i^4\nu_i h)$ for the density moment. We first consider the collisional term in the temperature equation (A.38), which represents thermal diffusion, before turning to the more involved calculation for the viscosity in (A.31).

A.1.6.2 Collisional thermal diffusion

The collisional term in the temperature equation (A.38) is

$$\begin{aligned}&\frac{2}{3n_i}\left\langle \sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{R}} \int d^3\boldsymbol{v}\ \frac{v^2}{v_{ti}^2}\left\langle C_l\left[h_{\boldsymbol{k}}e^{-i\boldsymbol{k}\cdot\boldsymbol{\rho}}\right]e^{i\boldsymbol{k}\cdot\boldsymbol{\rho}}\right\rangle\right\rangle_r \\ &\approx \frac{2}{3n_i}\sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\int d^3\boldsymbol{v}\ \frac{v^2}{v_{ti}^2}\left\langle C_l\left[h_{\boldsymbol{k}}e^{-i\boldsymbol{k}\cdot\boldsymbol{\rho}}\right]e^{i\boldsymbol{k}\cdot\boldsymbol{\rho}}\right\rangle,\end{aligned} \tag{A.44}$$

where, to order $\mathcal{O}(k_\perp^2\rho_i^2\nu_i h)$, we can expand

$$\begin{aligned}\left\langle C_l\left[h_{\boldsymbol{k}}e^{-i\boldsymbol{k}\cdot\boldsymbol{\rho}}\right]e^{i\boldsymbol{k}\cdot\boldsymbol{\rho}}\right\rangle &= C_l\left[h_{\boldsymbol{k}}^{(1)}\right] + \left\langle i\boldsymbol{k}\cdot\boldsymbol{\rho}C_l\left[(-i\boldsymbol{k}\cdot\boldsymbol{\rho})h_{\boldsymbol{k}}^{(0)}\right]\right\rangle \\ &-\frac{1}{2}\left\langle(\boldsymbol{k}\cdot\boldsymbol{\rho})^2\right\rangle C_l\left[h_{\boldsymbol{k}}^{(0)}\right] + \left\langle C_l\left[-\frac{1}{2}(\boldsymbol{k}\cdot\boldsymbol{\rho})^2h_{\boldsymbol{k}}^{(0)}\right]\right\rangle + \mathcal{O}(k_\perp^4\rho_i^4\nu_i h).\end{aligned} \tag{A.45}$$

Taking a temperature moment of (A.45) annihilates the first and fourth terms due to the conservation-of-energy property of the collision operator ($\int d^3\boldsymbol{v}\ v^2\ C_l[f] = 0$ for any f) and the third term vanishes because $C_l\left[h^{(0)}\right] = 0$. Just as in Newton *et al.* (2010), after performing the integration in (A.44), we find that the second term of (A.45) gives

$$\int d^3\boldsymbol{v}\ \frac{v^2}{v_{ti}^2}(i\boldsymbol{k}\cdot\boldsymbol{\rho})C_l[(-i\boldsymbol{k}\cdot\boldsymbol{\rho})h_{\boldsymbol{k}}^{(0)}] = -\frac{3}{2}n_i\chi k_\perp^2\frac{\delta T_{\boldsymbol{k}}}{T_i}, \tag{A.46}$$

where

$$\chi \equiv \frac{8}{9}\sqrt{\frac{2}{\pi}}\nu_i\rho_i^2 \tag{A.47}$$

and the ion-ion collision frequency ν_i is defined in (A.12). Thus, (A.44) is

$$-\chi\sum_{\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}}k_\perp^2\frac{\delta T_{\boldsymbol{k}}}{T_i} = \chi\nabla_\perp^2\frac{\delta T(\boldsymbol{r})}{T_i}. \tag{A.48}$$

The right-hand side of (A.38) is (A.48), hence we arrive at the temperature equation (2.10).

A.1.6.3 Collisional viscous damping

As we discussed before, under the ordering (2.8), the density moment of (A.15) is obtained at an order higher in $k_\perp^2\rho_i^2 \ll 1$ than the temperature moment. As a sanity check, note that taking a density moment ($\int d^3\boldsymbol{v}\ \langle . \rangle_r$) of (A.45) annihilates all terms: terms one and four vanish because the collision operator conserves particle number ($\int d^3\boldsymbol{v}\ C_l[f] = 0$ for any f), term two vanishes because of conservation of momentum ($\int d^3\boldsymbol{v}\ \boldsymbol{v}C_l[f] = 0$ for any f), and, as before, term three is identically zero due to the form of $h^{(0)}$. We thus need to expand (A.43) to order $\mathcal{O}(k_\perp^4\rho_i^4\nu_i h)$.

The collisional term of the density moment in (A.31) is

$$\sum_{\boldsymbol{k}}\int d^3\boldsymbol{v}\ \langle\langle C_l\big[h_{\boldsymbol{k}}e^{-i\boldsymbol{k}\cdot\boldsymbol{\rho}}\big]e^{i\boldsymbol{k}\cdot\boldsymbol{\rho}}\rangle e^{i\boldsymbol{k}\cdot\boldsymbol{R}}\rangle_r = \sum_{\boldsymbol{k}}\int d^3\boldsymbol{v}\ J_0\left(\frac{k_\perp v_\perp}{\Omega}\right)\langle C_l\big[h_{\boldsymbol{k}}e^{-i\boldsymbol{k}\cdot\boldsymbol{\rho}}\big]e^{i\boldsymbol{k}\cdot\boldsymbol{\rho}}\rangle e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$$
$$= \sum_{\boldsymbol{k}}\int d^3\boldsymbol{v}\ J_0\left(\frac{k_\perp v_\perp}{\Omega}\right)C_l\big[h_{\boldsymbol{k}}e^{-i\boldsymbol{k}\cdot\boldsymbol{\rho}}\big]e^{i\boldsymbol{k}\cdot\boldsymbol{\rho}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \equiv \sum_{\boldsymbol{k}} C_{D\boldsymbol{k}}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}. \tag{A.49}$$

Note that any term with an odd power of $\boldsymbol{k}\cdot\boldsymbol{\rho}$ is annihilated by gyroaveraging. Expanding (A.49) to order $\mathcal{O}(k_\perp^4\rho_i^4\nu_i h)$, we find

$$\begin{aligned} C_{D\boldsymbol{k}} = \int d^3\boldsymbol{v}\ \Bigg\{ & -\frac{1}{2}(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 C_l\Big[-\frac{1}{2}(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 h_{\boldsymbol{k}}^{(0)}\Big] \\ & -\frac{1}{2}(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 C_l\Big[h_{\boldsymbol{k}}^{(1)}\Big] - \frac{1}{6}i(\boldsymbol{k}\cdot\boldsymbol{\rho})^3 C_l\Big[(-i\boldsymbol{k}\cdot\boldsymbol{\rho})h_{\boldsymbol{k}}^{(0)}\Big] \\ & -\frac{k_\perp^2 v_\perp^2}{4\Omega^2}\left(\boldsymbol{k}\cdot\boldsymbol{\rho}C_l\Big[\boldsymbol{k}\cdot\boldsymbol{\rho}h_{\boldsymbol{k}}^{(0)}\Big] - \frac{1}{2}C_l\Big[(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 h_{\boldsymbol{k}}^{(0)}\Big] + C_l\Big[h_{\boldsymbol{k}}^{(1)}\Big]\right)\Bigg\}. \end{aligned} \tag{A.50}$$

Since the linearised collision operator is isotropic in velocity space and the gyrokinetic distribution h is gyroangle independent, it follows that $C[h^{(1)}]$ is also gyroangle independent. Therefore, we can write

$$\int d^3\boldsymbol{v}\ \frac{1}{2}(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 C_l\Big[h_{\boldsymbol{k}}^{(1)}\Big] = \int d^3\boldsymbol{v}\ \frac{1}{2}\langle(\boldsymbol{k}\cdot\boldsymbol{\rho})^2\rangle C_l\Big[h_{\boldsymbol{k}}^{(1)}\Big] = \int d^3\boldsymbol{v}\ \frac{k_\perp^2 v_\perp^2}{4\Omega^2}C_l\Big[h_{\boldsymbol{k}}^{(1)}\Big]. \tag{A.51}$$

Substituting (A.51) into (A.50), we obtain

$$C_{D\boldsymbol{k}} = \int d^3\boldsymbol{v} \left\{ \frac{1}{4}(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 C_l\Big[(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 h_{\boldsymbol{k}}^{(0)}\Big] - \frac{k_\perp^2 v_\perp^2}{2\Omega^2} C_l\Big[h_{\boldsymbol{k}}^{(1)}\Big] - \frac{1}{6}(\boldsymbol{k}\cdot\boldsymbol{\rho})^3 C_l\Big[\boldsymbol{k}\cdot\boldsymbol{\rho} h_{\boldsymbol{k}}^{(0)}\Big] \right.$$
$$\left. - \frac{k_\perp^2 v_\perp^2}{4\Omega^2}\boldsymbol{k}\cdot\boldsymbol{\rho} C_l\Big[\boldsymbol{k}\cdot\boldsymbol{\rho} h_{\boldsymbol{k}}^{(0)}\Big] + \frac{k_\perp^2 v_\perp^2}{8\Omega^2} C_l\Big[(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 h_{\boldsymbol{k}}^{(0)}\Big] \right\}. \tag{A.52}$$

The high-collisionality limit allowed us to obtain the form of $h^{(0)}$, but $h^{(1)}$ is unknown without inverting the collision operator. We can take advantage of the form of the GK equation (A.15) to lowest non-trivial order and express directly the required moment of $C[h^{(1)}]$ using moments of $C[h^{(0)}]$. Consider the $\int d^3\boldsymbol{v}(v_\perp^2 - 2v^2/3)$ moment of the GK equation up to order $\mathcal{O}(k_\perp^2\rho_i^2\nu_i h)$. In appendix A.1.4, where we took a v^2 moment, we found that all non-collisional terms in the temperature moment of (A.15) are isotropic in velocity space to order $\mathcal{O}(k_\perp^2\rho_i^2\nu_i h)$, hence they are annihilated by the operator $\int d^3\boldsymbol{v}(v_\perp^2 - 2v^2/3)$. Using the expansion (A.45) of the collision operator, we find

$$0 = \int d^3\boldsymbol{v} \left(v_\perp^2 - \frac{2}{3}v^2\right)\left\{\boldsymbol{k}\cdot\boldsymbol{\rho} C_l\Big[\boldsymbol{k}\cdot\boldsymbol{\rho} h_{\boldsymbol{k}}^{(0)}\Big] - \frac{1}{2}C_l\Big[(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 h_{\boldsymbol{k}}^{(0)}\Big] + C_l\Big[h_{\boldsymbol{k}}^{(1)}\Big]\right\}$$
$$= \int d^3\boldsymbol{v} \left\{ v_\perp^2 \boldsymbol{k}\cdot\boldsymbol{\rho} C_l\Big[\boldsymbol{k}\cdot\boldsymbol{\rho} h_{\boldsymbol{k}}^{(0)}\Big] - \frac{1}{2}v_\perp^2 C_l\Big[(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 h_{\boldsymbol{k}}^{(0)}\Big] \right.$$
$$\left. + v_\perp^2 C_l\Big[h_{\boldsymbol{k}}^{(1)}\Big] - \frac{2}{3}v^2 \boldsymbol{k}\cdot\boldsymbol{\rho} C_l\Big[\boldsymbol{k}\cdot\boldsymbol{\rho} h_{\boldsymbol{k}}^{(0)}\Big] \right\}. \tag{A.53}$$

