Hydrodynamic convection in accretion discs

Loren E. Held[*] and Henrik N. Latter
Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK

ABSTRACT

The prevalence and consequences of convection perpendicular to the plane of accretion discs have been discussed for several decades. Recent simulations combining convection and the magnetorotational instability have given fresh impetus to the debate, as the interplay of the two processes can enhance angular momentum transport, at least in the optically thick outburst stage of dwarf novae. In this paper, we seek to isolate and understand the most generic features of disc convection, and so undertake its study in purely hydrodynamical models. First, we investigate the linear phase of the instability, obtaining estimates of the growth rates both semi-analytically, using one-dimensional spectral computations, and analytically, using WKBJ methods. Next, we perform three-dimensional, vertically stratified, shearing box simulations with the conservative, finite-volume code PLUTO, both with and without explicit diffusion coefficients. We find that hydrodynamic convection can, in general, drive outward angular momentum transport, a result that we confirm with ATHENA, an alternative finite-volume code. Moreover, we establish that the sign of the angular momentum flux is sensitive to the diffusivity of the numerical scheme. Finally, in sustained convection, whereby the system is continuously forced to an unstable state, we observe the formation of various coherent structures, including large-scale and oscillatory convective cells, zonal flows, and small vortices.

Key words: accretion, accretion discs – convection – hydrodynamics – instabilities, turbulence.

1 INTRODUCTION

Thermal convection – the bulk motion of fluid due to an entropy gradient – transports heat, mass, and angular momentum. There is compelling analytical, numerical, experimental, and observational evidence for convection both in the inner regions of stars and in the outer core and mantle of the Earth (Spruit, Nordlund & Title 1990; Aurnou et al. 2015), but its application to astrophysical discs is less clear. In particular, convection perpendicular to the plane of a disc bears marked differences to convection in most stars and planets. For example, the gravitational acceleration reverses sign at the disc mid-plane, and the gas supports a strong background shear flow. A third difference concerns the heat source necessary to drive the unstable entropy gradient. In stars, this heat is provided by nuclear fusion, while in the Earth's outer core convection is driven by both thermal and compositional gradients. In discs, heating at the mid-plane is normally assumed to be supplied by accretion (and the associated viscous dissipation of heat). Turbulence arising from the magnetorotational instability and dissipation of large-scale density waves are two mechanisms that might convert orbital energy into the required heat. Note importantly that the fluid motions associ-

ated with these heating processes may in fact impede the onset of convection, and at the very least interact with it in non-trivial ways.

It has long been speculated that convection perpendicular to the plane of the disc influences the optically thick, high-accreting outburst phase of dwarf novae (for a succinct review, see Cannizzo 1993). More recently, simulations of magnetorotational turbulence in stratified discs have shown that the MRI is capable of generating a convectively unstable entropy gradient. Moreover, three-dimensional, shearing box simulations with zero-net vertical magnetic flux reveal an interplay between convection and the MRI that might enhance angular momentum transport, typically quantified by the dimensionless parameter α (Bodo et al. 2013; Hirose et al. 2014). A similar interplay between magnetorotational instability and convection might operate in the ionized inner regions of protoplanetary discs, or during FU Orionis outbursts (Bell & Lin 1994; Hirose 2015).

In addition to dwarf novae and the inner regions of protoplanetary discs, hydrodynamic convection might drive activity in the weakly ionized 'dead zones' of protoplanetary discs where the magnetorotational instability is unlikely to operate (Armitage 2011). Although irradiation by the central star usually results in an isothermal vertical profile in protoplanetary discs (Chiang & Goldreich 1997; D'Alessio et al. 1998), heating by spiral density waves driven by a planet might nevertheless generate regions exhibiting a convectively

* E-mail: leh50@cam.ac.uk

unstable gradient (Boley & Durisen 2006; Lyra et al. 2016). So too might the heating from the dissipation of the strong zonal magnetic fields associated with the Hall effect (Lesur, Kunz & Fromang 2014).

As mentioned, accretion may instigate convection. But convection might drive accretion itself. Motivated by considerations of the primordial solar nebula, Lin & Papaloizou (1980) constructed a simple hydrodynamic model of a cooling young protoplanetary disc contracting towards the mid-plane. They showed that vertical convection arises if the opacity increases sufficiently fast with temperature, with the heating source initially provided by the gravitational contraction. But if convection is modelled via a mixing length theory, with an effective viscosity, the process might *self-sustain*: convective eddies could extract energy from the background orbital shear, and viscous dissipation of that energy might replace the initial gravitational contraction as a source of heat for maintaining convection.

In order to go beyond mixing length theory, Ruden, Papaloizou & Lin (1988) analysed linear axisymmetric modes in a thin, polytropic disc in the shearing box approximation and estimated that mixing of gas within convective eddies could result in values of $\alpha \sim 10^{-3}$–10^{-2}. They cautioned, however, that axisymmetric convective cells could not by themselves exchange angular momentum: such an exchange would have to be facilitated either by non-axisymmetric modes or by viscous dissipation of axisymmetric modes.

Following these early investigations, a debate ensued that centred not so much on the size of α but rather on its sign. Dissipation of *non-linear axisymmetric* convective cells was investigated by Kley, Papaloizou & Lin (1993), who performed quasi-global simulations (spanning about 100 stellar radii) of axisymmetric viscous discs with radiative heating and cooling. They measured an inward flux of angular momentum, but warned that this might be due to the imposed axial symmetry and to their relatively high viscosity. Convective shearing waves were first investigated by Ryu & Goodman (1992), who concluded that linear *non-axisymmetric* perturbations would result in a net *inward* angular momentum flux at sufficiently large time. These results were questioned however by Lin, Papaloizou & Kley (1993), who examined analytically and numerically a set of *localized* linear non-axisymmetric disturbances in global geometry. They demonstrated that these modes could transport angular momentum outwards in some cases, and opined that the inward transport of angular momentum reported by Ryu & Goodman (1992) was an artefact of the shearing box approximation.

Interest in hydro convection waned after local non-linear 3D compressible simulations showed that it resulted in *inward* rather than outward angular momentum transport. Using the finite-difference code ZEUS and rigid, isothermal vertical boundaries, Stone & Balbus (1996, hereafter SB96) initialized inviscid, fully compressible, and vertically stratified shearing box simulations with a convectively unstable vertical temperature profile, and measured a time-averaged value of $\alpha \sim -4.2 \times 10^{-5}$. Inward angular momentum transport was also reported by Cabot (1996), who ran simulations similar to those of SB96 but included full radiative transfer and a relatively high explicit viscosity. Analytical arguments for the inward transport of angular momentum by convection were presented by SB96 which, crucially, assumed axisymmetry, especially in the pressure field. However, in a rarely cited paper, Klahr, Henning & Kley (1999) presented fully compressible, three-dimensional, global simulations including explicit viscosity and radiative transfer of hydrodynamic convection in discs that gave some indication that non-linear convection in discs actually assumes non-axisymmetric patterns.

Some 15 years later, the claims of SB96 and of Cabot (1996) were called into question: fully local shearing box simulations of Boussinesq hydrodynamic convection in discs using the spectral code SNOOPY and employing explicit diffusion coefficients indicated that the sign of angular momentum transport due to vertical convection can be reversed provided the Rayleigh number (the ratio of buoyancy to kinematic viscosity and thermal diffusivity) is sufficiently large (Lesur & Ogilvie 2010). It would appear then that, in certain circumstances, vertical convection can drive outward angular momentum transport after all.

Our aim in this paper is to explore the competing results and claims concerning hydrodynamic convection in accretion discs, in particular to determine the sign and magnitude of angular momentum transport. We also isolate and characterize other generic features of convection that should be shared by multiple disc classes (dwarf novae, protoplanetary discs, etc.) and by different driving mechanisms (MRI turbulence, spiral shock heating, etc.). We do so both analytically and through numerical simulations, working in the fully compressible, vertically stratified shearing box approximation. We restrict ourselves to an idealized hydrodynamic set-up, omitting magnetic fields and complicated opacity transitions, and in so doing ensure our results are as general as possible.

The structure of the paper is as follows. First, in Section 2 we introduce the basic equations and provide a brief overview of the numerical code, the numerical parameters and set-up, and our main diagnostics. In Section 3, we investigate the linear behaviour of the convective instability, employing both analytical WKBJ and semi-analytical spectral methods to calculate the growth rates and the eigenfunctions. In Section 4, we explore the non-linear regime through *unforced* simulations, in which the convection is not sustained and is permitted to decay after non-linear saturation. Because this unforced convection is a transient phenomenon which might depend on the initial conditions, in Section 5 we explore the non-linear regime through simulations of *forced* convection, in which the internal energy is relaxed to its initial, convectively unstable, state to mimic, possibly, the action of background MRI turbulent dissipation and optically thick radiative cooling. Finally, Section 6 summarizes and discusses the results.

2 METHODS

2.1 Governing equations

We work in the shearing box approximation (Goldreich & Lynden-Bell 1965; Hawley, Gammie & Balbus 1995; Latter & Papaloizou 2017), which treats a local region of a disc as a Cartesian box located at some fiducial radius $r = r_0$ and orbiting with the angular frequency of the disc at that radius $\Omega_0 \equiv \Omega(r_0)$. A point in the box has Cartesian coordinates (x, y, z) that are related to the cylindrical coordinates (r, ϕ, z) through $x = r - r_0$, $y = r_0(\phi - \phi_0 - \Omega_0 t)$, and $z = z$. The equations of gas dynamics are now

$$\partial_t \rho + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{1}$$

$$\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} = -\frac{1}{\rho} \nabla P - 2\Omega_0 \mathbf{e}_z \times \mathbf{u} + \mathbf{g}_{\text{eff}} + \frac{1}{\rho} \nabla \cdot \mathbf{T}, \tag{2}$$

$$\partial_t(\rho e) + \mathbf{u} \cdot \nabla(\rho e) = -\gamma \rho e \nabla \cdot \mathbf{u} + \mathbf{T} : \nabla \mathbf{u} + \xi \nabla^2 T + \Lambda_{\text{relax}}, \tag{3}$$

corresponding to conservation of mass, momentum, and thermal energy, respectively. These equations are closed with a perfect gas equation of state $P = (\mathcal{R}/\mu)\rho T$ or, equivalently, $P = e(\gamma - 1)\rho$ in terms of the specific internal energy e. Here, P is the thermal pressure

of the fluid, T is the temperature, ρ is mass density, $\mathcal{R}$ is the gas constant, and μ is the mean molecular weight. The adiabatic index (ratio of specific heats) is denoted by γ and is typically taken to be $\gamma = 5/3$ in all our simulations.[1] The effective gravitational potential is embodied in the tidal acceleration $\mathbf{g}_{\mathrm{eff}} = 2q\Omega_0^2 x\hat{\mathbf{x}} - \Omega_0^2 z\hat{\mathbf{z}}$, where q is the shear rate. For Keplerian discs $q = 3/2$ a value we adopt throughout the paper.

Other terms in the equations include the viscous stress tensor $T \equiv 2\rho\nu\mathbf{S}$. Here, ν is the kinematic viscosity and $\mathbf{S} \equiv (1/2)[\nabla\mathbf{u} + (\nabla\mathbf{u})^{\mathrm{T}}] - (1/3)(\nabla \cdot \mathbf{u})\mathbf{I}$ is the traceless shear tensor. The thermal conductivity is ξ, but in our simulations we specify thermal diffusivity χ rather than ξ. The former is related to the latter via $\chi = \xi/(c_p\rho)$, where c_p is the specific heat capacity at constant pressure.

In some of our simulations, we include heating and cooling through a *thermal relaxation* term $\Lambda_{\mathrm{relax}} \equiv (\rho e - \rho_0 e_0)/\tau_{\mathrm{relax}}$. This relaxes the thermal energy in each cell back to its initial state $e_0(z) \equiv e(z, t = 0)$ on a time-scale given by τ_{relax}. We refer to simulations run with $\Lambda_{\mathrm{relax}} \neq 0$ as simulations of *forced* compressible convection. Conversely, simulations with $\Lambda_{\mathrm{relax}} = 0$ are referred to as *unforced*. The thermal relaxation time-scale τ_{relax} is taken to be the inverse of the linear growth rate of the convective instability.