Here a few terms have dropped out due to the energy-conservation properties of the collision operator. Extracting from (A.53) an expression for $v_\perp^2 C_l\Big[h_{\boldsymbol{k}}^{(1)}\Big]$ and substituting this expression into (A.52), we obtain an expression for $C_{D\boldsymbol{k}}$ involving only $h^{(0)}$:

$$C_{D\boldsymbol{k}} = \int d^3\boldsymbol{v} \left\{ \frac{1}{4}(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 C_l\Big[(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 h_{\boldsymbol{k}}^{(0)}\Big] - \frac{1}{6}(\boldsymbol{k}\cdot\boldsymbol{\rho})^3 C_l\Big[\boldsymbol{k}\cdot\boldsymbol{\rho} h_{\boldsymbol{k}}^{(0)}\Big] \right.$$
$$+ \frac{k_\perp^2 v_\perp^2}{4\Omega^2}\boldsymbol{k}\cdot\boldsymbol{\rho} C_l\Big[\boldsymbol{k}\cdot\boldsymbol{\rho} h_{\boldsymbol{k}}^{(0)}\Big] - \frac{k_\perp^2 v^2}{3\Omega^2}\boldsymbol{k}\cdot\boldsymbol{\rho} C_l\Big[\boldsymbol{k}\cdot\boldsymbol{\rho} h_{\boldsymbol{k}}^{(0)}\Big]$$
$$\left. - \frac{k_\perp^2 v_\perp^2}{8\Omega^2} C_l\Big[(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 h_{\boldsymbol{k}}^{(0)}\Big] \right\}. \tag{A.54}$$

We proceed to evaluate these integrals:

$$\int d^3\boldsymbol{v}\ \frac{1}{4}(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 C_l\Big[(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 h_{\boldsymbol{k}}^{(0)}\Big] = -n_i\chi\rho_i^2 k_\perp^4\left(\frac{3}{20}\varphi_{\boldsymbol{k}} + \frac{3}{16}\frac{\delta T_{\boldsymbol{k}}}{T_i}\right), \tag{A.55}$$

$$\int d^3\boldsymbol{v}\ \frac{1}{6}(\boldsymbol{k}\cdot\boldsymbol{\rho})^3 C_l\Big[\boldsymbol{k}\cdot\boldsymbol{\rho} h_{\boldsymbol{k}}^{(0)}\Big] = -n_i\chi\rho_i^2 k_\perp^4\frac{3}{20}\frac{\delta T_{\boldsymbol{k}}}{T_i}, \tag{A.56}$$

$$\int d^3\boldsymbol{v}\ \frac{k_\perp^2 v_\perp^2}{4\Omega^2}(\boldsymbol{k}\cdot\boldsymbol{\rho}) C_l\Big[\boldsymbol{k}\cdot\boldsymbol{\rho} h_{\boldsymbol{k}}^{(0)}\Big] = -n_i\chi\rho_i^2 k_\perp^4\frac{3}{10}\frac{\delta T_{\boldsymbol{k}}}{T_i}, \tag{A.57}$$

$$\int d^3\boldsymbol{v}\ \frac{k_\perp^2 v^2}{3\Omega^2}(\boldsymbol{k}\cdot\boldsymbol{\rho}) C_l\Big[\boldsymbol{k}\cdot\boldsymbol{\rho} h_{\boldsymbol{k}}^{(0)}\Big] = -n_i\chi\rho_i^2 k_\perp^4\frac{1}{2}\frac{\delta T_{\boldsymbol{k}}}{T_i}, \tag{A.58}$$

$$\int d^3\boldsymbol{v}\ \frac{k_\perp^2 v_\perp^2}{8\Omega^2} C_l\Big[(\boldsymbol{k}\cdot\boldsymbol{\rho})^2 h_{\boldsymbol{k}}^{(0)}\Big] = -n_i\chi\rho_i^2 k_\perp^4\left(\frac{3}{80}\varphi_{\boldsymbol{k}} + \frac{3}{64}\frac{\delta T_{\boldsymbol{k}}}{T_i}\right). \tag{A.59}$$

Combining all of these, we get the following expression for the collision term on the right-hand side of (A.31):

$$\sum_{\boldsymbol{k}} C_{D\boldsymbol{k}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = -\sum_{\boldsymbol{k}} n_i\chi\frac{\rho_i^2}{2}k_\perp^4\left(\frac{9}{40}\varphi_{\boldsymbol{k}} - \frac{67}{160}\frac{\delta T_{\boldsymbol{k}}}{T_i}\right)e^{i\boldsymbol{k}\cdot\boldsymbol{r}} = -\frac{1}{2}n_i\chi\rho_i^2\nabla_\perp^4\left(a\varphi - b\frac{\delta T}{T_i}\right), \tag{A.60}$$

where $a = 9/40$ and $b = 67/160$. The right-hand side of (A.31) is (A.60)$/\tau n_i$, hence follows the ion-density equation (2.9).

A.2 2D conservation laws

All calculations in this section are done to order $\mathcal{O}(k_\perp^4\rho_i^4)$ and for $k_\parallel = 0$.

The nonlinearly conserved free energy in electrostatic GK is given by

$$W = \sum_s \int d^3\boldsymbol{r} \int d^3\boldsymbol{v}\ \frac{T_s\delta f_s^2}{2F_s}, \tag{A.61}$$

where s labels the particle species and F_s is the corresponding Maxwellian equilibrium distribution.

Using the modified adiabatic electron response (A.13), the electron contribution to (A.61) is

$$\int d^3\boldsymbol{r} \int d^3\boldsymbol{v}\ \frac{T_e\delta f_e^2}{2F_e} = \frac{1}{2}T_i n_i \int d^3\boldsymbol{r}\ \tau\varphi'^2. \tag{A.62}$$

The ion contribution to W requires some work. Using (A.1), viz., $\delta f_i = h - \varphi F_i$, we obtain

$$\int d^3\boldsymbol{r} \int d^3\boldsymbol{v}\ \frac{T_i\delta f_i^2}{2F_i} = \int d^3\boldsymbol{r} \int d^3\boldsymbol{v}\ \frac{1}{2}T_i\left(\frac{\langle h^2\rangle_r}{F_i} + \varphi^2 F_i - 2\varphi\langle h\rangle_r\right). \tag{A.63}$$

Using the quasineutrality condition (A.16) and the gyroaverage expansion (A.6), we find

$$\int d^3\boldsymbol{r} \int d^3\boldsymbol{v}\ \frac{T_i\langle h^2\rangle_r}{2F_i} = \frac{1}{2}T_i n_i \int d^3\boldsymbol{r}\ \left[\frac{3}{2}\frac{\delta T}{T_i} + \varphi^2 + 2\tau\varphi'^2 - \frac{1}{2}\varphi\rho_i^2\nabla_\perp^2\left(\varphi + \frac{\delta T}{T_i}\right)\right], \tag{A.64}$$

$$\int d^3\boldsymbol{r} \int d^3\boldsymbol{v}\ \varphi\langle h\rangle_r = \varphi^2 + \tau\varphi'^2. \tag{A.65}$$

Finally, substituting these expressions into (A.63) and combining with (A.62), we get

$$W = \frac{1}{2}T_i n_i \int d^3\boldsymbol{r}\ \left[\frac{3}{2}\left(\frac{\delta T}{T_i}\right)^2 + \tau\varphi'^2 - \frac{1}{2}\varphi\rho_i^2\nabla_\perp^2\left(\varphi + \frac{\delta T}{T_i}\right)\right]. \tag{A.66}$$

To lowest order in $k_\perp^2\rho_i^2$, the free energy is, therefore,

$$W = \frac{3}{2}T_i n_i \int d^3\boldsymbol{r}\ \frac{1}{2}\left(\frac{\delta T}{T_i}\right)^2, \tag{A.67}$$

as promised in §2.2. Using (2.22) and (2.23), it is straightforward to show that, up to multiplicative constants related to the normalisations (2.17), (A.67) satisfies (2.28).

The 2D ($k_\parallel = 0$) GK equation for species s has an additional conserved quantity (Schekochihin *et al.*, 2009), given by

$$I_s = \frac{T_s}{2F_s}\int d^3\boldsymbol{R}\ \langle \delta f_s\rangle_R^2 = \frac{T_s}{2F_s}\int d^3\boldsymbol{R}\ \left(h_s - \frac{Z_s e}{T_s}\langle\phi\rangle_R F_s\right)^2. \tag{A.68}$$

In the 2D model (2.22) and (2.23), only the ions are assumed to have 2D dynamics — indeed, the modified adiabatic electron response (2.4) arises as a consequence precisely of the fast parallel streaming of the electrons. Thus, for the ions,

$$\begin{aligned} I \equiv I_i &= \frac{T_i}{2F_i}\int d^3\boldsymbol{R}\ (h - \langle\varphi\rangle_R F_i)^2 \\ &= \frac{1}{2}T_i\int d^3\boldsymbol{R}\ \left\{\left(\frac{\delta T}{T_i}\right)^2\left(\frac{v^2}{v_{ti}^2} - \frac{3}{2}\right)^2 F_i + \frac{\delta T}{T_i}\left(\frac{v^2}{v_{ti}^2} - \frac{3}{2}\right)h^{(1)}\right. \\ &\left.\quad + 2\frac{\delta T}{T_i}\left(\frac{v^2}{v_{ti}^2} - \frac{3}{2}\right)\left[\tau\varphi' - \frac{1}{4}\rho_i^2\nabla_\perp^2\left(\varphi + \frac{\delta T}{T_i}\right) - \frac{1}{4}\frac{v_\perp^2}{v_{ti}^2}\rho_i^2\nabla_\perp^2\varphi\right]F_i\right\}. \end{aligned} \tag{A.69}$$

Note that I is a function of velocity $\boldsymbol{v}$. In order to eliminate the unknown $h^{(1)}$, we can integrate I:

$$\int d^3\boldsymbol{v}\ I = \frac{1}{2}T_i n_i\int d^3\boldsymbol{r}\ \left[\frac{3}{2}\left(\frac{\delta T}{T_i}\right)^2 - \frac{1}{2}\frac{\delta T}{T_i}\rho_i^2\nabla_\perp^2\varphi\right]. \tag{A.70}$$

Subtracting this from the free energy (A.66), we obtain

$$W - \int d^3\boldsymbol{v}\ I = \frac{1}{2}T_i n_i\int d^3\boldsymbol{r}\ \left(\tau\varphi'^2 - \frac{1}{2}\varphi\rho_i^2\nabla_\perp^2\varphi\right) = \frac{1}{2}T_i n_i\tau\int d^3\boldsymbol{r}\ \left[\varphi'^2 + \rho_s^2(\boldsymbol{\nabla}_\perp\varphi)^2\right], \tag{A.71}$$

which is the conserved quantity in (2.26). This can be viewed as a version of the electrostatic GK invariant (Schekochihin *et al.*, 2009)

$$Y \equiv W - \sum_s \int d^3\boldsymbol{v}\ I_s, \tag{A.72}$$

but without the electron contribution I_e, which is not conserved because the electrons do not obey $k_\parallel = 0$.