2.2 Important parameters and instability criteria

The onset of thermal convection is controlled by the local *buoyancy frequency*, defined as

$$N_B^2 = z\left[\frac{1}{\gamma}\frac{\partial \ln P}{\partial z} - \frac{\partial \ln \rho}{\partial z}\right]\Omega_0^2. \tag{4}$$

But convection is opposed by the effects of viscosity and thermal diffusivity, with the ratio of destabilizing and stabilizing processes quantified by the *Rayleigh number*

$$\mathrm{Ra} \equiv \frac{N_B^2 H^4}{\nu\chi}, \tag{5}$$

where H is the disc's scale height (formally defined later). On the other hand, the ratio of buoyancy to shear is expressed through the *Richardson number*

$$\mathrm{Ri} \equiv \frac{N_B^2}{q^2\Omega_0^2}. \tag{6}$$

For convection perpendicular to the plane of an accretion disc, the buoyancy force is perpendicular to the direction of the background shear. Thus, the Richardson number in our set-up is more an expression for the scaled intensity of N_B^2. Finally, the *Prandtl number* is defined as the ratio of kinematic viscosity to thermal diffusivity

$$\mathrm{Pr} \equiv \frac{\nu}{\chi}. \tag{7}$$

When the explicit diffusion coefficients are neglected (ν, $\chi = 0$), convective instability occurs when the buoyancy frequency is negative (i.e. in regions where $N_B^2 < 0$). When explicit viscosity and thermal diffusivity are included, however, convective instability requires both that $N_B^2 < 0$ *and* that the Rayleigh number exceeds some critical value $\mathrm{Ra_c}$. Though in real astrophysical discs the viscosity is negligible, the inclusion of magnetic fields complicates the stability criterion, as does the presence of preexisting turbulence (as

---

[1] Technically $\gamma = 5/3$ is valid for *ionized* hydrogen. Since all our simulations are hydrodynamic it might be more appropriate to use $\gamma = 7/5$, which is valid for molecular hydrogen, but the differences are small and do not affect our results.

might be supplied by the MRI) which itself may diffuse momentum and heat. In the latter case, it may be possible to define a turbulent 'Rayleigh number', which must be sufficiently large so that convection resists the disordered background flow. In any case, the sign of N_B^2 is certainly insufficient to assign convection to MRI-turbulent flows, as is often done in recent work.

2.3 Numerical set-up

2.3.1 Codes

Most of our simulations are run with the conservative, finite-volume code PLUTO (Mignone et al. 2007). Unless stated otherwise, we employ the Roe Riemann solver, 3rd order in space WENO interpolation, and the 3rd order in time Runge–Kutta algorithm. Other configurations are explored in Section 4. To allow for longer time-steps, we employ the FARGO scheme (Mignone et al. 2012). When explicit viscosity ν and thermal diffusivity χ are included, we further reduce the computational time by using the Super-Time-Stepping (STS) scheme (Alexiades, Amiez & Gremaud 1996). Note that due to the code's conservative form kinetic energy is not lost to the grid but converted directly into thermal energy. Ghost zones are used to implement the boundary conditions.

We use the built-in shearing box module in PLUTO (Mignone et al. 2012). Rather than solving equations (1)–(3) (primitive form), PLUTO solves the governing equations in *conservative form*, evolving the total energy equation rather than the thermal energy equation, where the total energy density of the fluid is given by $E = (1/2)\rho u^2 + \rho e$ (kinetic + internal). In PLUTO the thermal relaxation term $\Lambda_{\mathrm{relax}} = (\rho e - \rho_0 e_0)/\tau_{\mathrm{relax}}$ is not implemented directly in the total energy equation. Instead, it is included on the right-hand side of the thermal energy equation, which is then integrated (in time) analytically.

Finally, we have verified the results of our fiducial simulation (presented in Section 4.1) with the conservative finite-volume code ATHENA (Stone et al. 2008).

2.3.2 Initial conditions and units

We use two different convectively unstable profiles to initialize our simulations. The first exhibits a Gaussian temperature profile

$$T = T_0\exp\left[-\frac{z^2}{\beta H^2}\right], \tag{8}$$

where T_0 is mid-plane temperature and β is a dimensionless tuning parameter. See Fig. 1 for all associated thermal profiles. For comparison, we also use the profile introduced by SB96 in which the temperature follows a power law

$$T = T_0 - Az^p, \tag{9}$$

where A and p are parameters, with $p = 3/2$ usually (see Fig. 2). Further details of both profiles are given in Appendix A. The equilibria are convectively unstable within a confined region (of size L_c) about the disc mid-plane, and convectively stable outside this region. We have checked that these vertical profiles are indeed good equilibrium solutions in our numerical set-up by running a series of simulations without perturbations; these show that after 55 orbits the initial profiles are unchanged with velocity fluctuation amplitudes typically less than 10^{-3} at the vertical boundaries and less than 10^{-7} at the mid-plane.

The background velocity is given by $\mathbf{u} = -(3/2)\Omega_0 x\,\mathbf{e}_y$. At initialization, we usually perturb all the velocity components with random noise exhibiting a flat power spectrum. The perturbations
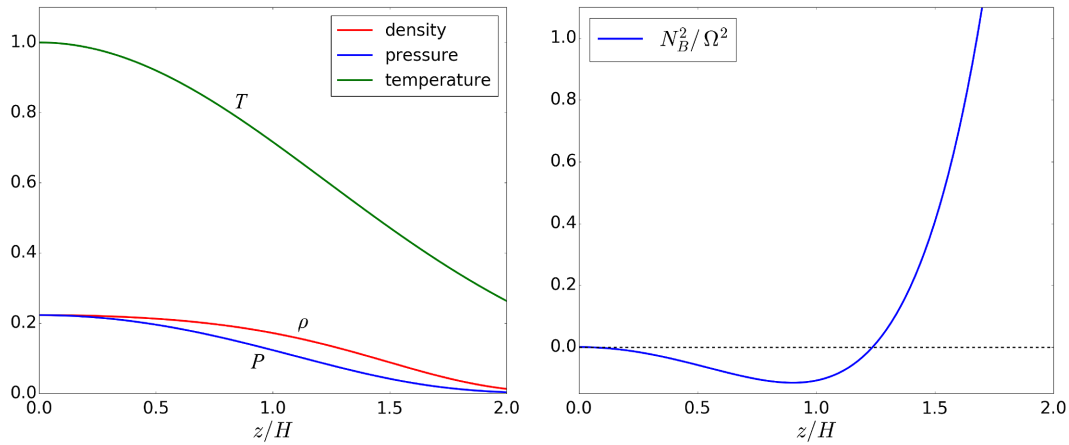
Figure 1. Vertical disc structure for a Gaussian temperature profile in a disc of height $2H$ about the mid-plane. *Left*: vertical profiles for density ρ, pressure P, and temperature T. *Right*: vertical profile of the buoyancy frequency N_B^2/Ω^2. For clarity, only half of the vertical domain is shown. The profile parameters are $\{T_0 = 1.0, \rho_0 = 1.0, \beta = 3.0\}$ and the adiabatic index is $\gamma = 5/3$.
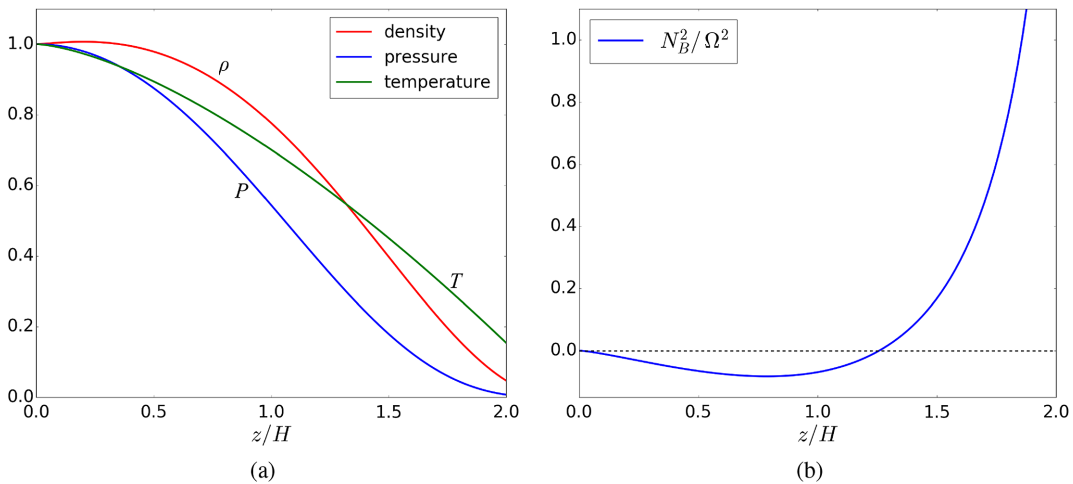

Figure 2. Vertical disc structure of Stone & Balbus (1996) in a disc of height $2H$ about the mid-plane. *Left*: vertical profiles for density ρ, pressure P, and temperature T. *Right*: vertical profile of the buoyancy frequency N_B^2/Ω^2. For clarity, only half of the vertical domain is shown. The profile parameters are $\{T_0 = 1.0, \rho_0 = 1.0, g = 5.0\}$ and the adiabatic index is $\gamma = 5/3$.

$\delta\mathbf{u}$ have maximum relative amplitude of about 5×10^{-5} and can be either positive or negative. In order to investigate specifically the nature of linear axisymmetric convective modes, we initialize several PLUTO simulations with linear axisymmetric modes calculated semi-analytically rather than with random noise (see Section 3.5).

Time units are selected so that $\Omega_0 = 1$. From now, the subscript on the angular frequency is dropped if it appears. The length unit is chosen so that the initial mid-plane sound speed $c_s = 1$, which in turn defines a constant reference scale height $H \equiv c_s/\Omega_0 = 1$. Note, however, that the sound speed is generally a function of both space and time.

2.3.3 Box size and resolution

We measure box size in units of scale height H, defined above. We employ resolutions of 64^3, 128^3, 256^3, and 512^3 in boxes of size $4H \times 4H \times 4H$, which correspond to resolutions of 16, 32, 64, and 128 grid cells per scale height H in all directions. However, in simulations in which we force convection, we employ a 'wide-box' ($6H \times 6H \times 4H$) with a resolution of $256 \times 256 \times 256$ which

corresponds to about 43 grid cells per H in the horizontal direction and 64 grid cells per H in the vertical direction.

2.3.4 Boundary conditions

We use shear-periodic boundary conditions in the *x*-direction (see Hawley et al. 1995) and periodic boundary conditions in the *y*-direction. In the vertical direction, we keep the ghost zones associated with the thermal variables in isothermal hydrostatic equilibrium (the temperature of the ghost zones being kept equal to the temperature of the vertical boundaries at initialization) in the manner described in Zingale et al. (2002). For the velocity components, we use mostly standard *outflow* boundary conditions in the vertical direction, whereby the vertical gradients of all velocity components are zero, and variables in the ghost zones are set equal to those in the active cells bordering the ghost zones. In a handful of our simulations, we also use free-slip and periodic boundary conditions to test the robustness of our results.

2.4 Diagnostics

2.4.1 Averaged quantities

We follow the time evolution of various volume-averaged quantities (e.g. kinetic energy density, thermal energy density, Reynolds stress, mass density, thermal pressure, and temperature). For a quantity X, the volume-average of that quantity is denoted $\langle X \rangle$ and is defined as

$$\langle X \rangle(t) \equiv \frac{1}{V} \int_V X(x, y, z, t) \mathrm{d}V, \tag{10}$$

where V is the volume of the box.

If we are interested only in the vertical structure of a quantity X then we average over the x- and y-directions, only. The horizontal average of that quantity is denoted $\langle X \rangle_{xy}$ and is defined as

$$\langle X \rangle_{xy}(z, t) \equiv \frac{1}{A} \int_A X(x, y, z, t) \mathrm{d}A, \tag{11}$$

where A is the horizontal area of the box. Horizontal averages over different coordinate directions (e.g. over the y- and z-directions) are defined in a similar manner.

We are also interested in averaging certain quantities (e.g. the Reynolds stress) over time. The temporal average of a quantity X is denoted $\langle X \rangle_t$ and is defined as

$$\langle X \rangle_t(x, y, z) \equiv \frac{1}{\Delta t} \int_{t_i}^{t_f} X(x, y, z, t) \mathrm{d}t, \tag{12}$$

where we integrate from some initial time t_i to some final time t_f and $\Delta t \equiv t_f - t_i$.

2.4.2 Reynolds stress, alpha, and energy densities

In accretion discs, the radial transport of angular momentum is related to the xy-component of the Reynolds stress, defined as

$$R_{xy} \equiv \rho u_x \delta u_y, \tag{13}$$

where $\delta u_y \equiv u_y + q\Omega x$ is the fluctuating part of the y-component of the total velocity u_y. $R_{xy} > 0$ corresponds to positive (i.e. radially outward) angular momentum transport, while $R_{xy} < 0$ corresponds to negative (i.e. radially inward) angular momentum transport.

The Reynolds stress is related to the classic dimensionless parameter α, but the reader should note that various definitions of α exist in the literature and care must be taken when comparing measurements of α quoted in different papers. We define α as

$$\alpha \equiv \frac{\langle R_{xy} \rangle}{q \langle P \rangle}. \tag{14}$$

The kinetic energy density of a fluid element is defined as

$$E_{\mathrm{kin}} \equiv \frac{1}{2} \rho u^2, \tag{15}$$

where u is the magnitude of the *total* velocity of a fluid element. Often we will plot the vertical kinetic energy density, in which case u in the above is replaced by u_z.

The total energy (kinetic plus thermal) is *not* conserved in the shearing box even when there are no (explicit or numerical) dissipative effects: work done by the effective gravitational potential energy on the fluid and flux of angular momentum through the x-boundaries due to the Reynolds stress act as sources for the total energy. The upper and lower vertical boundaries also let energy (and other quantities) leave the box.

2.4.3 Vertical profiles of horizontally averaged quantities

Finally, we track the vertical profiles of horizontally averaged pressure, density, temperature, and Reynolds stress. Of special interest is the buoyancy frequency, which is calculated from the pressure and density data by finite differencing the formula

$$\frac{N_B^2}{\Omega^2} = z \left[\frac{1}{\gamma} \frac{\mathrm{d} \ln \langle P \rangle_{xy}}{\mathrm{d}z} - \frac{\mathrm{d} \ln \langle \rho \rangle_{xy}}{\mathrm{d}z} \right]. \tag{16}$$

We also calculate the (horizontally averaged) vertical profiles of mass and heat flux. We define the mass flux as

$$F_{\mathrm{mass}} = \langle \rho u_z \rangle_{xy}, \tag{17}$$

and the heat flux as

$$F_{\mathrm{heat}} = \langle \rho u_z T \rangle_{xy}. \tag{18}$$

3 LINEAR THEORY

In this section, we investigate the linear phase of the axisymmetric convective instability both semi-analytically by solving a 1D boundary value/eigenvalue problem using spectral methods, and also analytically using WKBJ methods.

Our aim is to determine the eigenvalues (growth rates) and eigenfunctions (density, velocity, and pressure perturbations) of the axisymmetric convective instability in the shearing box as a function of radial wavenumber k_x. The convectively unstable background vertical profile used in all calculations is shown in Fig. 1. The profile is discussed in detail in Section A2. All calculations were carried out with profile parameters $T_0 = 1.0$, $\rho_0 = 1.0$, $\beta = 3.0$ and an adiabatic index of $\gamma = 5/3$. For our chosen background equilibrium, this corresponds to a Richardson number of Ri ~ 0.05.

We proceed as follows: first, we produce linearized equations for the perturbed variables; secondly, these are Fourier analysed so that $X'(x, z, t) = e^{\sigma t} e^{i k_x x} X'(z)$ for each of the perturbed fluid variables ρ', $\mathbf{u}'$, and p'. Here, σ can be complex (if σ is real and positive, it is referred to as a *growth rate*). The linearized equations are given by

$$\sigma \rho' = -i k_x \rho_0 u_x' - \partial_z (\rho_0 u_z'), \tag{19}$$

$$\sigma u_x' = 2\Omega u_y' - \frac{i k_x}{\rho_0} P', \tag{20}$$

$$\sigma u_y' = (q - 2)\Omega u_x', \tag{21}$$

$$\sigma u_z' = -\frac{1}{\rho_0} \partial_z P' + \frac{\rho'}{\rho_0^2} \partial_z P_0, \tag{22}$$

$$\sigma P' = -u_z' \partial_z P_0 - \gamma P_0 (i k_x u_x' + \partial_z u_z'). \tag{23}$$

Background fluid quantities are denoted $X_0 = X_0(z)$. For equilibrium ρ_0 and P_0 see Appendix A.

3.1 Spectral methods

We discretize the fluid variables on a vertical grid containing N_z points by expanding each eigenfunction as a linear superposition of Whittaker cardinal (sinc) functions. These are appropriate as all perturbations should decay to zero far from the convectively unstable region. This choice also saves us from explicitly imposing boundary conditions. The discretized equations are next gathered up into a single algebraic eigenvalue problem (Boyd 2001). Finally, we solve this matrix equation numerically using the QR algorithm
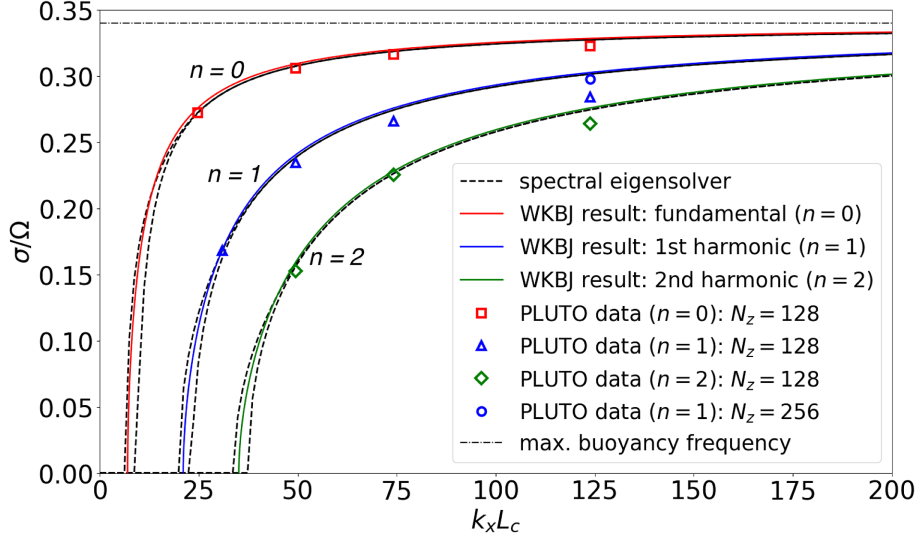
Figure 3. Growth rates as a function of radial wavenumber k_x (scaled by the width of the convectively unstable region L_c). The solid lines were calculated analytically using WKBJ methods and the dashed lines were calculated semi-analytically using pseudo-spectral methods. The squares, triangles, and diamonds correspond to measurements taken from PLUTO simulations run at a vertical resolution of $N_z = 128$. The blue circle was taken from a PLUTO simulation run at $N_z = 256$.

to obtain the growth rates σ and eigenfunctions for a given radial wavenumber k_x. The code that computes the eigensolution we often refer to as our 'spectral eigensolver'.

3.2 WKBJ approach

Alternatively, equations (19)–(23) can be combined into a second-order ODE of the form

$$\frac{\mathrm{d}^2 U}{\mathrm{d} z^2} + k_z^2(z) U = 0, \tag{24}$$

where

$$U = \left(\frac{\gamma P_0}{\sigma^2 + \kappa^2 + k_x^2 c_0^2} \right)^{1/2} u_z', \tag{25}$$

in which c_0 is the equilibrium sound speed and the 'vertical wavenumber' of the perturbations is given by

$$k_z^2(z) \approx -k_x^2 \frac{N_B^2(z) + \sigma^2}{\kappa^2 + \sigma^2}, \tag{26}$$

with κ the epicyclic frequency (Ruden et al. 1988).

We obtain approximate solutions to equation (24) analytically using WKBJ methods in the limit k_x large. Equation (24) has two turning points at $k_z^2(z) = 0$, which occur when $-N_B^2(z) = \sigma^2$. Physically, one turning point occurs at the disc mid-plane ($z = 0$), while the other turning point occurs at the boundary of the convectively unstable region $z = L_c$. When $-N_B^2(z) > \sigma^2$ then $k_z^2(z) > 0$, so the solutions of equation (24) are spatially oscillatory, otherwise they are evanescent. The unstable modes, in which we are interested, are confined and oscillatory within the convectively unstable region, and decay exponentially in space outside.

The standard WKBJ solution is given by

$$U(z) \sim c_1 \exp\left(+\mathrm{i} \int^z k_z(z') \mathrm{d} z' \right) + c_2 \exp\left(-\mathrm{i} \int^z k_z(z') \mathrm{d} z' \right),$$

where c_1 and c_2 are constants. By matching the interior (oscillatory) solution on to the exterior (exponentially decaying) solution, we

obtain the eigenvalue equation (dispersion relation):

$$\int_{z_0}^{z_1} k_z(z) \mathrm{d} z = \left(n + \frac{1}{2} \right) \pi \quad n = 0, 1, 2, 3, \cdots, \tag{27}$$

where lower and upper bounds of integration z_0 and z_1, respectively, are related by the implicit equation $N_B^2(z_0) = N_B^2(z_1) = -\sigma^2$, and n is the mode number. Substituting for $k_z^2(z)$ using equation (26), we obtain an implicit equation relating the radial wavenumber k_x and the growth rate σ which we solve numerically via a root-finding algorithm.

3.3 Eigenvalues

For a given radial wavenumber k_x, the eigenvalues calculated with the spectral eigensolver come in pairs that lie close together in the complex plane. These correspond physically to the *even* and *odd* modes of the convective instability. As described in Section 3.4, even modes are symmetric about the mid-plane, whereas the odd modes are antisymmetric about the mid-plane. In Fig. 3, we plot as black dashed curves the first few eigenvalue branches as functions of k_x, these corresponding to vertical quantum numbers $n = 0, 1,$ and 2. Note that in the figure k_x has been normalized by the width of the convectively unstable region L_c (given by equation A5 in Appendix A2).

As the radial wavenumber k_x is increased, the difference between the growth rates of the even and odd modes vanishes, and the growth rates tend asymptotically to the maximum absolute value of the buoyancy frequency (indicated by a horizontal dot–dashed line). In other words, the maximum growth rate is limited by the depth of the buoyancy-frequency profile. The fastest growing modes are at arbitrarily large k_x (small radial wavenumber λ_x), and thus manifest as thin elongated structures. Thicker structures are not favoured as they comprise greater radial displacements that are resisted by the radial angular momentum gradient.

Superimposed on the spectrally computed values are the WKBJ growth rates (solid lines). WKBJ methods cannot distinguish be-
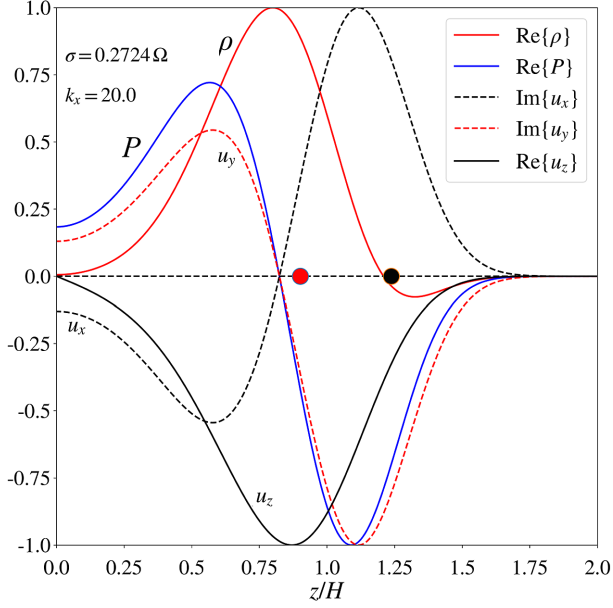
Figure 4. Vertical profiles of the perturbations $\text{Re}\{\rho'(z)\}$, $\text{Re}\{u'_z(z)\}$, $\text{Re}\{P'(z)\}$, and of $\text{Im}\{u'_x(z)\}$ and $\text{Im}\{u'_y(z)\}$ in the upper half plane $z \in [0, 2]H$, for the $n = 0$ odd mode at $k_x = 20.0$. The imaginary and real parts not shown are effectively zero. The black dot marks the upper extent of the convectively unstable region L_c, while the red dot marks the most convectively unstable points (see equation A5). The eigenfunctions have been rescaled so that the maximum of each eigenfunction is unity.

tween even or odd modes, nonetheless we find very close agreement between the WKBJ and semi-analytical results for all radial wavenumbers, even at low k_x where the WKBJ approximation is, strictly speaking, invalid. As the radial wavenumber increases, the semi-analytical and WKBJ results converge. Note that associated with each mode branch is a minimum radial wavenumber $k_{x,\,\text{min}}$ below which both the numerical and WKBJ growth rates are zero. This feature is just visible in the bottom left-hand side of Fig. 3.

We must stress that, because the fastest growing modes are on the shortest possible scales, inviscid simulations of convection are problematic, at least in the early (linear) phase of the evolution. It is here that the simulated behaviour may depend on the varying ways that different numerical schemes deal with grid diffusion, something we explore in Section 4.3. Only with resolved physical viscosity can this problem be overcome. However, in the fully developed non-linear phase of convection the short scale linear axisymmetric modes may be subordinate and this less an issue.

3.4 Vertical structure of the eigenfunctions

The reader is reminded that here we have an unstably stratified fluid in a medium in which the gravitational field smoothly changes sign. This contrasts to the more commonly studied situation of convection at the surface of the Earth (or within the solar interior). It hence is of interest to study the structure of the convective cells in some detail. We consider first the fundamental ($n = 0$) odd mode with a radial wavenumber $k_x = 20$.

In Fig. 4, we plot the vertical profiles of the non-zero components of the eigenfunction. To provide greater visual clarity, we have rescaled the perturbations so that the maximum amplitude of each perturbation is unity: thus, we can more easily observe the vertical structure of the perturbations, but not deduce their relative amplitudes. In Fig. 5, we plot their real parts in the xz-plane with the correct amplitudes. We find that $|u'_z| > |u'_y| \sim |u'_x| \gg |\rho'| \gg |P'|$. Thus, vertical velocity perturbations are greatest in magnitude, while pressure perturbations are smaller than vertical velocity perturbations by two orders of magnitude, indicating that the convective cells are roughly in pressure balance with the surroundings, i.e. that they are behaving adiabatically.