In order to obtain the third conserved quantity, we go back to (A.69) and consider

$$\int d^3\boldsymbol{v}\ \left(\frac{v^2}{v_{ti}^2} - \frac{3}{2}\right)^{-1} I = T_i n_i\tau\int d^3\boldsymbol{r}\ \left(\varphi'\frac{\delta T}{T_i} - \frac{1}{2}\frac{\delta T}{T_i}\rho_s^2\nabla_\perp^2\frac{\delta T}{T_i} - \frac{\delta T}{T_i}\rho_s^2\nabla_\perp^2\varphi\right). \tag{A.73}$$

Adding this to (A.71), we obtain

$$W - \int d^3\boldsymbol{v}\, \left[1 - \left(\frac{v^2}{v_{ti}^2} - \frac{3}{2}\right)^{-1}\right] I = T_i n_i \tau \int d^3\boldsymbol{r}\, \left[\frac{1}{2}\varphi'^2 + \frac{\delta T}{T_i}\varphi' + \frac{1}{2}\rho_s^2 \left(\boldsymbol{\nabla}_\perp \varphi + \boldsymbol{\nabla}_\perp \frac{\delta T}{T_i}\right)^2\right]. \tag{A.74}$$

This is the quantity that satisfies (2.27).

Appendix B

Tertiary instability

As mentioned in §3.2.2, the turbulence in the convection zones is driven by an instability that is localised to the ZF extrema. This localisation is due to the strong zonal shear in the shear zones that suppresses nonzonal perturbations in those regions. To study the stability of a zonally dominated state, we consider infinitesimal ITG perturbations over a background of strong ZF and zonal temperature:

$$\begin{aligned} &\varphi = \overline{\varphi} + \varphi',\ \varphi' \ll \overline{\varphi}, \\ &T = \overline{T} + T',\ T' \ll \overline{T}, \end{aligned} \tag{B.1}$$

and linearise (2.22) and (2.23) to obtain evolution equations for φ' and T'. We refer to the ITG modes governed by these linearised equations as "tertiary modes", and to their linear instability as the "tertiary instability" (in truth, this is just the primary ITG instability but for an equilibrium state modified by the zonal fields).

Using the decomposition (B.1) and dropping the nonzonal-nonzonal interaction terms in (2.22) and (2.23), we obtain the linearised tertiary-mode equations:

$$\begin{aligned} (\partial_t + \underbrace{\partial_x\overline{\varphi}\partial_y}_{①})\big(1-\nabla_\perp^2\big)\varphi' - (1\underbrace{-\partial_x^3\overline{\varphi}}_{②})\partial_y(\varphi'+T') + (\kappa_T \underbrace{-\partial_x\overline{T}}_{③})\partial_y\nabla_\perp^2\varphi' \\ + \underbrace{\big(\partial_x^2\overline{\varphi}\big)\partial_x\partial_y T' - \big(\partial_x^2\overline{T}\big)\partial_x\partial_y\varphi'}_{④} = -\chi\nabla^4(a\varphi' - bT'), \end{aligned} \tag{B.2}$$

$$(\partial_t + \underbrace{\partial_x\overline{\varphi}\partial_y}_{⑤})T' + (\kappa_T \underbrace{-\partial_x\overline{T}}_{⑥})\partial_y\varphi' = \chi\nabla^2 T'. \tag{B.3}$$

Let us examine (B.2) and (B.3) to gain some insight into the way in which zonal fields might affect the ITG instability. Terms "1" and "5" represent the advection of density and temperature perturbations by the ZF. In the locations of nonzero ZF shear ($\partial_x^2\overline{\varphi} \neq 0$), the zonal advection is responsible for shearing the turbulent eddies and thus suppressing turbulence. We do indeed find that the growing tertiary modes are localised where the zonal shear vanishes, $\partial_x^2\overline{\varphi} = 0$. Terms "3" and "6" reflect the modification of the background temperature gradient by the zonal temperature gradient $\partial_x\overline{T}$. Their presence suggests that a possible mechanism for ITG saturation (or mitigation) is to excite a zonal temperature gradient that cancels the background gradient, thus effectively eliminating the turbulent drive (see appendix B.4). Term "2" shows that the derivative of the zonal shear, $\partial_x^3\overline{\varphi}$, modifies the effective background magnetic field gradient. Terms "4" do not have an obvious simple interpretation that we know of.

Even though we can always embark on a 4-mode calculation similar to the one done in §3.1, it is, in fact, not useful in this case. In contrast to the linear regime, where there is always a well-defined fastest-growing primary mode, neither a monochromatic zonal profile nor a monochromatic tertiary mode are ever observed in simulations. Instead, we find ITG modes that are localised around the points of vanishing zonal shear.

B.1 ITG tertiary instability

Let us derive a heuristic approximation for the growth rate of the ITG instability localised around a local extremum of the ZF and of the zonal temperature gradient, i.e., around x_0 such that $\partial_x^2\overline{\varphi}(x_0) = \partial_x^2\overline{T}(x_0) = 0$. Note that terms "4" in (B.2) then vanish at x_0. Assuming that the tertiary modes are sufficiently localised, we can take $\partial_x^3\overline{\varphi}$ and $\partial_x\overline{T}$ to be constant (Rogers & Dorland, 2005). We can then repeat the linear calculation of §2.3, but now including the zonally modified gradients of the equilibrium magnetic field and background temperature, and obtain a dispersion relation for the tertiary modes. For simplicity, let us consider the collisionless ($\chi = 0$) modes. The resulting growth rate for a mode with poloidal wavenumber k_y is

$$\gamma_3 = \frac{k_y\sqrt{4(1+k_\perp^2)(1-\partial_x^3\overline{\varphi})(\kappa_T - \partial_x\overline{T}) - \left[1 - \partial_x^3\overline{\varphi} + (\kappa_T - \partial_x\overline{T})k_\perp^2\right]^2}}{2(1+k_\perp^2)}, \tag{B.4}$$

which is a modified version of (2.38).

Under the same assumptions that we used before to obtain the simplified collisionless ITG growth rate (2.39), viz., $\kappa_T - \partial_x \overline{T} \gg 1$ and $k \ll (\kappa_T - \partial_x \overline{T})^{-1/4}$, we find that the tertiary growth rate is

$$\gamma_3 \approx k_y \sqrt{(1 - \partial_x^3 \overline{\varphi})(\kappa_T - \partial_x \overline{T})}. \tag{B.5}$$

If there are no background gradients, then $\gamma_3 \approx k_y \sqrt{\partial_x^3 \overline{\varphi}(x_0) \partial_x \overline{T}(x_0)}$, as found by Rogers *et al.* (2000). To work out how wide this mode can be, note that we have ignored the effects of the nonzero zonal shear away from the point x_0. Taylor-expanding terms "1" and "5" in (B.2) and (B.3) around x_0, we have

$$ik_y \partial_x \overline{\varphi}(x) = ik_y \partial_x \overline{\varphi}(x_0) + \frac{1}{2} ik_y \partial_x^3 \overline{\varphi}(x_0)(x - x_0)^2 + \mathcal{O}\big((x - x_0)^3\big), \tag{B.6}$$

where the first term represents a Doppler shift in the frequency and the second captures the effect of the nonzero zonal shear. Dropping the latter is, therefore, valid only in an interval around x_0 such that

$$\begin{aligned} &\gamma_3 \gg \frac{1}{2} k_y \partial_x^3 \overline{\varphi}(x_0)(x - x_0)^2 \\ \implies &(x - x_0)^2 \ll \Delta^2 \equiv \frac{2\gamma_3}{k_y \partial_x^3 \overline{\varphi}(x_0)} \approx 2\sqrt{\frac{[1 - \partial_x^3 \overline{\varphi}(x_0)]\big[\kappa_T - \partial_x \overline{T}(x_0)\big]}{\big[\partial_x^3 \overline{\varphi}(x_0)\big]^2}}, \end{aligned} \tag{B.7}$$

where we used (B.5) for the final approximation. Rogers *et al.* (2000) and Rogers & Dorland (2005) found that the scale Δ is a good approximation for the radial width of the ITG tertiary mode in the case of no equilibrium gradients.

We find that the standard simplified picture of the ITG tertiary instability outlined above does not quite describe the observed tertiary modes. Namely, (i) we find a strong temperature-gradient-driven instability at ZF minima where $\partial_x^3 \overline{\varphi} > 1$ and (B.4) predicts no instability; and (ii) the instability at the ZF maxima is significantly (an order-of-magnitude) slower than predicted by (B.4). This is detailed in Appendix B.3.

B.2 KH tertiary instability

Recall that, as we showed in §2.1.5, (2.22) reduces to the Hasegawa-Mima equation in the case of $\kappa_T = 0$. This reduction is naturally achieved in the vicinity of a point

x_0 where $\partial_x^2\overline{\varphi}(x_0) = \partial_x^2\overline{T}(x_0) = 0$ and $\partial_x\overline{T}(x_0) = \kappa_T$. Then the temperature equation (B.3) decouples from (B.2) due to the cancellation (or "flattening") of the equilibrium temperature gradient by the zonal temperature. In that case, $T' = 0$ is a solution to (B.3) and (B.2) reduces to

$$(\partial_t + \overline{u}_y\partial_y)\big(1 - \nabla_\perp^2\big)\varphi' - (1 - \partial_x^2\overline{u}_y)\partial_y\varphi' = -a\chi\nabla^4\varphi', \tag{B.8}$$

which is the linearised Hasegawa-Mima equation for tertiary modes. This equation has a Kelvin-Helmholtz-like (KH) tertiary instability (Kim & Diamond, 2002; Numata *et al.*, 2007; St-Onge, 2017; Zhu *et al.*, 2018*a*). This KH instability is also localised around the radial locations of zero zonal shear and has a threshold roughly given by the necessary (but, in general, not sufficient) condition that $\partial_x^2\overline{u}_y - 1$ must change sign in the region of instability, known as the Rayleigh-Kuo criterion (Kuo, 1949; Zhu *et al.*, 2018*a*).

B.3 Tertiary instabilities of the zonal staircase

We now turn to a numerical investigation of the tertiary instability in the low-collisionality regime. We will consider the parameters and staircase profile of SimH ($\kappa_T = 0.36$, $\chi = 0.1$). The ITG turbulence trapped in the convection zones at the ZF extrema can differ substantially between ZF maxima and minima. For example, it is evident from figures 3.6 and 3.7 that the ZF minima harbour turbulence with a larger poloidal wavenumber compared to the turbulence at the ZF maxima.

Let us describe this difference quantitatively for SimL. Consider the poloidal Fourier transforms of our fields at a fixed radial location:

$$\varphi(x,y) = \sum_{k_y}\hat{\varphi}(x,k_y)e^{ik_y y}, \qquad T(x,y) = \sum_{k_y}\hat{T}(x,k_y)e^{ik_y y}. \tag{B.9}$$

Figure B.1 shows that the perturbations located at the ZF minima have a k_y spectrum peaked at a finite wavenumber, whereas the perturbations around the ZF maxima saturate at the largest available poloidal scale in the box.

An asymmetry between the ZF extrema was noticed already by McMillan *et al.* (2011), who, in their GK simulations, found that the ZF minima are less effective than the ZF maxima at stabilising turbulence. This difference has also been extensively studied using the Hasegawa-Mima equation (Zhu *et al.*, 2018*b*). We find that these differences are a consequence of the mechanisms that drive the tertiary modes. By inspecting equations (B.2) and (B.3), one can instantly identify the culprit of the asymmetry: the sign of $\partial_x^3\overline{\varphi}$ is different in the maxima (negative) and minima (positive) of the ZF. According to our analysis of the tertiary instabilities in Appendices B.1 and B.2, this has two distinct effects on the tertiary instability at the ZF extrema: one is related to the modification of the magnetic drift term and the second one to the existence of the KH tertiary at the ZF minima.

In order to investigate the tertiary instabilities that operate in the saturated state that we observe, we solve (2.22) and (2.23) numerically with an imposed static ZF profile extracted from the quasi-static zonal staircase. We do not impose a static zonal temperature profile because the observed flattening of the equilibrium temperature gradient suggests that the zonal temperature is an agent of saturation, rather than an instability mechanism.

B.3.1 Tertiary instability at ZF minima

Figure B.2 shows the behaviour of the tertiary instability at the ZF maxima and minima. Let us first discuss the ZF minima. There we observe a very fast ITG-like initial instability that features both φ' and T' growing at the rate $\gamma_3^{\text{fast}} \approx 0.1$. This is of the same order as the largest linear ITG growth rate $\gamma_{\max} \approx 0.2$ for the parameters of this simulation. The strong radial transport caused by the ITG eddies drives a zonal temperature perturbation that opposes the equilibrium temperature gradient, thus flattening it. This causes a "quasilinear" saturation of the ITG instability around $t = 50$ when the zonal temperature gradient reaches the level of the equilibrium gradient and the instability is largely quenched (see also Appendix B.4). Afterwards, φ' continues to grow exponentially, albeit at a much lower growth rate of $\gamma_3^{\text{quenched}} \approx 0.003$. The nonzonal temperature T' does not grow during this quenched-growth phase and remains about an order of magnitude lower than the electrostatic potential. As we discussed in Appendix B.2, the flattening of the temperature gradient ($\kappa_T \approx \partial_x\overline{T}$) and the comparatively low levels of temperature perturbations ($T' \ll \varphi'$) imply that φ'

satisfies the Hasegawa-Mima tertiary equation (B.8). We find that the Rayleigh-Kuo criterion $\partial_x^3\overline{\varphi} \gtrsim 1$ is satisfied at the ZF minima and so there is an unstable KH mode. Its growth rate ($\gamma_3^{\mathrm{HM}} \approx 0.008$) and poloidal spectrum are found to be similar to those of the quenched tertiary following the temperature-profile flattening (see figure B.1). We have found that if we do not evolve the zonal temperature and, thus, do not allow a flattening of the temperature gradient, the initial fast growth is never quenched.

In the vicinity of the ZF minima, we find $\partial_x^3\overline{\varphi} \gtrsim 1$ and (B.4) predicts no ITG mode there. In fact, we do find an ITG instability (the fast initial one in figure B.2) peaked exactly at x_0 where $\partial_x^2\overline{\varphi}(x_0) = 0$, even though $\partial_x^3\overline{\varphi}(x_0) > 1$. We have verified that this is an ITG instability and not a KH one by setting $\kappa_T = 0$ or by freezing the nonzonal temperature perturbations and solving the corresponding Hasegawa-Mima equation numerically. Either of these reduces the tertiary growth rate by an order of magnitude, down to the level of the KH instability, indicating that the fast mode is indeed an ITG instability. Even artificially scaling the ZFs in order to get $\partial_x^3\overline{\varphi} > 2$ in the ZF minima results in a fast ITG instability there. This tertiary mode is not the one described in Appendix B.1 and familiar from the work of Rogers *et al.* (2000). A further investigation of it is left for future work. Its effect in the quasi-stationary zonal-staircase-dominated state is to push the system in the vicinity of the ZF minima towards a state in which the zonal temperature perturbations cancel the equilibrium temperature gradient (see Appendix B.4).

B.3.2 Tertiary instability at ZF maxima

In contrast to the ZF minima, the ZF maxima are hosts to a wimpy ITG instability with a growth rate an order of magnitude smaller than that near the ZF minima: $\gamma_3^{\mathrm{wimpy}} = 0.006$ (see figure B.2, left panel). Setting $\kappa_T = 0$ eliminates this instability, confirming that it is, indeed, an ITG instability. The ZF satisfies $\partial_x^3\overline{\varphi} < 0$ at the ZF maxima, so there is no unstable KH mode there; this is confirmed numerically. The poloidal wavelength of the turbulence at the ZF maxima is significantly longer than that at the ZF minima. In fact, in SimL and SimH, it is determined by the poloidal box size L_y. By increasing L_y, we are able to obtain a saturated state with a well-defined wavelength at the ZF maxima that is smaller than the poloidal box size: see figure B.3.

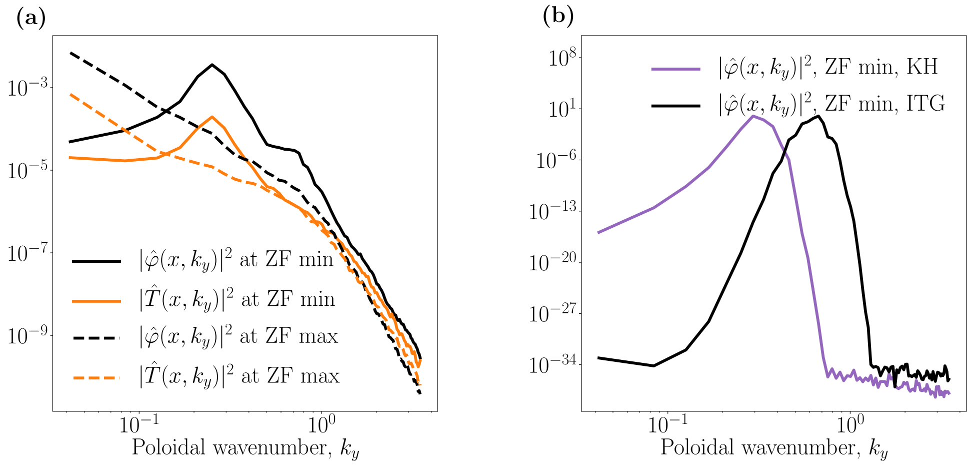


Figure B.1: **(a)** Time-averaged poloidal spectra $\langle|\hat{\varphi}(x,k_y)|^2\rangle$ (black) and $\langle|\hat{T}(x,k_y)|^2\rangle$ (orange), as defined by (B.9), at a fixed radial location x for a ZF minimum (solid) and a ZF maximum (dashed). The time average is performed over the entire time period shown in figure 3.9. A peak in the spectrum is observed around $k_y \approx 0.25$ for the turbulence at the ZF minima. This agrees with the wavelength evident in figure 3.7. In contrast, the ITG turbulence at the ZF maxima saturates at the largest available poloidal length in the box. **(b)** Poloidal spectra $\langle|\hat{\varphi}(x,k_y)|^2\rangle$ of the unstable modes at a ZF minimum for the KH instability that develops in the Hasegawa-Mima equation (purple) and the fast ITG instability without zonal temperature perturbations (black). The spectra have been normalised to the maximal value of 1. We see that both modes are approximately monochromatic, peaked around $k_y \approx 0.29$ for the KH mode and $k_y \approx 0.66$ for the ITG mode. The spectrum of saturated tubulence at the ZF minima peaks around $k_y \approx 0.25$, see (a). This supports the case for the poloidal wavenumber at the ZF minima to be determined by the KH instability there.

The suppression of the ITG instability at the ZF maxima might be due to the localisation width Δ, naïvely given by (B.7). Using the observed values of $\partial_x^3\overline{\varphi}$ and $\partial_x\overline{T}$ in (B.7), we find $\Delta \approx 1.8$ for the fastest mode of (B.4). This suggests that these ITG modes should have radial wavenumbers of the order of $k_x \sim 2\pi/\Delta \approx 3.5$, which are deep in the stable region for the parameters considered (see figure 2.2b). We leave the detailed investigation of these modes for future work. Their effect in the quasi-stationary zonal-staircase-dominated state is to seed turbulence in the shear zones through the emission of travelling structures (ferdinons) when the zonal shear in the shear zones has been sufficiently weakened by viscosity (see §3.2.1.2 and §3.2.3).

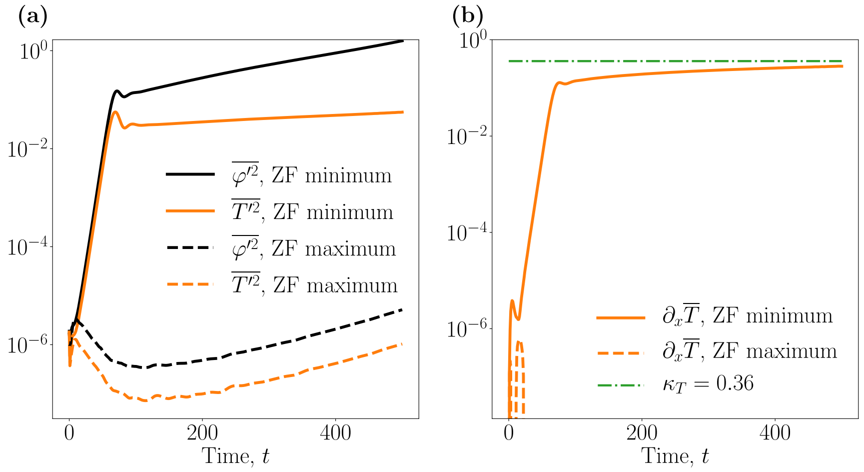


Figure B.2: Time evolution of the perturbations in the ZF extrema of the zonal staircase of SimH. **(a)** $\overline{\varphi'^2}$ (solid black) and $\overline{T'^2}$ (solid orange) at the ZF minimum at $x \approx 24.5$. Dashed lines show the same quantities at the ZF maximum at $x \approx 47$. See figure 3.6a for the ZF profile. **(b)** Zonal temperature gradient at the ZF minimum (the ZF maximum has a very low and negative zonal temperature gradient, not shown). After the initial fast instability, the zonal temperature gradient settles at $\partial_x\overline{T} \approx \kappa_T = 0.36$.