In Fig. 4, u'_z (solid black curve) is antisymmetric about the midplane, whereas the remaining perturbations are symmetric. Figs 4 and 5 clearly show that both odd and even modes consist of a radial chain of convective cells above and below the mid-plane and localized near the most convectively unstable point (denoted by the large red dot in the former figure). The peak in the amplitude of u'_z that occurs near this point tells us that the fluid elements reach
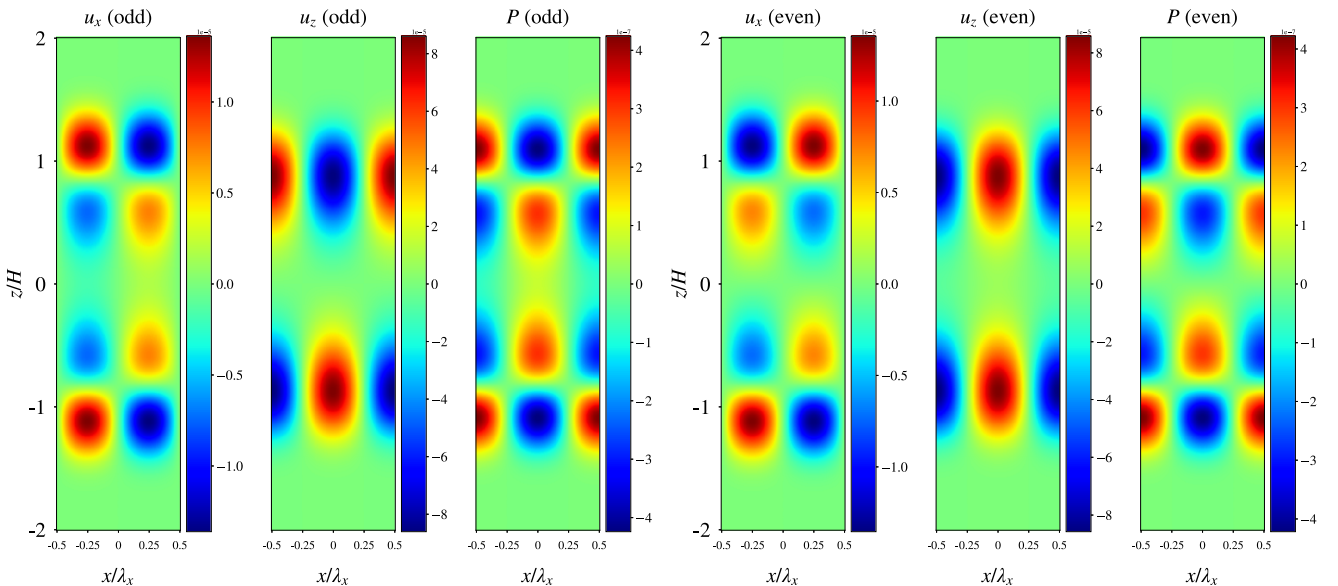

Figure 5. Two-dimensional profiles in the xz-plane of the real parts of the selected components of odd and even eigenfunctions for $n = 0$ and $k_x = 20.0$.

their maximum acceleration where the buoyancy force is greatest. As the fluid perturbations behave adiabatically, they adjust their pressure to maintain a balance with the background pressure: thus, where cool elements begin to rise (higher background pressure), the pressure perturbations (solid blue line) increase (i.e. $P' > 0$), and where hot elements begin to sink (lower background pressure), the pressure perturbations decrease (i.e. $P' < 0$). The vertical velocity perturbation (black solid line) is out of phase with the radial and azimuthal velocity perturbations (dashed lines), since vertical motion is converted into radial motion where the convective cells turn over (picture fluid motion at the top of a fountain).

The even modes possess a non-zero, but relatively small, u'_z at the mid-plane, and thus weakly couple the two sides of the disc. Such modes may permit some exchange of mass, momentum, and thermal energy across the mid-plane when reaching large amplitudes (in particular, see Section 5.2). In particular, falling fluid elements may 'overshoot' the midplane and thus circulate from one side of the disk to the other.

As the radial wavenumber is increased from $k_x = 20.0$ to 200.0 (very small radial wavelengths), the perturbations become increasingly localized about the most convectively unstable point(s). The localization in z however is not as narrow as the mode's radial wavelength. For larger k_x, in fact, each pair of even and odd modes changes their character and becomes entirely localized to either the upper or lower disc. Activity in the two halves of the disc is hence completely decoupled for such small-scale (but fast growing) disturbances.

3.5 Comparison of theory with simulations

In this section, we compare the growth rates previously calculated with those measured from PLUTO simulations initialized with the exact linear modes taken from the spectral eigensolver.

Each PLUTO simulation is run in a box of size $\lambda_x \times 0.625H \times 4H$. We maintain a fixed number of 32 grid cells per scale height in the z-direction, 16 grid cells per scale height in the y-direction, and 64 grid cells per *radial wavelength* in the x-direction. Thus, the radial wavelength of the modes is always well resolved in our simulations. No explicit diffusion coefficients (i.e. viscosity or thermal diffusivity) are used.

The results are plotted in Fig. 3 as squares (and one circle). Numerical values appear in Table 1. We find excellent agreement between simulation and theory: the relative error is < 1 per cent for the fundamental mode in the interval $k_x \in [20.0, 60.0]$ and for the first harmonic at $k_x = 25.0$. As the radial wavenumber is increased the simulation growth rates begin to deviate from the theoretical curves, although the percentage error remains less than 6 per cent even at large wavenumbers. This behaviour is expected due to the effects of numerical diffusion which acts to *decrease* the growth rates as the radial wavelength approaches the size of the vertical grid. Although the radial wavenumber λ_x is always well resolved in the simulations, the vertical wavelength becomes increasingly less so.

The numerical diffusion could be reduced by increasing the vertical resolution of the simulations. To check this we have rerun the PLUTO simulation, initialized with eigenfunctions corresponding to the first harmonic ($n = 1$) even mode with $k_x = 100$, at *twice* the vertical resolution, i.e. $N_z = 256$ instead of $N_z = 128$. At a resolution of $N_z = 128$, we measured a growth rate of $\sigma_{\mathrm{PLUTO}} = 0.2840\,\Omega$, corresponding to a percentage error of 5.8 per cent with the growth rate calculated semi-analytically ($\sigma_{\mathrm{EIG}} = 0.3016\,\Omega$). At a vertical resolution of $N_z = 256$, however, we measured a growth

Table 1. Comparison of eigensolver growth rates with growth rates measured in PLUTO simulations. The simulations were initialized with the eigenfunctions corresponding to each mode n and radial wavenumber k_x. All simulations except the one marked with a footnote were run at a resolution of $64 \times 10 \times 128$. The standard deviation on the growth rates σ_{PLUTO} measured in the simulations is typically in the range from ± 0.0002 to ± 0.0004.

Mode	Parity	k_x	σ_{EIG}	σ_{PLUTO}	Percentage error
$n = 0$	even	20.0	0.2727	0.2719	0.3
	odd	20.0	0.2724	0.2715	0.3
	even	40.0	0.3076	0.3059	0.6
	even	60.0	0.3186	0.3160	0.8
	even	100.0	0.3273	0.3223	1.6
$n = 1$	even	25.0	0.1694	0.1678	0.9
	odd	25.0	0.1643	0.1624	0.9
	even	40.0	0.2392	0.2343	1.1
	even	60.0	0.2747	0.2655	3.3
	even	100.0	0.3016	0.2840	5.8
	–	–	–	0.2974	1.4^a
$n = 2$	even	40.0	0.1593	0.1525	4.2
	odd	40.0	0.1580	0.1517	3.9
	even	60.0	0.2272	0.2251	0.9
	even	100.0	0.2749	0.2637	4

Note: ^{a} Simulation run with a resolution of $N_z = 256$ grid points in the vertical direction.

rate of $\sigma_{\mathrm{PLUTO}} = 0.2974\,\Omega$, corresponding to a percentage error of just 1.4 per cent. Thus, the growth rates measured in the simulations appear to converge to those calculated analytically and semi-analytically as the resolution of the simulations is increased.

Finally, in all simulations of the axisymmetric modes, we observed *inward* angular momentum transport. This is in agreement with the analytical argument presented in Stone & Balbus (1996) regarding axisymmetric flow.

4 SIMULATIONS OF UNFORCED COMPRESSIBLE CONVECTION

In this section, we describe fully compressible, three-dimensional simulations of hydrodynamic convection in the shearing box carried out with PLUTO. These simulations are 'unforced' in that they are initialized with a convectively unstable profile, but this unstable profile is not maintained by a separate process (e.g. turbulent heating via the magnetorotational instability). Thus, convection rearranges the profile, shifting it to a marginally stable state, and thus ultimately hydrodynamical activity dies out. While the unforced convection studied in this section is essentially a transient phenomenon that depends on initial conditions, and is sensitive to the numerical scheme (as we discuss in Sections 4.2 and 4.3), it serves as a good starting point for investigating the problem and illuminates a number of interesting features.

The key result of this section is that, even without the inclusion of explicit viscosity, we observe that hydrodynamic convection in the shearing box generally produces a *positive Reynolds stress*, and thus can drive outward transport of angular momentum. This is in direct contradiction to simulations carried out in ZEUS by SB96. In Section 4.1, we describe our fiducial simulations, and in Section 4.2 we compare ATHENA and PLUTO simulations. In Section 4.3, we describe the sensitivity of the sign of angular momentum transport to the numerical scheme. Finally, in Section 4.4, we investigate the
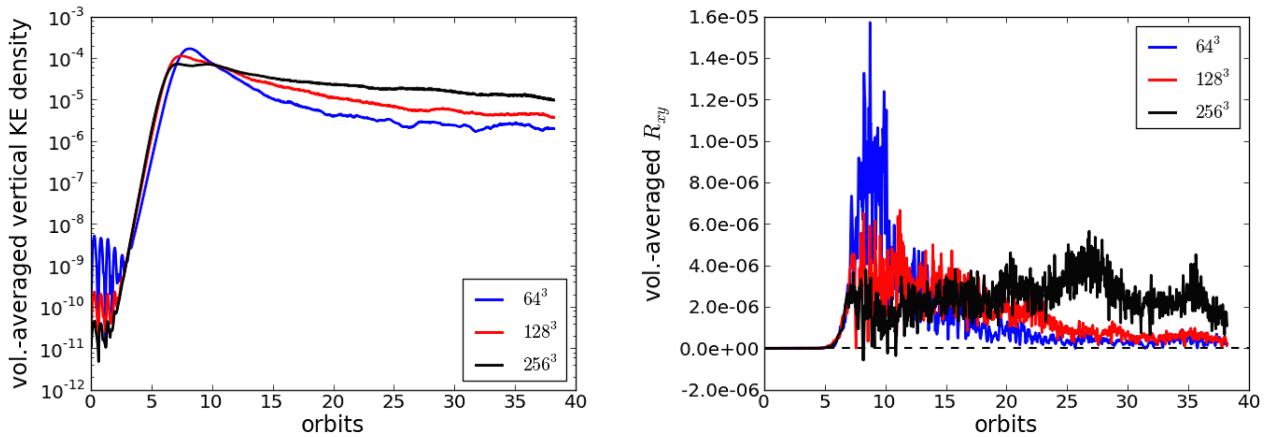
Figure 6. Left: semilog plot of the time evolution of the volume-averaged vertical kinetic energy density in PLUTO simulations at resolutions of 64^3 (blue line), 128^3 (red line), and 256^3 (black line). Right: time evolution of volume-averaged xy-component of the Reynolds stress tensor.

effects of including explicit diffusion coefficients in our simulations, thus connecting to the Boussinesq simulations of Lesur & Ogilvie (2010).

4.1 Fiducial simulations

4.1.1 Set-up and initialization of fiducial simulations

The simulations described in this section were run at resolutions of 64^3, 128^3, and 256^3 in boxes of size $4H \times 4H \times 4H$, where H is the scale height.[2] All the simulations were initialized with the convectively unstable vertical profiles for density and pressure described in Appendix A2 with profile parameters $T_0 = 1.0$, $\rho_0 = 1.0$, $\beta = 3.0$ and an adiabatic index of $\gamma = 5/3$ (see Fig. 1). The Stone & Balbus vertical profile SB96 was also trialled yielding very similar results, most of which have been omitted in the interests of space. Vertical outflow conditions were employed for our fiducial simulations but periodic and free-slip boundary conditions in the vertical direction produced the same behaviour. Random perturbations to the velocity components were seeded at initialization, with a maximum amplitude $|\delta\mathbf{u}| \sim 10^{-5}$. Finally, we adopt a very small but finite thermal diffusivity of $\chi = 2 \times 10^{-6}$ to facilitate conduction of thermal energy through the vertical boundaries and to aid code stability. No explicit non-adiabatic heating, cooling, or thermal relaxation is included in the simulations described in this section: therefore, convection gradual dies away after non-linear saturation.

4.1.2 Time evolution of averaged quantities

In the left-hand panel of Fig. 6, we track the time evolution of the volume-averaged kinetic energy density associated with the vertical velocity for simulations run at resolutions of 64^3, 128^3, and 256^3, respectively.

Initially, all simulations exhibit small-amplitude oscillations due to internal gravity waves excited in the convectively stable region at initialization. After some three orbits, the linear phase of the convective instability begins in earnest, characterized by exponential

growth of the perturbation amplitudes. During this phase, internal energy is converted into the kinetic energy of the rising and sinking fluid motion that comprises the convective cells. As the resolution is increased, shorter scale modes are permitted to grow. Because they are the most vigorous, the growth rate (proportional to the slope of the kinetic energy density) is slightly larger at better resolution.

The linear phase ends some 10 orbits into the simulation, and the flow becomes especially disordered. The peak in the vertical kinetic energy density occurs at this point, but this peak decreases with resolution, something we discuss below. After non-linear saturation, the kinetic energy decreases gradually: the convective cells redistribute thermal energy and mass, thus shifting the thermal profile from a convectively unstable to a marginal state, and ultimately the convective motions die out. After about 38 orbits, the kinetic energy density has dropped to about one hundredth of its value at non-linear saturation in the lower resolution runs. The level of numerical diffusivity has an appreciable effect in damping activity after non-linear saturation, with the decrease in vertical kinetic energy successively smaller over the same period of time in the 64^3, 128^3, and 256^3 simulations, respectively.