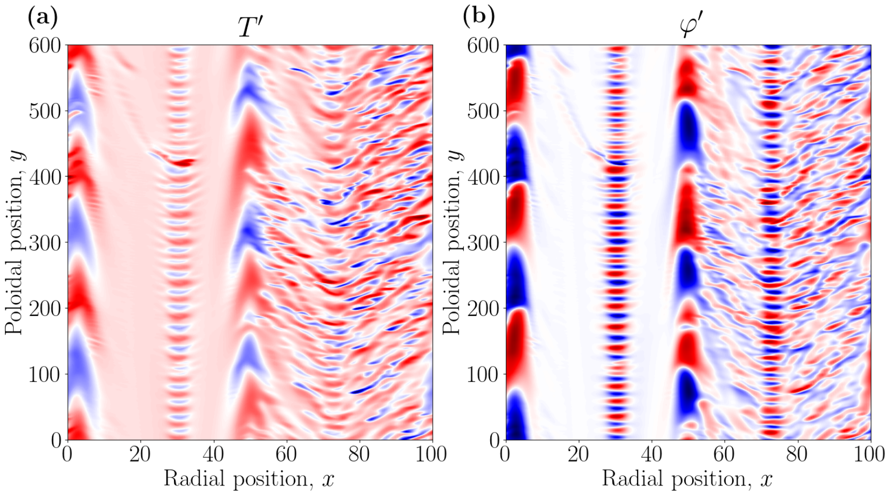


Figure B.3: Snapshots of T and φ perturbations for SimH parameters, viz., $\kappa_T = 0.36$, $\chi = 0.1$, but for $L_y = 600$. Evident is a discrepancy in the poloidal wavenumbers of the nonzonal perturbations at the ZF maxima ($x \approx 5, 50$) and ZF minima ($x \approx 30, 75$). The dominant wavenumber at the ZF minima is $k_y \approx 0.25$, just as in the shorter box (see figure B.1), and $k_y \approx 0.03$ at the ZF maxima. The poloidal wavelength at the ZF maxima is shorter than the poloidal size of the box.

B.4 Zonal-temperature saturation and equilibrium-gradient flattening

Is it inevitable that the temperature profile is flattened in the regions of vigorous instability? Intuitively, one might expect this to be a consequence of the high level of radial transport by the nonzonal (φ') eddies. In that case, the flattening ought to depend on some measure of the strength of the radial $\boldsymbol{E} \times \boldsymbol{B}$velocity. Let us translate this intuition into equations.

The zonal part of (2.23) is

$$\partial_t \overline{T} = \partial_x \overline{(T\partial_y \varphi + \chi \partial_x T)} = -\partial_x (Q_t + Q_d), \tag{B.10}$$

where $Q_t = -\overline{T\partial_y \varphi}$ and $Q_d = -\chi \partial_x \overline{T}$ are the local turbulent and diffusive radial heat fluxes, respectively. This is a conservation equation for the zonal temperature perturbations. The fact that the zonal temperature stays constant $\partial_t \overline{T} = 0$ during the quiescent periods implies a balance of the local radial heat fluxes:

$$\partial_x (Q_t + Q_d) = 0 \implies Q_t = -Q_d + Q \implies \partial_x \overline{T} = \frac{Q_t - Q}{\chi}, \tag{B.11}$$

where the integration constant Q is the total box-averaged radial turbulent heat flux defined by (2.29). Indeed, this balance is observed in the quiescent zonal staircase (see figure B.4).

Consider the steady state ($\partial_t = 0$) solution of (B.3). Since we are interested in the behaviour of the temperature perturbations in the ZF extrema, where the zonal shear vanishes, we drop the effect of zonal shear (formally speaking, we transform to a frame moving with the local ZF velocity). We find

$$0 = \partial_t T' = \chi \nabla^2 T' - (\kappa_T - \partial_x \overline{T}) \partial_y \varphi'. \tag{B.12}$$

Substituting (B.11) into (B.12) gives us

$$(\chi \kappa_T - Q_t + Q) \partial_y \varphi' = \chi^2 \nabla^2 T'. \tag{B.13}$$

Multiplying both sides of (B.13) by the radial $\boldsymbol{E} \times \boldsymbol{B}$velocity $u_x = -\partial_y \varphi'$ and performing a poloidal average yields

$$(\chi \kappa_T - Q_t + Q) \overline{u_x^2} = \chi^2 \overline{\left(\frac{\nabla^2 T'}{T'} \right) T' \partial_y \varphi'} \approx \chi^2 Q_t k_\perp^2, \tag{B.14}$$

where we have approximated the thermal diffusion as that of a locally monochromatic temperature perturbation T' with a wavenumber $k_\perp$. Then

$$Q_t = \frac{Q + \chi\kappa_T}{1 + \chi^2 k_\perp^2/\overline{u_x^2}}. \tag{B.15}$$

Substituting (B.15) into the rightmost expression in (B.11) yields the following prediction for the zonal temperature gradient

$$\partial_x \overline{T} = \frac{\kappa_T - Q\chi k_\perp^2/\overline{u_x^2}}{1 + \chi^2 k_\perp^2/\overline{u_x^2}}. \tag{B.16}$$

Therefore, for $Q > 0$, steady-state saturation implies $\partial_x \overline{T} < \kappa_T$. In the case of strong radial transport, $\overline{u_x^2} \gg \chi^2 k_\perp^2$, we obtain

$$Q_t \approx Q + \chi\kappa_T, \quad \partial_x \overline{T} \approx \kappa_T. \tag{B.17}$$

Thus, perturbations with a strong radial $\boldsymbol{E}\times\boldsymbol{B}$velocity flatten out the temperature gradient by generating a strong zonal temperature perturbation, as intuitively expected. Indeed, (B.15) agrees with the numerical data for the heat flux at the ZF minima (see figure B.5), where we find the fast ITG tertiary instability that saturated via flattening of the equilibrium temperature gradient, as discussed in Appendix B.3.1.

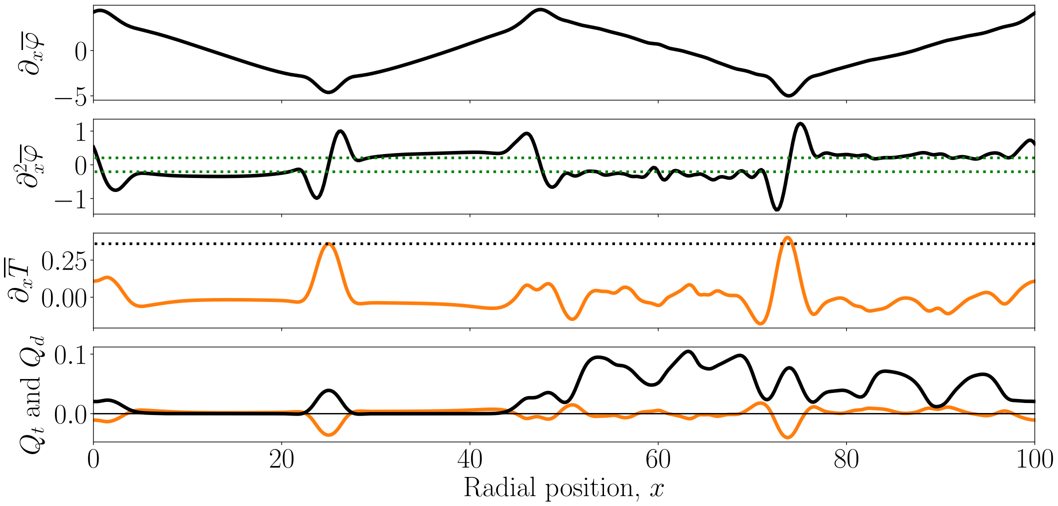


Figure B.4: Same as figure 3.6c, but with plots of the turbulent $Q_t = -\overline{T\partial_y\varphi}$ (black) and diffusive $Q_d = -\chi\partial_x\overline{T}$ (orange) heat fluxes as well. We can see that the balance (B.11) between Q_t and Q_d holds in the saturated staircase (left half of the domain), but does not in the turbulent regions (right half of the domain), where the turbulence generates a significant turbulent heat flux.