The behaviour of the kinetic energy aligns relatively well with our expectations. The Reynolds stress, on the other hand, is more interesting. The evolution of the xy-component of the Reynolds stress is plotted in the right-hand panel of Fig. 6. The stress is small, but perhaps its most striking feature is its positivity over the entire duration of all three simulations.

Perhaps most surprising is the positivity of the stress during the *linear* phase of the instability. We might expect that the axisymmetric modes (which send angular momentum inwards) dominate this period of the evolution, as non-axisymmetric disturbances only have a finite window of growth (about an orbit) before they are sheared out and dissipated by the grid. But the positivity of the stress suggests that instead it is the shearing waves that are the dominant players in the linear phase. Visual inspection of the velocity fields confirms that the flow is significantly non-axisymmetric, and we find several examples of strong shearing waves 'shearing through' $k_x = 0$ during this phase. It would appear these waves transport angular momentum primarily outward as they evolve from leading to trailing, behaviour that is in fact consistent with Ryu & Goodman (1992), who find that inward transport only occurs at sufficiently long times when the waves are strongly trailing and hence effectively axisymmetric. (Even then the stress is extremely oscillatory.)

---

[2] A simulation run at a resolution 512^3 is included in the list of fiducial simulations in Table B1 (see Appendix B) but is not described in this section for brevity.
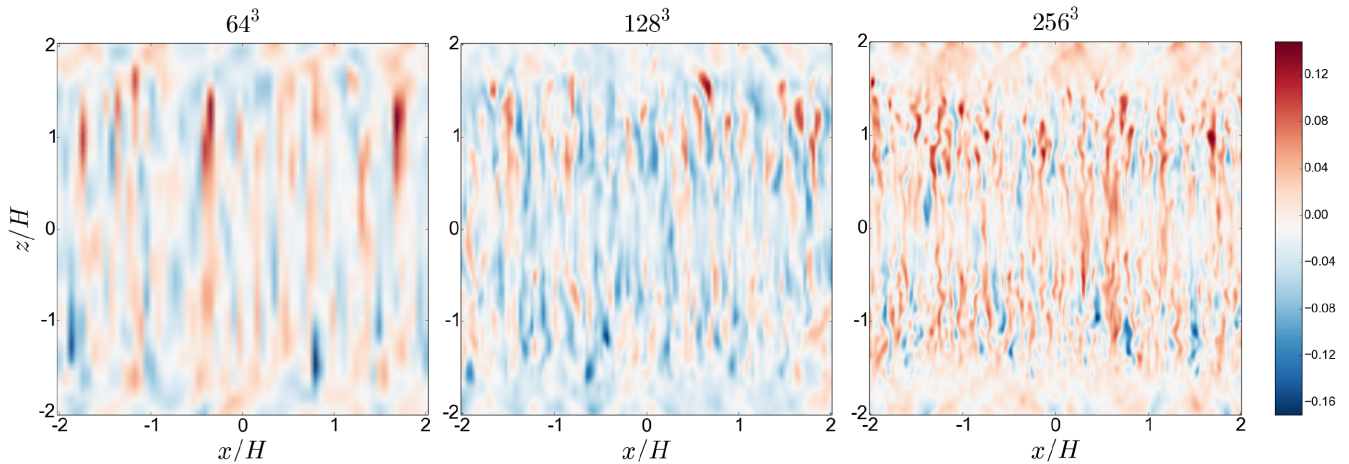
Figure 7. From left to right: snapshots of u_z in the xz-plane taken at non-linear saturation (just after the linear phase) at resolutions of 64^3, 128^3, and 256^3, respectively. Blue denotes $u_z < 0$ (sinking fluid) and red denotes $u_z > 0$ (rising fluid).

Why do the shearing waves dominate over the axisymmetric modes? One hypothesis is that 3D white noise seeds fast growing shearing waves that can outcompete the axisymmetric convective instability over their permitted window of growth. If this were to be true then the simulated non-linear regime is achieved before the dominant shearing waves become strongly trailing and begin to send angular momentum inward. Given that the typical time-scale for shearing out is roughly a few orbits at best, this is marginal but not impossible. An alternative, more plausible, and more troubling hypothesis is that our inviscid numerical code misrepresents trailing shearing waves: more specifically, the code is artificially reseeding fresh leading waves from strongly trailing waves ('aliasing'; Geoffroy Lesur, private communication). As a consequence, shearing waves outcompete the axisymmetric modes because they can shear through several times. If true, this is certainly concerning. But we hasten to add that this phenomenon should only be problematic in the low-amplitude linear phase; once the perturbations achieve large amplitudes, the aliasing will be subsumed under *physical* mode–mode interactions.

The linear phase ends in a spike in the stress. Time-averages of the volume-averaged value of R_{xy} from orbit 5 to orbit 15, a period spanning non-linear saturation, show that the Reynolds stress decreases as the resolution increases (see Table B1 in Appendix B). The time- and volume-averaged values of R_{xy} over a period spanning non-linear saturation (orbits 5–15) are $+4.8 \times 10^{-6}$, $+2.7 \times 10^{-6}$, and $+1.7 \times 10^{-6}$ in the 64^3, 128^3, and 256^3 simulations, respectively. In terms of the turbulent alpha-viscosity, this corresponds (from equation 14) to $\alpha \sim +2.3 \times 10^{-5}$, $+1.5 \times 10^{-5}$, and 9.6×10^{-6}, respectively. The dependence on the numerical viscosity can be explained by appealing to secondary shear instabilities (see the next subsection).

The volume-averaged density remains roughly constant during the linear phase of the instability, followed by a drop after non-linear saturation as mass is lost through the lower and upper boundaries. Although the convectively unstable region at initialization is confined to the region $|z| < L_c \sim 1.11H$, and is therefore well within the box, convective overshoot deposits mass (and heat) outside the convectively unstable region. The overall decrease in density over the duration of the simulation is small, but appears to increase with greater resolution. The overall percentage change in the density is -1.4, -3.4, and -8.8 per cent in the 64^3, 128^3 and 256^3, simulations, respectively.

4.1.3 Structure of the flow

The development of convective instability and the associated convective cells are best observed through snapshots of the vertical component of the velocity in the xz-plane, shown at resolutions of 64^3, 128^3, and 256^3 just after non-linear saturation in Fig. 7. A full set of convective cells is clearly visible at all resolutions. They are thin, filamentary structures several grid cells wide comprising alternating negative and positive velocities (updrafts and downdrafts).

The maximum vertical Mach number of the flow around non-linear saturation in the 256^3 run is about $M_z \sim 0.15$, with the largest vertical Mach numbers generally being measured near the vertical boundaries (where the temperature – and therefore the sound speed – is lowest).

The higher resolution simulations indicate that the plumes develop a wavy or buckling structure as they rise or sink, indicating the onset of a secondary shear instability. It is likely that the buckling of the convective plumes by these 'parasitic modes' limits the amplitudes of the linear modes, and ultimately leads to their breakdown. At lower resolution, numerical diffusion smooths out the secondary shear modes, and so the convective plumes reach large amplitudes before breaking down (blue curve in the right-hand panel of Fig. 6). At high resolution, the shear modes are not so impeded and make short work of these structures (black curve in the same figure).
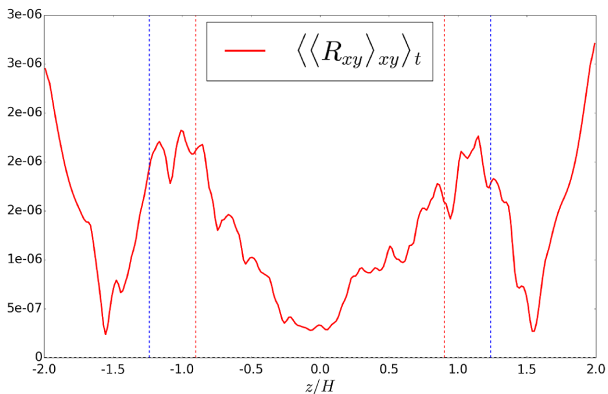
4.1.4 Vertical heat and mass flux

Fig. 8(a) shows the vertical profiles of horizontally averaged heat and mass fluxes taken from snapshots just after non-linear saturation from the simulation with resolution 256^3. For clarity, we have time-averaged the horizontally averaged vertical mass and heat flux profiles between orbits 5 and 10, a period spanning non-linear saturation (see Fig. 6).

Negative (positive) values for the fluxes for $z < 0$ and positive (negative) values for the fluxes for $z > 0$ correspond to transport of heat and mass away from (towards) the mid-plane. Overall, the heat flux is away from the mid-plane, peaking in the vicinity of the most convectively unstable points at initialization (indicated by the vertical dashed red lines in Fig. 8). In addition, there is some mass flux towards the mid-plane within $-H < z < H$. Thus, convection is transporting mass and heat such as to erase the convectively unstable stratification, as expected. Note that the positive peaks in the heat
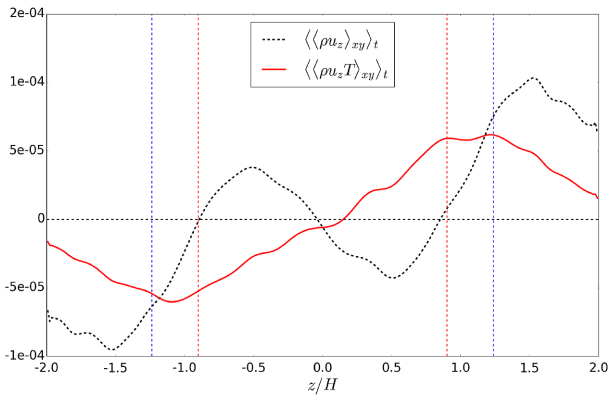
(a)

(b)

Figure 8. (a) Vertical profiles of the vertical heat (solid red line) and mass flux (dashed black line). The outer vertical dashed (blue) lines mark the boundaries of the convectively unstable region at initialization, while the inner vertical dashed (red) lines mark the most convectively unstable points at initialization. (b) Vertical profile of the xy-component of the Reynolds stress tensor R_{xy}. In all cases, the results have been time-averaged between orbits 5 and 10, spanning non-linear saturation.

flux near the most convectively unstable points are similar to those observed by Hirose et al. (2014) in the thermally dominated region of the disc (i.e. $P_{thermal} > P_{magnetic}$; cf. the middle-panel of Fig. 5 in their paper).

Beyond $|z| > H$, the mass flux is directed away from the mid-plane, peaking just beyond the point that marks the boundary of the convectively unstable region. This outward mass flux might be due to convective overshoot, although we expect this effect to vanish if averaged over a suitable time interval.

Finally, in Fig. 8(b) we show the vertical profile of the horizontally averaged R_{xy} taken from the 256^3 simulation and time-averaged between orbits 5 and 10. The Reynolds stress is clearly positive over all of the vertical domain, peaking just beyond the most convectively unstable points.[3] Thus, the bulk of outward angular momentum transport occurs where the convective cells begin to turn over, resulting in radial mixing of the gas. Note, however, that we also observe rather large positive stresses near the vertical boundaries, with the stress at the vertical boundaries about 1.5 times the peak stress in the remainder of the domain.

---

[3]Note however that at any given instant in time, we generally do observe regions where the Reynolds stress is negative.
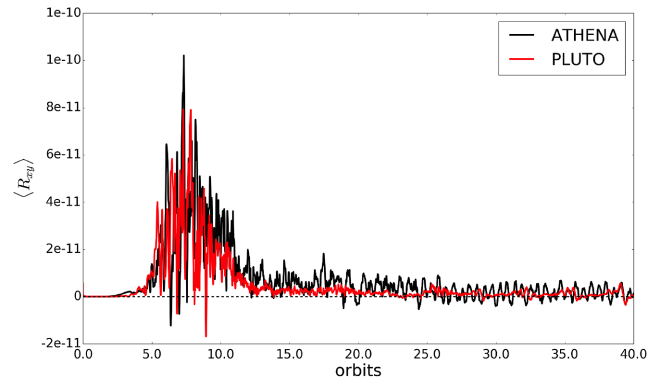

Figure 9. Comparison of time evolution of volume-averaged Reynolds stress taken from a simulation run in ATHENA (black) with the same quantity taken from a simulation run in PLUTO (red).

4.2 Simulation of compressible hydro convection in ATHENA

In the previous section, we found that hydrodynamic convection in a vertically stratified shearing box in PLUTO (without explicit viscosity) can drive outward angular momentum transport. Here, we verify this result using the finite-volume code ATHENA (Stone et al. 2008; Stone & Gardiner 2010). To facilitate as close a comparison as possible between the two codes and also to the ZEUS runs of SB96, we initialize both codes with Stone & Balbus's convectively unstable profile (see Appendix A1). Explicit diffusion coefficients were omitted and vertical periodic boundary conditions implemented. Both simulations were run in a shearing box of size $4H \times 4H \times 4H$ and at a resolution of $64 \times 64 \times 64$. PLUTO and ATHENA offer somewhat different suites of numerical schemes: here, we settle on a combination that is slightly more diffusive than that employed in our fiducial simulations because this combination allows for as close a match as possible. Specifically, the numerical scheme employed in ATHENA is (second-order) piecewise linear interpolation on primitive variables, the HLLC Riemann solver, and MUSCL-Hancock integration. In PLUTO, we use (second-order) piecewise linear interpolation on primitive variables together with a Van Leer limiter function, the HLLC Riemann solver, and MUSCL-Hancock integration. The angular frequency and the sound speed at initialization in both simulations were set to $\Omega = 10^{-3}$ and $c_s = 10^{-3}$, respectively.