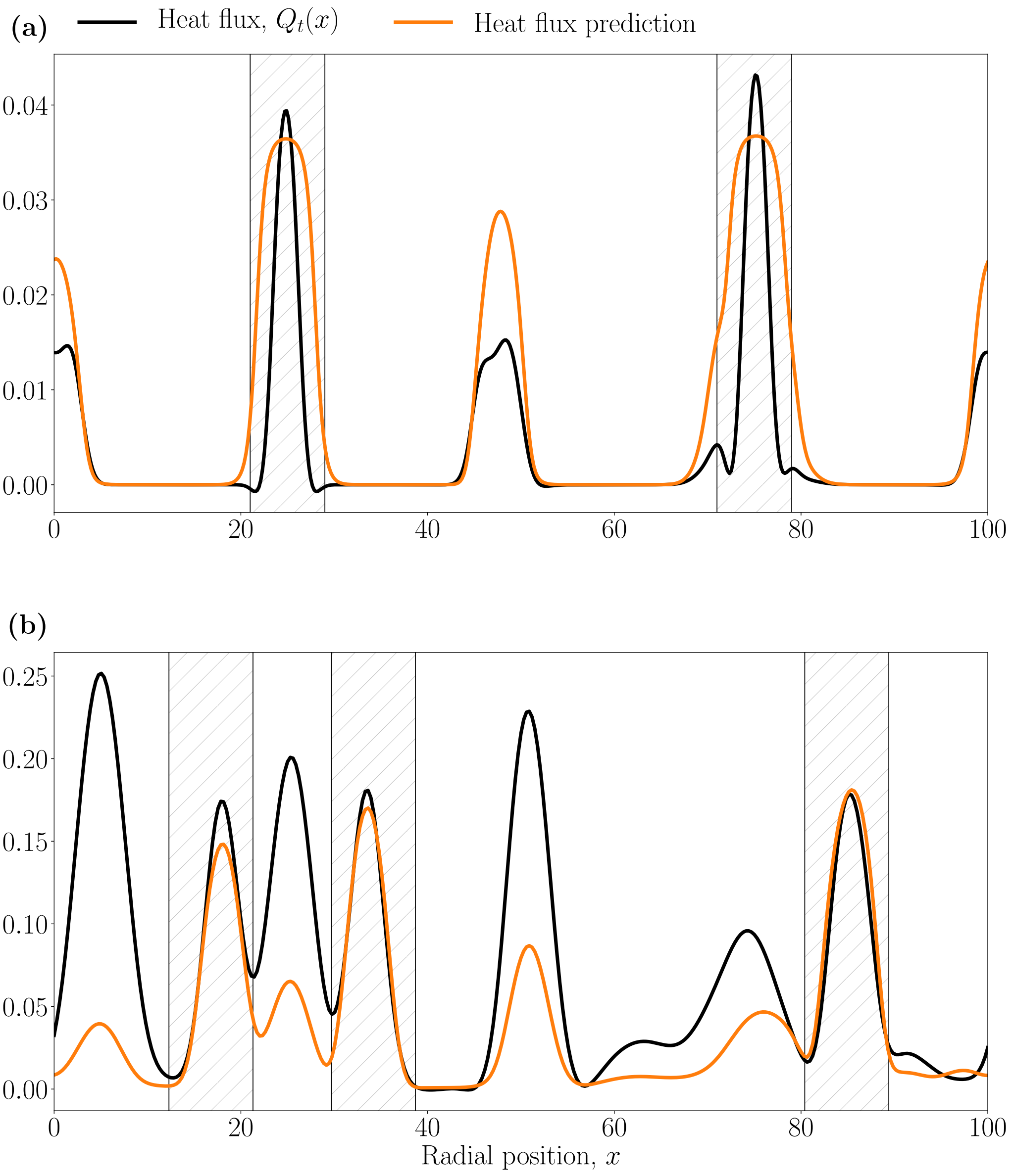


Figure B.5: Comparison of the estimate (B.15) (orange) and observed turbulent heat flux (black) of the saturated zonal staircase. To choose $k_\perp$, we estimated k_x and k_y using the observed radial width δ of the convection zones and the poloidal spectrum of the turbulence there. The shaded areas highlight the ZF minima and have a width δ. **(a)** Data from SimH with $\delta = 8$ (so $k_x = 2\pi/8 \approx 0.8$) and $k_y = 0.25$, corresponding to the spectral peak in figure B.1. This gives $k_\perp \approx 0.8$. **(b)** Same as SimH, but with increased collisionality $\chi = 1$. We estimated $\delta = 9$ (so $k_x \approx 0.7$) and $k_y = 0.25$, corresponding to the peak of the poloidal spectrum at the ZF minima in that simulation. This gives $k_\perp \approx 0.7$. The agreement is better for the higher value of collisionality.

Appendix C

Slab-ITG

C.1 Slab-ITG instability condition

Here we derive the instability boundaries (2.49) for the dispersion relation (2.48). Note that the left-hand side of (2.48) is a cubic polynomial in $\hat{\omega}_{\boldsymbol{k}}$ with one positive and two negative roots, while the right-hand side is a simple quadratic propotional to $\hat{\omega}_{\boldsymbol{k}}^2$ (see figure C.1). First, if $\hat{\gamma}_{\boldsymbol{k}}^2 < 0$, then the right-hand side is a concave parabola and it is geometrically evident that there will always be three intersections of the parabola and the cubic, and so there are no unstable solutions. On the other hand, if $\hat{\gamma}_{\boldsymbol{k}}^2 > 0$, then these two curves cross three times if and only if the cubic left-hand side is larger than the quadratic right-hand side at $\hat{\omega}_{\boldsymbol{k}} = \hat{\omega}_{\boldsymbol{k}}^{(0)} < 0$ where the two curves share an identical slope. We differentiate (2.48) to find that $\hat{\omega}_{\boldsymbol{k}}^{(0)}$ is the negative solution to

$$\left(\hat{\omega}_{\boldsymbol{k}}^{(0)}\right)^2 + \left[\frac{2}{3} - \frac{4k_\perp^2\hat{\gamma}_{\boldsymbol{k}}^2}{3(1+k_\perp^2)}\right]\hat{\omega}_{\boldsymbol{k}}^{(0)} - \frac{1}{3}\frac{k_\parallel^2}{1+k_\perp^2} = 0. \tag{C.1}$$

Using (C.1) to substitute for $\hat{\omega}_{\boldsymbol{k}}^{(0)}$, the instability condition that the left-hand side of (2.48) be larger than its right-hand side is then found to be equivalent to

$$\hat{\omega}_{\boldsymbol{k}}^{(0)} < \hat{\omega}_{\boldsymbol{k}}^{\max} \equiv -\hat{k}_\parallel^2 \frac{4(1+k_\perp^2) + k_\perp^2\hat{\gamma}_{\boldsymbol{k}}^2}{3\hat{k}_\parallel^2(1+k_\perp^2) + (1+k_\perp^2 - 2k_\perp^2\hat{\gamma}_{\boldsymbol{k}}^2)^2}. \tag{C.2}$$

Since $\hat{\omega}_{\boldsymbol{k}}^{(0)}$ is the negative solution to the quadratic (C.1) and $\hat{\omega}_{\boldsymbol{k}}^{\max} < 0$, (C.2) can be true if and only if the quadratic (C.1) is negative when we substitute $\hat{\omega}_{\boldsymbol{k}}^{\max}$ for $\hat{\omega}_{\boldsymbol{k}}^{(0)}$. Performing that substitution and simplifying the resulting expression yields an involved quadratic inequality for $\hat{k}_\parallel^2$ whose solution is the interval $\hat{k}_\parallel^2 \in (\hat{k}_{\parallel,-}^2, \hat{k}_{\parallel,+}^2)$, where $\hat{k}_{\parallel,\pm}^2$ are given by (2.49).

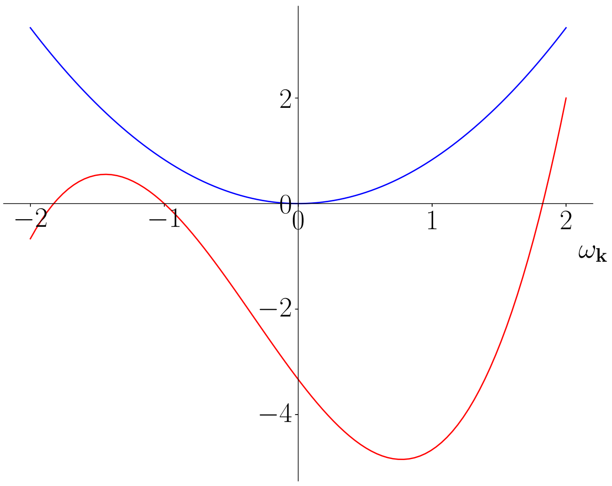


Figure C.1: The left-hand (red) and right-hand (blue) sides of the slab-ITG dispersion relation (2.48) for $\hat{k}_\parallel = 2$, $k_\perp^2 = 0.2$. There is only one real solution, so there exists a complex one with positive imaginary part. Thus, there are linearly unstable modes for $\hat{k}_\parallel = 2$, $k_\perp^2 = 0.2$.

C.2 Slab-ITG instability with general gradients and low collisionality

We aim to solve the dispersion relation of (4.12)–(4.14) in the $k_\parallel \sim k_\perp^2 \gg 1$ limit, neglecting the magnetic-drift contributions, and ordering collisionality as $\chi k_\perp^3 \sim \kappa_T$. Ignoring the magnetic-drift term $-\partial_y(\varphi + T)$ (as it is subdominant for the small-scale slab-ITG modes, see §2.3.2.2), we find the following dispersion relation

$$\omega_{\boldsymbol{k}}(1+k_\perp^2) - \frac{k_\parallel^2}{\omega_{\boldsymbol{k}} + id\chi k_\perp^2}\left(1 + \frac{\boldsymbol{\kappa}_T \cdot \boldsymbol{k}}{\omega_{\boldsymbol{k}} + i\chi k_\perp^2}\right) + k_\perp^2 \boldsymbol{\kappa}_T \cdot \boldsymbol{k} - \boldsymbol{\kappa}_n \cdot \boldsymbol{k} + ia\chi k_\perp^4 - \frac{ibk_\perp^4 \boldsymbol{\kappa}_T \cdot \boldsymbol{k}}{\omega_{\boldsymbol{k}} + i\chi k_\perp^2} = 0. \tag{C.3}$$

As already mentioned in §4.2.3, the dispersion relation must be invariant under $\boldsymbol{k} \mapsto -\boldsymbol{k}$, $\omega_{\boldsymbol{k}} \mapsto -\omega_{\boldsymbol{k}}^*$, so, without loss of generality, we assume that $\boldsymbol{\kappa}_T \cdot \boldsymbol{k} > 0$. We then write (C.3) as

$$\left[\hat{\omega}_{\boldsymbol{k}}^2(1 + k_\perp^2) - \hat{k}_\parallel^2\right](\hat{\omega}_{\boldsymbol{k}} + 1) = 2k_\perp^2 \hat{\gamma}_{\boldsymbol{k}}^2 \hat{\omega}_{\boldsymbol{k}}^2 - i\hat{\chi}_{\boldsymbol{k}} k_\perp^4 \left[\frac{d\hat{k}_\parallel^2}{k_\perp^2}\left(1 + \frac{1}{\hat{\omega}_{\boldsymbol{k}}}\right) + \frac{\hat{k}_\parallel^2}{k_\perp^2 \hat{\omega}_{\boldsymbol{k}}} + a\hat{\omega}_{\boldsymbol{k}}^2 - b\hat{\omega}_{\boldsymbol{k}}\right], \tag{C.4}$$

where, in addition to the definitions in §2.3.2, we have defined the following quantities:

$$\hat{\gamma}_{\boldsymbol{k}} \equiv \sqrt{\frac{(\boldsymbol{\kappa}_T + \boldsymbol{\kappa}_n) \cdot \boldsymbol{k}}{2k_\perp^2 \boldsymbol{\kappa}_T \cdot \boldsymbol{k}}}, \quad \hat{\chi}_{\boldsymbol{k}} \equiv \frac{\chi}{\boldsymbol{\kappa}_T \cdot \boldsymbol{k}}. \tag{C.5}$$

Here $\hat{\gamma}_{\boldsymbol{k}}$ is the largest collisionless growth rate.