We find that angular momentum transport is directed *outwards* in both codes, demonstrating that the outward transport of angular momentum by hydrodynamic convection in the non-linear phase is robust to a change of code. Fig. 9 compares the time evolution of the volume-averaged Reynolds stress taken from the ATHENA simulation with that taken from the PLUTO simulation. Both simulations exhibit exponential growth followed by non-linear saturation, together with the development of convective cells (not shown). These results also contrast with the SB96 runs with ZEUS and thus demonstrate that the positive transport reported in the previous sections is not special to the Gaussian temperature profile.

Although for this particular combination of schemes the overall result is the same, we noticed that different schemes resulted in qualitative differences between the two codes. For example, when we use the HLLC solver, piecewise parabolic reconstruction (PPM), and Corner-Transport-Upwind (CTU) integration, ATHENA exhibits delayed onset of instability, a more gradual drop in kinetic energy density following non-linear saturation, and a slower drop in angular momentum transport compared to PLUTO. When we use the Roe
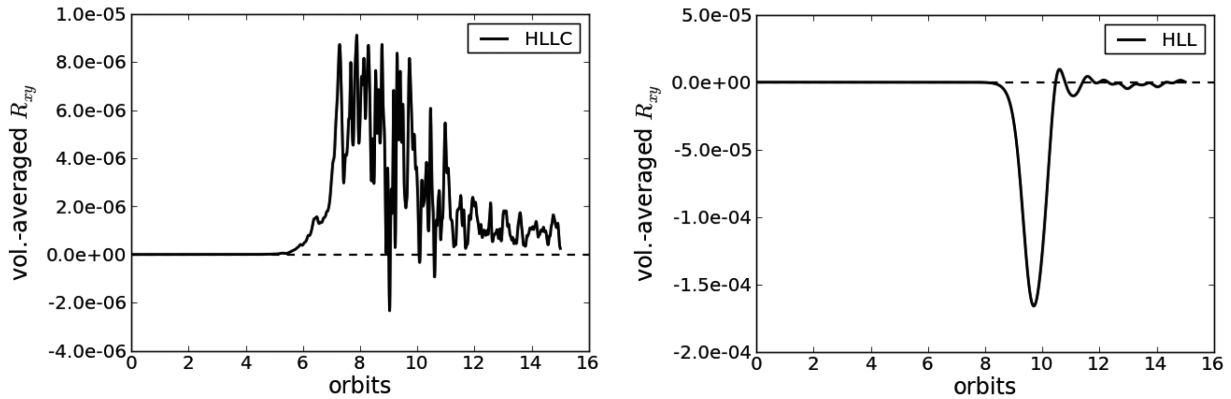
Figure 10. Left: time evolution of volume-averaged Reynolds stress tensor taken from a PLUTO simulation run with the less diffusive HLLC solver. Right: the same, but taken from a simulation run with the more diffusive HLL solver. The change in the sign of the stress tensor provides compelling evidence that whether convection can transport angular momentum inwards or outwards depends on the diffusivity of the underlying numerical scheme.

solver, PPM reconstruction, and CTU integration, the Reynolds stress is highly oscillatory in time in both codes, indicative, perhaps, of numerical instability. The different behaviour (in both codes) based on which combination of numerical schemes is chosen is worrying, although we emphasize that the differences are probably accentuated by the transient nature of the evolution and its sensitivity to the initial conditions, and the fact that the linear phase is partially controlled by grid diffusion. The sensitivity of our results to the numerical scheme is investigated in more detail in the following section.

4.3 Sensitivity of sign of angular momentum transport to numerical scheme

An important result of Sections 4.1–4.2 is that purely hydrodynamic convection in PLUTO and in ATHENA resulted in $R_{xy} > 0$, i.e. outward angular momentum transport. This is in disagreement with the ZEUS results of SB96, who reported a Reynolds stress of $R_{xy} < 0$. Given that ZEUS is a non-conservative, finite-difference code, our hypothesis for explaining the discrepancy is that ZEUS run at comparatively low resolution (as in SB96) is sufficiently diffusive that it imposes an artificial near-axisymmetry on the flow (which will send angular momentum inward). In this section, we report our attempts to assess the numerical diffusivity of various algorithms and their impact on convection.

First, we reran the fiducial simulations described in Section 4.1.1 but with different combinations of numerical schemes. We found that the sign of angular momentum transport is indeed sensitive to our choice of scheme. This is best illustrated in Fig. 10: the left-hand panel shows the time evolution of the volume-averaged R_{xy} taken from a simulation run with the HLLC Riemann solver, while the right-hand panel shows the same quantity but taken from a simulation run with the (more diffusive) HLL solver. Both simulations exhibit similar exponential growth in the vertical kinetic energy during the linear phase of the instability, and the velocity field shows the development of convective cells in both cases, but the sign of the Reynolds stress is radically different. The HLL run is also far more laminar and axisymmetric. We next repeated the HLL runs with higher resolutions, up to 256^3, but with no change in the sign of R_{xy}. It must be stressed that going to higher resolutions does not necessarily help in the problem of convection; this is because the fastest growing linear modes are always near the grid scale and

hence it is impossible (in the linear phase at least) to escape grid effects.

The results of different combinations of schemes are summarized in Table B2. They indicate that the sign of R_{xy} appears to be robust to changes in the interpolation and time-stepping schemes, but sensitive to the Riemann solver. In particular, the less diffusive Riemann solvers (Roe and HLLC) gave $R_{xy} > 0$, while the more diffusive Riemann solvers (HLL and a simple Lax–Friedrichs solver) gave $R_{xy} < 0$. Altogether we explored 12 different configurations of Riemann solver, interpolation scheme, and time-stepping method. These ranged from the least diffusive set-up, which was identical to that employed in the simulations described in the previous section (a Roe Riemann solver, third-order-in-space WENO interpolation, and third-order-in-time Runge–Kutta time-stepping), to the most diffusive set-up (a simple Lax–Friedrich's Riemann solver, second-order-in-space linear interpolation, and second-order-in-time Runge–Kutta time-stepping).

Thus, we have compiled evidence that, as suspected, angular momentum transport due to convection can be sensitive to the diffusivity of the numerical set-up. Oversmoothing of the flow by diffusive Riemann solvers such as HLL or by codes such as ZEUS, which employs artificial viscosity and finite-differencing of the pressure term, imposes a spurious axisymmetry on the flow. This enforced axisymmetry ensures that the flow transports angular momentum inwards.

4.4 Viscous simulations

Given the concerns raised in the last section regarding numerical schemes, as well as the fact that the fastest growing inviscid modes are on the shortest scales, we expand our study to include explicit viscosity (while retaining thermal diffusivity). A properly resolved viscous simulation should exhibit none of the numerical problems encountered above, and the fastest growing mode occurs on a well-defined scale above the (resolved) viscous scale. Our main aim in this section is to test whether the results of our fiducial simulations are solid: mainly, if angular momentum transport can be positive in the presence of viscosity. Additionally, the inclusion of explicit diffusion coefficients enables us to investigate the Rayleigh number dependence of fully compressible hydrodynamic convection in the shearing box, and thus connect to previous work by Lesur & Ogilvie (2010).

We carry out a suite of simulations at a resolution mainly of 256^3 investigating the effects on hydrodynamic convection when the Rayleigh number Ra is increased, but the Prandtl number Pr is fixed at 2.5.[4] Thus, viscosity and thermal diffusivity each have the same magnitude in any given simulation, and we decrease both in order to increase the Rayleigh number. The Rayleigh numbers of the simulations are Ra $= 10^5$, 10^6, 10^7, 10^8, 10^9, and 10^{10}. (Note that the simulations undertaken at the two highest Ra are probably underresolved, as explored later.) The Richardson number at initialization is fixed by the initial vertical profile, which is described in detail in Section 4.1.1 (and shown in Fig. 1). For the profile parameters chosen, the Richardson number at initialization is Ri $\sim$ 0.05. Further details of the simulations are given in Table B3 in Appendix B.

4.4.1 Rayleigh number dependence

As the Rayleigh number is increased from low to high values, the system proceeds through the same sequence of states found by Lesur & Ogilvie (2010). We observe no instability for Ra $= 10^5$. At Ra $= 10^6$ instability occurs but the flow appears relatively laminar and axisymmetric; in particular, the Reynolds stress is *negative* throughout the linear and non-linear phases. We conclude that the critical Rayleigh number for the onset of convection lies in the range $10^5 < \mathrm{Ra_c} < 10^6$. At Ra $= 10^7$ the instability is more vigorous and the flow field significantly more chaotic and non-axisymmetric in the non-linear phase, at which point the Reynolds stress has become positive. We conclude that the critical Ra at which the sign of R_{xy} flips lies between 10^6 and 10^7. At higher Ra the flow appears even more turbulent and non-axisymmetric. It is a relief that in the non-linear phase of the instability we find agreement with the inviscid simulations of previous sections at sufficiently high Ra.

Our critical Rayleigh number for the onset of convection is considerably higher than in Lesur & Ogilvie (2010), who report a value of $\mathrm{Ra_c} = 6900$, but this is perhaps explained by their much larger Richardson number ($=0.2$) that remains constant throughout their box. On the other hand, the critical Ra at which R_{xy} changes sign is much closer to ours, which suggests that the breakdown into non-axisymmetry may be controlled by secondary shear instabilities that are more sensitive to the viscosity than the convective driving.

While the transport of angular momentum is unambiguously outwards after non-linear saturation when Ra $\geq 10^7$, it is almost always inwards during the linear phase, as is illustrated by the red curve in Fig. 11. Moreover, during this early stage, the simulation is dominated by the axisymmetric modes of Section 3. This disagrees with our inviscid fiducial simulations (see the discussion in Section 4.1.2), which are non-axisymmetric from the outset. We confess that our viscous simulations are more in tune with our physical intuition of how the system should evolve: (a) in the linear phase growing axisymmetric modes outcompete shearing waves, (b) at sufficiently large amplitudes the modes are subject to secondary shear instabilities in the *xz*-plane that buckle the rising and falling plumes, (c) perhaps concurrently or on a short time later (and viscosity permitting), secondary *non-axisymmetric* instabilities also attack the plumes because they exhibit significant shear in the *xy*-plane as well, and (d) at this point, the flow degenerates into something more disordered, and importantly, *non-axisymmetric* and
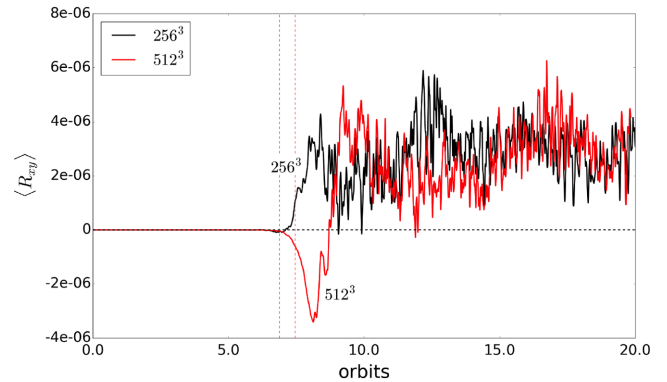

Figure 11. Time evolution of volume-averaged *xy*-component of Reynolds stress tensor from two simulations run at resolutions of 256^3 (black curve) and 512^3 (red curve) at a Rayleigh number of Ra $= 10^9$. The vertical dashed lines mark the end of the linear phase in the 256^3 (black) and 512^3 (red) simulations, respectively.

the Reynolds stress flips sign. We conclude that viscosity preferentially damps shearing waves vis-a-vis axisymmetric modes or effectively kills off the artificial aliasing of shearing waves in inviscid simulations (if this is present).

4.4.2 Convergence with resolution

A final issue is whether our viscous simulations are adequately resolved. More specifically: above what critical Ra is a grid of 256^3 points inadequate? We conducted simulations at Ra $= 10^8$ with 256^3 and 512^3 grid points and found generally good agreement between the two. The linear growth rates were almost identical and the ultimate non-linear state statistically similar. The only noticeable difference was in the peak Reynolds stress, which was somewhat larger in the higher resolution run. Overall, we conclude that 256^3 grid points are adequate to resolve a simulation with Ra $= 10^8$.

Things start to deteriorate at a Rayleigh number of Ra $= 10^9$. In Fig. 11, we plot the time evolution of the *xy*-component of the volume-averaged Reynolds stress for two simulations run at 256^3 (black) and 512^3 (red). The lower resolution run possesses no extended period of negative R_{xy}, in contrast to the runs at Ra $= 10^8$. We speculate that at this resolution physical viscosity is subordinate to the grid and non-axisymmetric disturbances are artificially enhanced, probably via aliasing. At 512^3, the stress is definitely negative in the linear phase and there is a strong negative peak shortly afterwards. The physical viscosity is now permitted to work properly and appears to prohibit artificial non-axisymmetric disturbances. As a consequence, the axisymmetric modes preserve their control of the simulation for significantly longer and the simulation is in accord with those at lower Ra. More reassuring is the non-linear phase a little later in which the two flows closely resemble each other. We conclude that in the linear phase 256^3 is insufficient to describe a simulation of Ra $= 10^9$, but in the later non-linear phase it is probably adequate.