As we discussed in §2.3.2.2, the linearly unstable solutions lie close to $\hat{k}_\parallel = k_\perp$ and are given by $\hat{\omega}_{\boldsymbol{k}} = -1 + \delta\hat{\omega}_{\boldsymbol{k}}$, where $\delta\hat{\omega}_{\boldsymbol{k}} \sim \mathcal{O}(1/k_\perp) \ll 1$. Substituting into (C.4) $\hat{\omega}_{\boldsymbol{k}} = -1 + \delta\hat{\omega}_{\boldsymbol{k}}$ and $\hat{k}_\parallel = k_\perp + \delta\hat{k}_\parallel$, where $\delta\hat{\omega}_{\boldsymbol{k}}/\hat{\omega}_{\boldsymbol{k}} \sim \delta\hat{k}_\parallel/\hat{k}_\parallel \sim \mathcal{O}(1/k_\perp) \ll 1$, we find

$$\delta\hat{\omega}_{\boldsymbol{k}} = -\frac{\delta\hat{k}_\parallel}{2k_\perp} \pm \sqrt{\frac{\delta\hat{k}_\parallel^2}{4k_\perp^2} - \hat{\gamma}_{\boldsymbol{k}}^2 + \frac{i(a+b-1)\hat{\chi}_{\boldsymbol{k}} k_\perp^2}{2}} + \mathcal{O}\big(k_\perp^{-2}\big). \tag{C.6}$$

We claim that $\delta\hat{\omega}_{\boldsymbol{k}}$ achieves its largest imaginary part when $\delta\hat{k}_\parallel = 0$. This is easiest to see if we consider the map $z \mapsto \sqrt{z + ic}$, where $z, c \in \mathbb{R}$. Let us define $x = \mathrm{Re}\big(\sqrt{z+ic}\big)$ and $y = \mathrm{Im}\big(\sqrt{z+ic}\big)$. We claim that y monotonically decreases with increasing z on the branch where $y > 0$. Indeed, we can square $x + iy = \sqrt{z+ic}$ to find that x and y satisfy $x^2 - y^2 = z$ and $2xy = c$, i.e., x and y lie on the intersection of two hyperbolae. One of these is fixed by c, and the other produces a family on non-intersecting hyperbolae as z varies. It is then geometrically evident that y decreases monotonically with z: see figure C.2. Therefore, (C.6) attains its largest imaginary part, i.e., the largest growth rate, when $\delta\hat{k}_\parallel = 0$. Moreover, as $x = c/2y$, the sign of the real part of $\sqrt{z+ic}$ for the branch with $y > 0$ is determined by the sign of c.

Then, using (2.19), we obtain

$$\frac{T_{\boldsymbol{k}}}{\varphi_{\boldsymbol{k}}} = \frac{\boldsymbol{\kappa}_T \cdot \boldsymbol{k}}{\omega_{\boldsymbol{k}} + i\chi k_\perp^2} = \frac{1}{\hat{\omega}_{\boldsymbol{k}}} + \mathcal{O}\big(k_\perp^{-2}\big) = \frac{1}{-1+\delta\hat{\omega}_{\boldsymbol{k}}} + \mathcal{O}\big(k_\perp^{-2}\big) = -1 - \delta\hat{\omega}_{\boldsymbol{k}} + \mathcal{O}\big(k_\perp^{-2}\big), \tag{C.7}$$

where we have dropped the $i\chi k_\perp^2$ term from the denominator because $\chi k_\perp^2 \sim |\boldsymbol{\kappa}_T| k_\perp^{-1} \sim \omega_{\boldsymbol{k}} k_\perp^{-2}$ is small. Substituting (C.6) into (C.7) then gives

$$\frac{T_{\boldsymbol{k}}}{\varphi_{\boldsymbol{k}}} = -1 - \sqrt{-\hat{\gamma}_{\boldsymbol{k}}^2 + \frac{i(a+b-1)\hat{\chi}_{\boldsymbol{k}} k_\perp^2}{2}} + \mathcal{O}\big(k_\perp^{-2}\big) \tag{C.8}$$

for the linearly unstable mode with the largest growth rate. The sign of the real part of $T_{\boldsymbol{k}}/\varphi_{\boldsymbol{k}}$ for this most unstable mode is then given by the sign of $a + b - 1$.

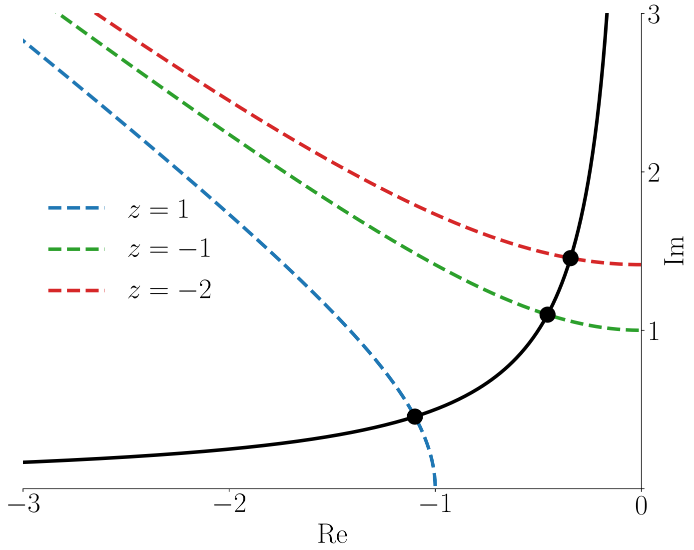


Figure C.2: A visualisation of the branch of $\sqrt{z + ic}$ with positive imaginary part for $c = -1$. The black curve is the hyperbola $2xy = c$ and the coloured curves show $x^2 - y^2 = z$ for three different values of z. The corresponding values of $\sqrt{z + ic}$ are given by the intersection points marked with black circles. We see that increasing z brings the intersection point "down" the black hyperbola and hence decreases the imaginary part of $\sqrt{z + ic}$. If c is positive, then the black curve sits in the first quadrant and $\sqrt{z + ic}$ has a positive real part.

References

Abel, IG & Cowley, SC 2013 Multiscale gyrokinetics for rotating tokamak plasmas: II. Reduced models for electron dynamics. *New J. Phys* **15**, 023041.

Abel, IG, Plunk, GG, Wang, E, Barnes, M, Cowley, SC, Dorland, W & Schekochihin, AA 2013 Multiscale gyrokinetics for rotating tokamak plasmas: fluctuations, transport and energy flows. *Rep. Prog. Phys* **76**, 116201.

Abiteboul, J 2012 Turbulent and neoclassical toroidal momentum transport in tokamak plasmas. Theses, Aix-Marseille Université.

Barnes, M, Parra, FI & Schekochihin, AA 2011 Critically balanced ion temperature gradient turbulence in fusion plasmas. *Phys. Rev. Lett.* **107**, 115003.

Beer, MA 1995 Gyrofluid Models of Turbulent Transport in Tokamaks. PhD thesis, Princeton University.

Burrell, KH 1997 Effects of ExB velocity shear and magnetic shear on turbulence and transport in magnetic confinement devices. *Phys. Plasmas* **4**, 1499.

Calvo, I & Parra, FI 2015 Radial transport of toroidal angular momentum in tokamaks. *Plasma Phys. Control. Fusion* **57**, 075006.

Catto, PJ 2019 Practical gyrokinetics. *J. Plasma Phys.* **85**, 925850301.

Catto, PJ & Simakov, AN 2004 A drift ordered short mean free path description for magnetized plasma allowing strong spatial anisotropy. *Phys. Plasmas* **11**, 90.

Catto, PJ & Simakov, AN 2005 A new, explicitly collisional contribution to the gyroviscosity and the radial electric field in a collisional tokamak. *Phys. Plasmas* **12**, 114503.

Colyer, GJ, Schekochihin, AA, Parra, FI, Roach, CM, Barnes, MA, Ghim, Y-c & Dorland, W 2017 Collisionality scaling of the electron heat flux in ETG turbulence. *Plasma Phys. Control. Fusion* **59**, 055002.

Coppi, B, Rosenbluth, MN & Sagdeev, RZ 1967 Instabilities due to temperature gradients in complex magnetic field configurations. *Phys. Fluids* **10**, 582.

Cowley, SC, Kulsrud, RM & Sudan, R 1991 Considerations of ion-temperature-gradient-driven turbulence. *Phys. Fluids B* **3**, 2767.

Diamond, PH, Itoh, S-I, Itoh, K & Hahm, TS 2005 Zonal flows in plasma—a review. *Plasma Phys. Control. Fusion* **47**, R35.

Dif-Pradalier, G, Diamond, PH, Grandgirard, V, Sarazin, Y, Abiteboul, J, Garbet, X, Ghendrih, Ph, Strugarek, A, Ku, S & Chang, CS 2010 On the validity of the local diffusive paradigm in turbulent plasma transport. *Phys. Rev. E* **82**, 025401.

Dif-Pradalier, G, Hornung, G, Garbet, X, Ghendrih, Ph, Grandgirard, V, Latu, G & Sarazin, Y 2017 The ExB staircase of magnetised plasmas. *Nuclear Fusion* **57**, 066026.

Dif-Pradalier, G, Hornung, G, Ghendrih, Ph, Sarazin, Y, Clairet, F, Vermare, L, Diamond, PH, Abiteboul, J, Cartier-Michaud, T, Ehrlacher, C, Estève, D, Garbet, X, Grandgirard, V, Gürcan, Ö D, Hennequin, P, Kosuga, Y, Latu, G, Maget, P, Morel, P, Norscini, C, Sabot, R & Storelli, A 2015 Finding the elusive ExB staircase in magnetized plasmas. *Phys. Rev. Lett.* **114**, 085004.

Dimits, AM, Bateman, G, Beer, MA, Cohen, BI, Dorland, W, Hammett, GW, Kim, C, Kinsey, JE, Kotschenreuther, M, Kritz, AH, Lao, LL, Mandrekas, J, Nevins, WM, Parker, SE, Redd, AJ, Shumaker, DE, Sydora, R & Weiland, J 2000 Comparisons and physics basis of tokamak transport models and turbulence simulations. *Phys. Plasmas* **7**, 969.

Dorland, W & Hammett, GW 1993 Gyrofluid turbulence models with kinetic effects. *Phys. Fluids B* **5**, 812.

Dorland, W, Jenko, F, Kotschenreuther, M & Rogers, BN 2000 Electron temperature gradient turbulence. *Phys. Rev. Lett.* **85**, 5579.

Drake, JF, Guzdar, PN & Hassam, AB 1988 Streamer formation in plasma with a temperature gradient. *Phys. Rev. Lett.* **61**, 2205.

Frieman, EA & Chen, L 1982 Nonlinear gyrokinetic equations for low-frequency electromagnetic waves in general plasma equilibria. *Phys. Fluids* **25**, 502.

Frisch, U 1995 *Turbulence: The Legacy of A. N. Kolmogorov*. CUP.

Gentle, KW & He, H 2008 Texas helimak. *Plasma Sci. Technol.* **10**, 284.

Hammett, GW, Beer, MA, Dorland, W, Cowley, SC & Smith, SA 1993 Developments in the gyrofluid approach to tokamak turbulence simulations. *Plasma Phys. Control. Fusion* **35**, 973.

Hasegawa, A & Mima, K 1978 Pseudo-three-dimensional turbulence in magnetized nonuniform plasma. *Phys. Fluids* **21**, 87.

Hasegawa, A & Wakatani, M 1983 Plasma edge turbulence. *Phys. Rev. Lett.* **50**, 682.

Helander, P & Sigmar, DJ 2002 *Collisional Transport in Magnetized Plasmas*. CUP.

Horton, W, Choi, D & Tang, WM 1981 Toroidal drift modes driven by ion pressure gradients. *Phys. Fluids* **24**, 1077.

Horton, W & Hasegawa, A 1994 Quasi-two-dimensional dynamics of plasmas and fluids. *Chaos* **4**, 227.

Hugill, J 1983 Transport in tokamaks – a review of experiment. *Nucl. Fusion* **23**, 331.

Jenko, F, Dorland, W, Kotschenreuther, M & Rogers, BN 2000 Electron temperature gradient driven turbulence. *Phys. Plasmas* **7**, 1904.