5 STRUCTURES IN FORCED COMPRESSIBLE CONVECTION

In Section 4, we initialized our simulations with a convectively unstable vertical profile, but otherwise did not include any source of heating or cooling. Convection (and to a much lesser extent, conduction) transferred heat and mass vertically so as to zero the

---

[4]We have also repeated some of the simulations at a Prandtl number of unity. The differences with the Pr $= 2.5$ simulations are nominal.
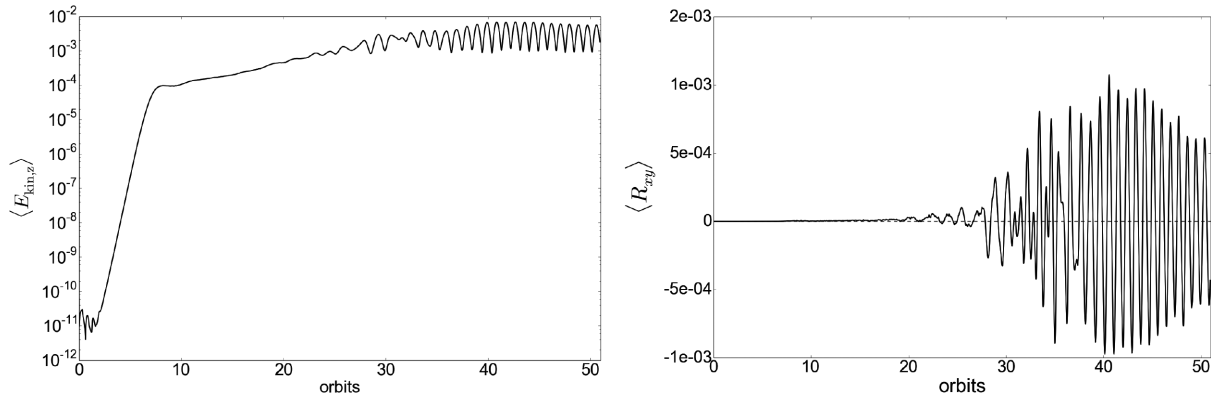
Figure 12. Left: semilog plot of time evolution of volume-averaged vertical kinetic energy density in a simulation with thermal relaxation. Right: time evolution of volume-averaged *xy*-component of the Reynolds stress tensor. The thermal relaxation time was set to $\tau_{\mathrm{relax}} = 4.2\,\Omega^{-1}$.

buoyancy frequency and send the box into a convectively stable equilibrium. As a result, we were only able to probe non-linear convection for a short period of time. Now we aim to continually sustain convection in order to explore this phase in greater depth. In the absence of self-sustaining convection, this means we have to maintain the convectively unstable profile artificially.

5.1 Set-up

SB96 perpetuated convection by forcing the temperature at the mid-plane to adjust to its value at initialization. This strategy, however, raises problems in conservative codes, such as PLUTO and ATHENA, because the energy injection at the mid-plane has no way to leave the box, except through the distant vertical boundaries (in ZEUS numerical losses on the grid supply a turbulence-dependent 'cooling', in addition to the cooling facilitated through the fixed-temperature boundaries). In practice, we found that the disc heats up to an inordinate level and, more importantly, it settles into a marginally unstable state, rather than a driven convective state.

Rather than forcing the code in this way, we mimic the effects of *both* heating and cooling through thermal relaxation. The idea is to add a source term to the energy equation such that the vertical internal energy profile relaxes to its value at equilibrium on a time-scale τ_{relax}. Although this technique is artificial, it serves as a very basic tool for approximating the effects of realistic heating and cooling, as might be supplied by MRI turbulence and radiative transfer. The time-scale τ_{relax} then would be the characteristic time that the radiative MRI system achieves thermodynamical quasi-equilibrium. In our simulations, however, we take the relaxation time-scale to be equal to the linear growth time of convection.

In addition, to mitigate the effects of mass outflows, we incorporate a mass source term to the simulations. At the end of the *n*th time-step we calculated the total vertical mass-loss through the upper and lower vertical boundaries. We then added this mass back into the box in the $(n + 1)$th time-step with the same vertical profile used to initialize the density (cf. equation A3).

We implement the thermal relaxation term by making slight modifications to PLUTO's built-in optically thin cooling module. The thermal energy equation is updated during a substep to take into account user-defined sources of heating and cooling. The resolution employed in the simulation described in this section is $256 \times 256 \times 256$ in a box of size $6H \times 6H \times 4H$, corresponding to about 43 grid cells per scale height in the *x*- and *y*-directions and 64 grid cells per scale height in the vertical direction. The numerical set-up,

boundary conditions and initial conditions are the same as those describe in Section 4.1.1. The thermal relaxation time is taken to be equal to the inverse of the growth rate of the convective instability which we measured to be $\sigma = 0.2404\,\Omega$. Thus, the thermal relaxation time is $\tau_{\mathrm{relax}} = 4.2\,\Omega^{-1}$, i.e. the internal energy is relaxed back to its equilibrium profile on about 0.7 orbits. We have also included an explicit viscosity and thermal diffusivity of $\nu = 1.075 \times 10^{-5}$ corresponding to a Rayleigh number of $\mathrm{Ra} = 10^9$ and a Prandtl number of $\mathrm{Pr} = 2.5$.

Finally, we have repeated the simulations in a cube of resolution 64^3 and size $4H \times 4H \times 4H$, as well as at a resolution of $128 \times 128 \times 64$ in a box of size $6H \times 6H \times 4H$, and also with periodic boundary conditions in the vertical direction, and observed similar results. We have also run a simulation with the HLL solver to partially explore any code dependence.

5.2 Large-scale oscillatory cells

In the left-hand panel of Fig. 12 we show the time evolution of the volume-averaged kinetic energy density associated with the vertical velocity. As in Section 4, exponential growth in the linear phase is followed by non-linear saturation. The forced simulations, however, do not subsequently decay. Instead the vertical kinetic energy increases at a slower rate until about orbit 20, at which point oscillations in the kinetic energy density begin to develop. The cycles increase in frequency and amplitude until about orbit 41 at which point the system settles into a quasi-equilibrium, with the volume-averaged vertical kinetic energy density oscillating about $\langle E_{\mathrm{kin},\,z}\rangle \sim 2 \times 10^{-3}$ and the period remaining steady at $\Delta t = 0.88$ orbits.

Associated with the oscillations are large fluctuations in the *xy*-component of the Reynolds stress tensor (shown in the right-hand panel of Fig. 12). Instantaneous fluctuations in $\langle R_{xy}\rangle$ and in $\langle \alpha \rangle$ may be either positive or negative, but the time-averaged values over this cyclical phase (orbit 20 to the end of the simulation) are $R_{xy} \sim +9.9 \times 10^{-6}$ and $\alpha \sim +3.9 \times 10^{-5}$, respectively. Furthermore, comparing the oscillations in the vertical kinetic energy density to the fluctuations in $\langle R_{xy}\rangle$ and also to the volume-averaged gas pressure $\langle P \rangle$ (not shown here) it is apparent that *peaks* in the kinetic energy density are correlated with both $\langle R_{xy}\rangle < 0$ and troughs in $\langle P \rangle$, while *troughs* in the kinetic energy density are correlated with $\langle R_{xy}\rangle > 0$ and peaks in $\langle P \rangle$.

A clearer picture of the behaviour of the system emerges when we study the structure of the flow during the cyclical phase. Fig. 13
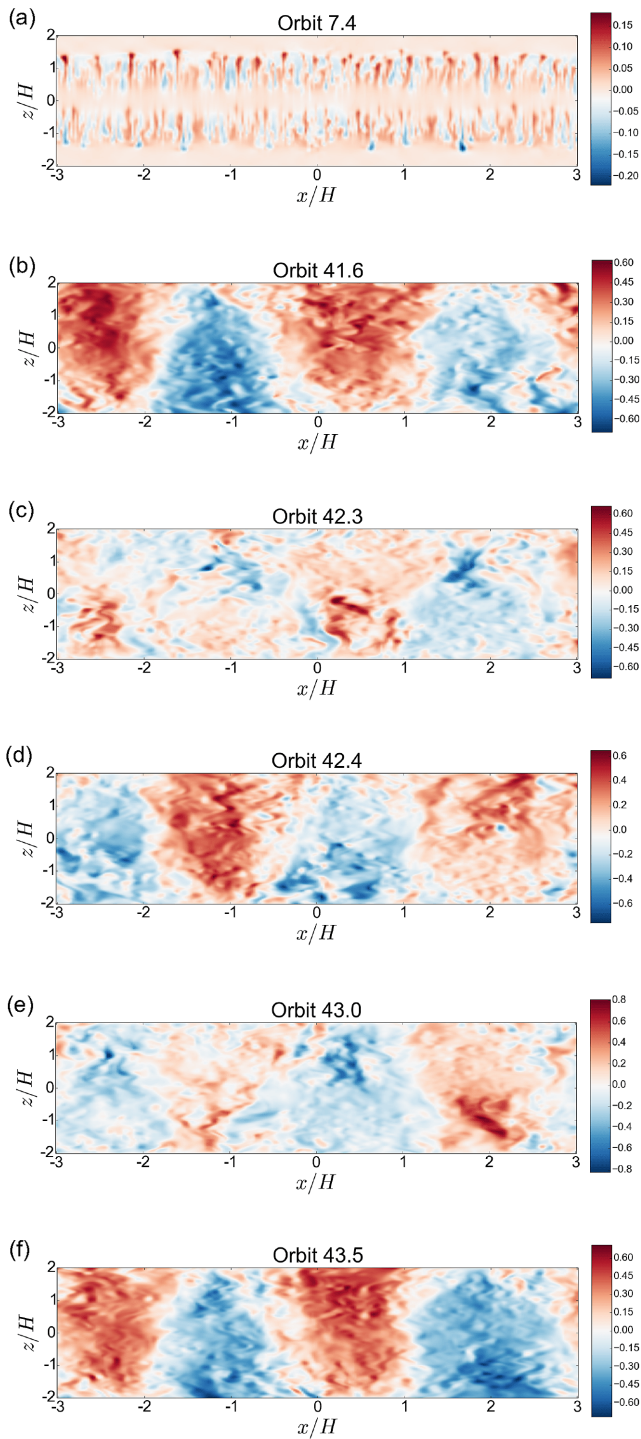
Figure 13. Snapshots at different times of the z-component of the velocity in the xz-plane taken from a simulation in which convection was sustained using thermal relaxation. Narrow convective cells shown in (a) just after non-linear saturation merge to form large scale structures shown in (b), which are destroyed (c) and recreated (d) in a cyclical manner with the opposite rotation.

shows snapshots of the z-component of the velocity in the xz-plane. The first panel is taken just after the end of the linear phase (orbit 7.4); the subsequent panels are taken at five successive peaks and

troughs in the kinetic energy. As the kinetic energy rises in the non-linear phase, the thin convective cells with radial wavelengths $\lambda_x \sim 5\,\Delta x \sim 0.177H$, where Δx is the size of grid a cell in the x-direction, slowly begin to merge into larger coherent structures which couple the two halves of the disc together. By orbit 20 (start of the cyclical phase), the radial wavelength of the convective cells has increased to $\lambda_x \sim H$, and by the time the quasi-steady equilibrium state sets in (at around orbit 40) the radial wavelength of the convective cells is $\lambda_x \sim 3.4H$.

Comparing the snapshots of u_z to the peaks and troughs in the kinetic energy, it becomes apparent that peaks in the convective energy are associated with the large-scale axisymmetric convective cells, hence $\alpha < 0$. Troughs in the kinetic energy are associated with destruction of those convective cells and with positive stress (outward angular momentum transport). Thus, we are observing large-scale convective eddies that appear to be created and destroyed cyclically. Furthermore, it is evident from Fig. 13 that after the eddies are destroyed, they are re-formed with the opposite rotation.

The reader may be alarmed that the large eddies extend all the way to the vertical boundaries of the domain, and indeed there is an uncomfortable level of mass-loss during this phase. To check that the flows are not an artefact of our box size, we ran additional simulations with a vertical domain of $\pm 3H$. A density floor had to be activated in such runs, which unfortunately compromised our thermal equilibrium and complicated the onset of convection early in the simulation. However, they ultimately settled on to the cyclical state described above, but now the large-scale eddies peter out before reaching the vertical boundaries. As a result, the mass-loss drops to negligible amounts. This confirms that these flows are physical and robust, though only marginally contained within our smaller boxes.

Next, to explore any code dependence of this final outcome, we have rerun the simulation with the more diffusive HLL solver. We observe a negative Reynolds stress during the linear phase and well into the non-linear phase. Associated with this is a remarkably axisymmetric flow field, as confirmed by viewing slices of the pressure perturbation δP in the xy-plane at different times. The simulation, none the less, enters the cyclical phase during which this axisymmetry is broken. As expected there is a flip in the sign of the Reynolds stress from negative to positive. The behaviour thereafter mirrors that of the simulation of forced compressible convection run with the Roe solver: the cyclical creation and destruction of large-scale convective cells and an oscillatory Reynolds stress. We conclude that for forced runs, the ultimate quasi-steady state depends negligibly on the numerical algorithm.