Kim, E & Diamond, PH 2002 Dynamics of zonal flow saturation in strong collisionless drift wave turbulence. *Phys. Plasmas* **9**, 4530.

Kinsey, JE, Waltz, RE & Candy, J 2005 Nonlinear gyrokinetic turbulence simulations of ExB shear quenching of transport. *Phys. Plasmas* **12**, 062302.

Kobayashi, S & Rogers, BN 2012 The quench rule, Dimits shift, and eigenmode localization by small-scale zonal flows. *Phys. Plasmas* **19**, 012315.

Kotschenreuther, M, Dorland, W, Beer, MA & Hammett, GW 1995*a* Quantitative predictions of tokamak energy confinement from first-principles simulations with kinetic effects. *Phys. Plasmas* **2**, 2381.

Kotschenreuther, M, Rewoldt, G & Tang, WM 1995*b* Comparison of initial value and eigenvalue codes for kinetic toroidal plasma instabilities. *Comput. Phys. Commun.* **88**, 128.

Kuo, H 1949 Dynamic instability of two-dimensional nondivergent flow in a barotropic atmosphere. *J. Meteorol.* **6**, 105.

Liewer, PC 1985 Measurements of microturbulence in tokamaks and comparisons with theories of turbulence and anomalous transport. *Nucl. Fusion* **25**, 543.

Majda, AJ, Qi, D & Cerfon, AJ 2018 A flux-balanced fluid model for collisional plasma edge turbulence: Model derivation and basic physical features. *Phys. Plasmas* **25**, 102307.

McMillan, BF, Hill, P, Bottino, A, Jolliet, S, Vernay, T & Villard, L 2011 Interaction of large scale flow structures with gyrokinetic turbulence. *Phys. Plasmas* **18**, 112503.

McMillan, BF, Jolliet, S, Tran, TM, Villard, L, Bottino, A & Angelino, P 2009 Avalanchelike bursts in global gyrokinetic simulations. *Phys. Plasmas* **16**, 022310.

McMillan, BF, Pringle, CCT & Teaca, B 2018 Simple advecting structures and the edge of chaos in subcritical tokamak plasmas. *J. Plasma Phys.* **84**, 905840611.

Mikhailovskii, AB & Tsypin, VS 1971 Transport equations and gradient instabilities in a high pressure collisional plasma. *Plasma Phys.* **13**, 785.

Newton, SL, Cowley, SC & Loureiro, NF 2010 Understanding the effect of sheared flow on microinstabilities. *Plasma Phys. Control. Fusion* **52**, 125001.

Numata, R, Ball, R & Dewar, RL 2007 Bifurcation in electrostatic resistive drift wave turbulence. *Phys. Plasmas* **14**, 102312.

Ongena, J, Koch, R, Wolf, R & Zohm, H 2016 Magnetic-confinement fusion. *Nature Phys.* **12**, 398.

Orszag, SA 1971 On the elimination of aliasing in finite-difference schemes by filtering high-wavenumber components. *J. Atmos. Sci.* **28**, 1074.

Parker, JB 2016 Dynamics of zonal flows: failure of wave-kinetic theory, and new geometrical optics approximations. *J. Plasma Phys.* **82**, 595820602.

Parker, JB & Krommes, JA 2013 Zonal flow as pattern formation. *Phys. Plasmas* **20**, 100703.

Parker, JB & Krommes, JA 2014 Generation of zonal flows through symmetry breaking of statistical homogeneity. *New J. Phys.* **16**, 035006.

Parra, FI, Barnes, M & Peeters, AG 2011 Up-down symmetry of the turbulent transport of toroidal angular momentum in tokamaks. *Phys. Plasmas* **18**, 062501.

Parra, FI & Catto, PJ 2009 Vorticity and intrinsic ambipolarity in turbulent tokamaks. *Plasma Phys. Control. Fusion* **51**, 095008.

Parra, FI & Catto, PJ 2010 Non-physical momentum sources in slab geometry gyrokinetics. *Plasma Phys. Control. Fusion* **52**, 085011.

Plunk, GG & Bañón Navarro, A 2017 Nonlinear growth of zonal flows by secondary instability in general magnetic geometry. *New J. Phys* **19**, 025009.

Plunk, GG, Cowley, SC, Schekochihin, AA & Tatsuno, T 2010 Two-dimensional gyrokinetic turbulence. *J. Fluid Mech.* **664**, 407.

Pogutse, OP 1968 Magnetic drift instability in a collisionless plasma. *Plasma Physics* **10**, 649.

Qi, D, Majda, AJ & Cerfon, AJ 2019 A flux-balanced fluid model for collisional plasma edge turbulence: Numerical simulations with different aspect ratios. *Phys. Plasmas* **26**, 082303.

Rath, F, Peeters, AG, Buchholz, R, Grosshauser, SR, Migliano, P, Weikl, A & Strintzi, D 2016 Comparison of gradient and flux driven gyro-kinetic turbulent transport. *Phys. Plasmas* **23**, 052309.

Rath, S & Sridhar, S 1992 Core instability of elliptic drift vortices. *Phys. Fluids B* **4**, 1367.

Ricci, P, Rogers, BN & Dorland, W 2006 Small-scale turbulence in a closed-field-line geometry. *Phys. Rev. Lett.* **97**, 245001.

Rogers, BN & Dorland, W 2005 Noncurvature-driven modes in a transport barrier. *Phys. Plasmas* **12**, 062511.

Rogers, BN, Dorland, W & Kotschenreuther, M 2000 Generation and stability of zonal flows in ion-temperature-gradient mode turbulence. *Phys. Rev. Lett.* **85**, 5336.

Rudakov, LI & Sagdeev, RZ 1961 On the instability of a nonuniform rarefied plasma in a strong magnetic field. *Dokl. Akad. Nauk SSSR* **138**, 581.

Ruiz, DE, Glinsky, ME & Dodin, IY 2019 Wave kinetic equation for inhomogeneous drift-wave turbulence beyond the quasilinear approximation. *J. Plasma Phys.* **85**, 905850101.

Ruiz, DE, Parker, JB, Shi, EL & Dodin, IY 2016 Zonal-flow dynamics from a phase-space perspective. *Phys. Plasmas* **23**, 122304.

Schekochihin, AA, Cowley, SC, Dorland, W, Hammett, GW, Howes, GG, Quataert, E & Tatsuno, T 2009 Astrophysical gyrokinetics: kinetic and fluid turbulent cascades in magnetized weakly collisional plasmas. *Astrophys. J. Suppl. S.* **182**, 310.

Srinivasan, K & Young, WR 2012 Zonostrophic instability. *J. Atmospheric Sci.* **69**, 1633.

St-Onge, DA 2017 On non-local energy transfer via zonal flow in the Dimits shift. *J. Plasma Phys.* **83**, 905830504.

Strintzi, D & Jenko, F 2007 On the relation between secondary and modulational instabilities. *Phys. Plasmas* **14**, 042305.

Sugama, H & Horton, W 1997 Transport processes and entropy production in toroidally rotating plasmas with electrostatic turbulence. *Phys. Plasmas* **4**, 405.

Sugama, H & Horton, W 1998 Nonlinear electromagnetic gyrokinetic equation for plasmas with large mean flows. *Phys. Plasmas* **5**, 2560.

Sugama, H, Okamoto, M, Horton, W & Wakatani, M 1996 Transport processes and entropy production in toroidal plasmas with gyrokinetic electromagnetic turbulence. *Phys. Plasmas* **3**, 2379.

Terry, PW & Horton, W 1983 Drift wave turbulence in a low-order k space. *Phys. Fluids* **26**, 106.

Vallis, GK 2017 *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation*, 2nd edn. CUP.

Villard, L, Angelino, P, Bottino, A, Brunner, S, Jolliet, S, McMillan, BF, Tran, TM & Vernay, T 2013 Global gyrokinetic ion temperature gradient turbulence simulations of ITER. *Plasma Phys. Control. Fusion* **55**, 074017.

Villard, L, McMillan, BF, Sauter, O, Hariri, F, Dominski, J, Merlo, G, Brunner, S & Tran, TM 2014 Turbulence and zonal flow structures in the core and L-mode pedestal of tokamak plasmas. *J. Phys. Conf. Ser.* **561**, 012022.

Waltz, RE 1988 Three-dimensional global numerical simulation of ion temperature gradient mode turbulence. *Phys. Fluids* **31**, 1962.

Waltz, RE, Dewar, RL & Garbet, X 1998 Theory and simulation of rotational shear stabilization of turbulence. *Phys. Plasmas* **5**, 1784.

Waltz, RE, Kerbel, GD & Milovich, J 1994 Toroidal gyro-Landau fluid model turbulence simulations in a nonlinear ballooning mode representation with radial modes. *Phys. Plasmas* **1**, 2229.

Wootton, AJ, Carreras, BA, Matsumoto, H, McGuire, K, Peebles, WA, Ritz, CP, Terry, PW & Zweben, SJ 1990 Fluctuations and anomalous transport in tokamaks. *Phys. Fluids B* **2**, 2879.

van Wyk, F, Highcock, EG, Field, AR, Roach, CM, Schekochihin, AA, Parra, FI & Dorland, W 2017 Ion-scale turbulence in MAST: anomalous transport, subcritical transitions, and comparison to BES measurements. *Plasma Phys. Control. Fusion* **59**, 114003.

van Wyk, F, Highcock, EG, Schekochihin, AA, Roach, CM, Field, AR & Dorland, W 2016 Transition to subcritical turbulence in a tokamak plasma. *J. Plasma Phys.* **82**, 905820609.

Zhou, Y, Zhu, H & Dodin, IY 2019 Formation of solitary zonal structures via the modulational instability of drift waves. *Plasma Phys. Control. Fusion* **61**, 075003.

Zhou, Y, Zhu, H & Dodin, IY 2020 Solitary zonal structures in subcritical drift waves: a minimum model. *Plasma Phys. Control. Fusion* **62**, 045021.

Zhu, H, Zhou, Y & Dodin, IY 2018*a* On the Rayleigh–Kuo criterion for the tertiary instability of zonal flows. *Phys. Plasmas* **25**, 082121.

Zhu, H, Zhou, Y & Dodin, IY 2018*b* On the structure of the drifton phase space and its relation to the Rayleigh–Kuo criterion of the zonal-flow stability. *Phys. Plasmas* **25**, 072121.

Zhu, H, Zhou, Y & Dodin, IY 2019 Nonlinear saturation and oscillations of collisionless zonal flows. *New J. Phys.* **21**, 063009.

Zhu, H, Zhou, Y & Dodin, IY 2020*a* Theory of the tertiary instability and the Dimits shift from reduced drift-wave models. *Phys. Rev. Lett.* **124**, 055002.

Zhu, H, Zhou, Y & Dodin, IY 2020*b* Theory of the tertiary instability and the dimits shift within a scalar model. *J. Plasma Phys.* **86**, 905860405.