5.3 Zonal Flows

In Fig. 14, we plot, side-by-side, space–time diagrams in the xt-plane of the yz-averaged pressure perturbation $\delta P(x, t) \equiv (\langle P \rangle_{yz} - \langle P \rangle)/\langle P \rangle$ and of the yz-averaged perturbation to the y-component of the velocity $\langle \delta u_y \rangle_{yz}(x, t) \equiv \langle u_y - 1.5\,\Omega\,x \rangle_{yz}$. It is immediately evident from Fig. 14 that around the onset of the cyclical phase (orbit 20) alternating streaks in δP and δu_y begin to develop. Respectively, these mark alternating bands (in x) of high ($\delta P > 0$) and low ($\delta P < 0$) pressure, and of super-Keplerian ($\delta u_y > 0$) and sub-Keplerian ($\delta u_y < 0$) flow.

From Fig. 14, it is clear that the pressure and velocity perturbations are 90 degrees out of phase, which is characteristic of *zonal flows*. We hesitate, however, to claim that these structures are in geostrophic balance (i.e. that $\partial_x P' \sim \rho_0 \,\Omega\,\delta u_y$) because, as is evident from the space–time diagrams, the flows are not stationary in time but appear to fluctuate over about one orbit – the same
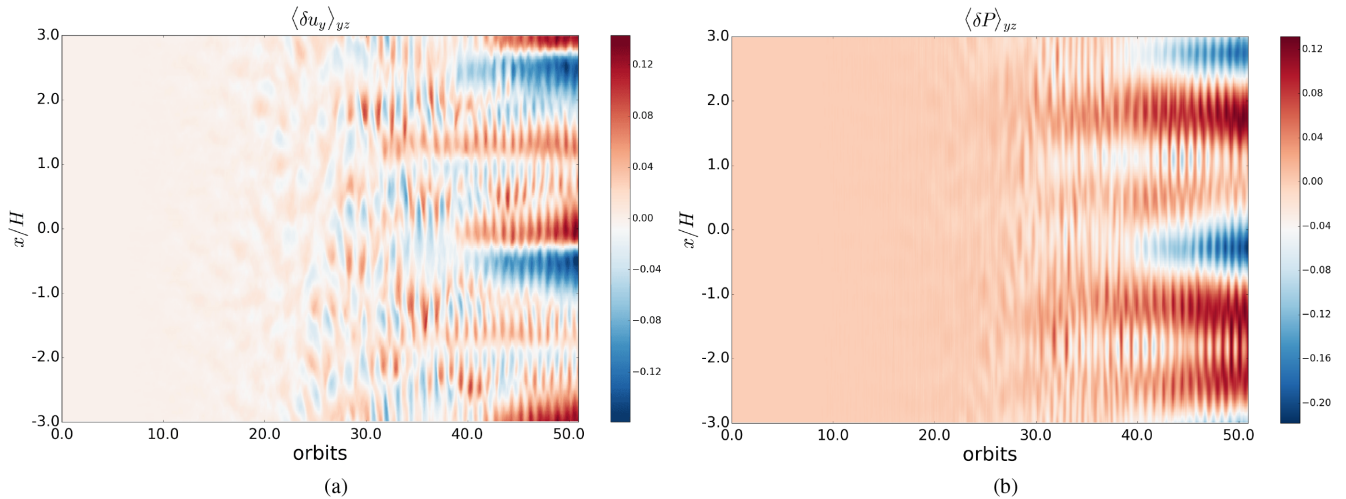
Figure 14. Space–time diagrams of *yz*-averaged *y*-component of the perturbed velocity δu_y (left) and of the pressure perturbation δP (right). (The perturbed pressure is defined in the first paragraph of Section 5.3).

time-scale over which the large convective cells are created and destroyed. Indeed, the creation and destruction of the zonal flows track precisely that of the large-scale convective cells, showing clearly that the two phenomena are connected. The axisymmetry observed during the formation of large-scale convective cells and zonal flows is consistent with the inward transport of angular momentum observed while these structures remain coherent: axisymmetric convective modes dominate during the lifetime of these structures and transport angular momentum inwards.

In Fig. 15 (right-hand panels), we plot snapshots in the *xy*-plane taken during the cyclical phase of the perturbation of the *y*-component of the velocity δu_y and of the perturbed pressure δP. The axisymmetric structure of the zonal flows is clearly visible in panels (b) and (d), as is the $\pi/2$ phase difference between the pressure perturbation δP and the perturbed *y*-component of the velocity δu_y.

5.4 Vortices

Because the zonal flows consist of alternating bands of sub- and super-Keplerian motion – with strong shear between the bands – we expect that they could give rise to vortices via the Kelvin–Helmholtz instability (modified by rotation and stratification). In Fig. 16, we plot a snapshot in the *xy*-plane taken during the cyclical phase of the the *z*-component of the vertical component of vorticity ω_z. Small but coherent anticyclonic blobs of vorticity are observed during the cyclical phase, and these occur precisely where the flow transitions from sub-Keplerian to super-Keplerian (i.e. at the edges of the zonal flows). Typically, only anticyclonic vortices are observed, which is consistent with the fact that cyclonic vortices tend to be more unstable in Keplerian shear flows.

The vortices tend to be elongated at the start of the cyclical phase (around orbit 20), with an aspect ratio of about 0.4, and grow increasingly circular with time; in fact, once quasi-steady equilibrium has been reached at around orbit 41, many of the vortices have aspect ratios approaching unity. They appear to have limited three-dimensional extent, and we did not observe any vortices that remained coherent for depths exceeding half a scale height.

5.5 Discussion

It is tempting to link our results to the large-scale emergent structures in recent simulations of rotating hydrodynamic convection with uniform (rather than Keplerian) rotation and in the Boussinesq (rather than compressible) regime (Julien et al. 2012, Rubio et al. 2014, Favier, Silvers & Proctor 2014; Guervilly, Hughes & Jones 2014). Both Favier et al. (2014) and Guervilly et al. (2014) observe growth in the vertical component of the kinetic energy at high Rayleigh numbers (Ra $\sim 10^7$–10^9), as we do. In addition, Favier et al. (2014) and Guervilly et al. (2014) notice the formation of depth-invariant large-scale vortices in the *xy*-plane of their simulations (i.e. in the plane perpendicular to the rotation axis).

There are both qualitative and quantitative differences between our results and those derived from uniformly rotating, Boussinesq convection. Although the Rayleigh number of our simulation with thermal relaxation was Ra $= 10^9$, which is consistent with the Rayleigh numbers at which large-scale vortices were observed in the Boussinesq simulations, we observe large-scale *convective cells* (in the *xz*-plane) rather than large-scale vortices (in the *xy*-plane). We do observe anticyclonic vortices in the *xy*-plane, but these are small and do not appear to have any three-dimensional extent compared to the depth invariant vortices observed in the Boussinesq simulations with uniform rotation. It is possible that they are prevented from merging and thus growing in size because of the turbulence associated with repeated destruction of the large-scale convective cells. The strong shear might also inhibit their growth. Alternatively, our large-scale axisymmetric cells could be interpreted as vortical structures sheared out by the disk's differential rotation.

Due to the cyclical nature of the large-scale convective cells, it is tempting to link our results to the intermittent convection reported in the MRI shearing box simulations of Hirose et al. (2014) and Coleman et al. (2018). In our runs, the forcing is due to explicit thermal relaxation, whereas in the runs of Hirose et al. (2014) the forcing is self-consistently provided by MRI heating and radiative cooling. Thus, the forcing and in particular the time-scales associated with the forcing are rather different. However, both thermal relaxation and the MRI limit cycles are similar in that they lead to a cyclical build-up of heat and its subsequent purge through vertical convective transport, with the role of opacity in the simulations of Hirose et al. (2014) being simply to modulate this cycle. Thus,
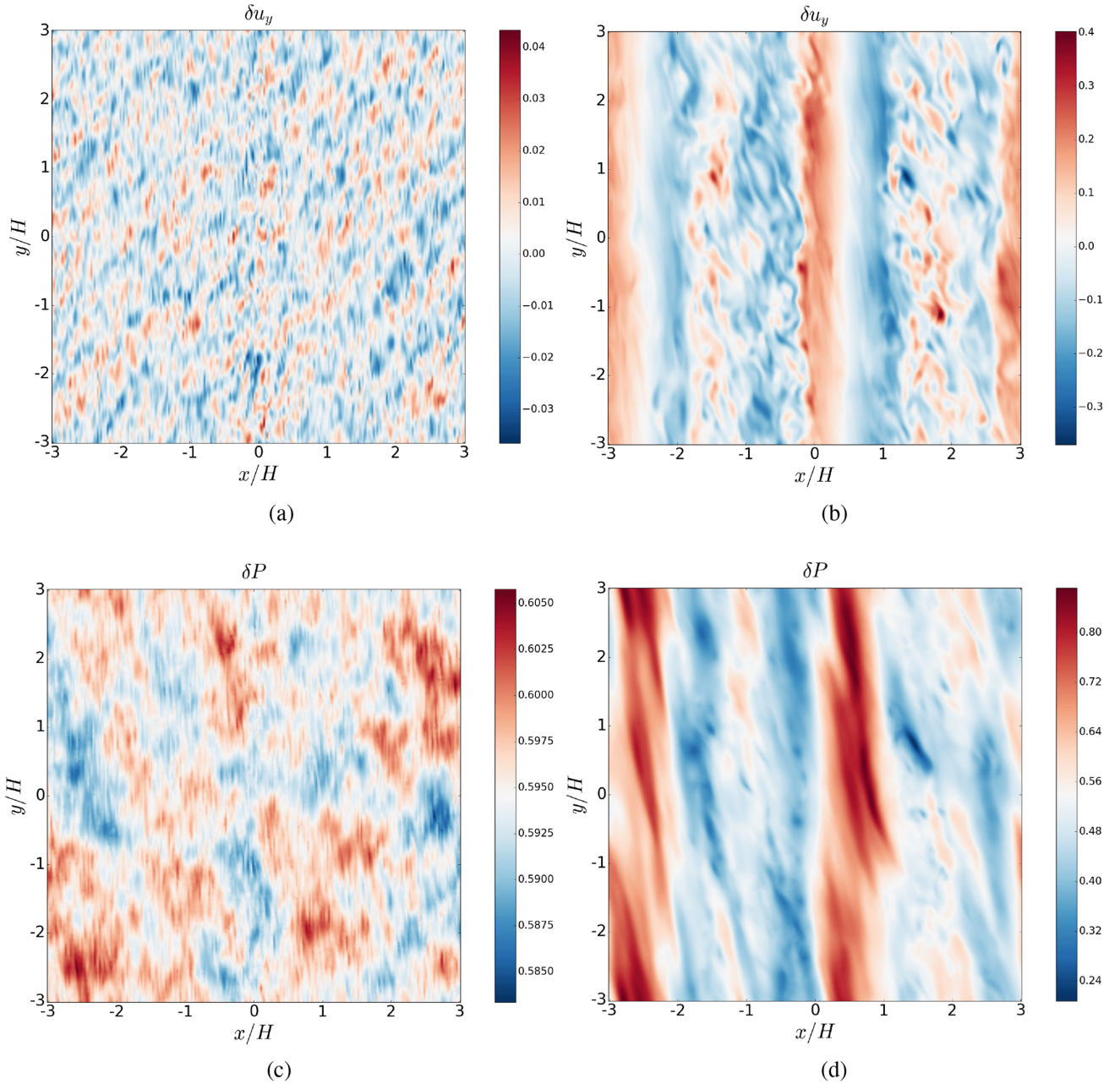
Figure 15. Two-dimensional slices in the *xy*-plane of the perturbation to the *y*-component of the velocity δu_y, and of the fractional pressure perturbation δP from a simulation of forced compressible convection at $z \approx 0.5H$. Left column: slices taken from a snapshot just after non-linear saturation (orbit 7.4). Right column: slices taken from a snapshot generated during the cyclical phase (orbit 41.6).

our results demonstrate that strongly non-linear hydrodynamic turbulent convection has a cyclical nature that might be generic, and that might therefore be robust to the inclusion of more realistic thermodynamics.

6 CONCLUSIONS

Motivated by recent radiation magnetohydrodynamic shearing box simulations that indicate that an interaction between convection and the magnetorotational instability in dwarf novae can enhance angular momentum transport, we have studied the simpler case of purely hydrodynamic convection, both analytically and through three-dimensional, fully compressible simulations in PLUTO.

For the linear phase of the instability, we find agreement between the growth rates of axisymmetric modes calculated theoretically and those measured in the simulations to within a percentage error of < 1 per cent, thus providing a useful check on our PLUTO code. The linear eigenmodes are worth examining not only to help understand the physical nature of convection in discs, but also because they may appear in some form during the non-linear phase of the evolution, especially on large scales, and during intermittent or cyclical convection.

We then explored the non-linear saturation of the instability, both when convection is continually forced and when it is allowed to reshape the background gradients so that it ultimately dies out. We focussed especially on the old problem of whether hydrodynamic
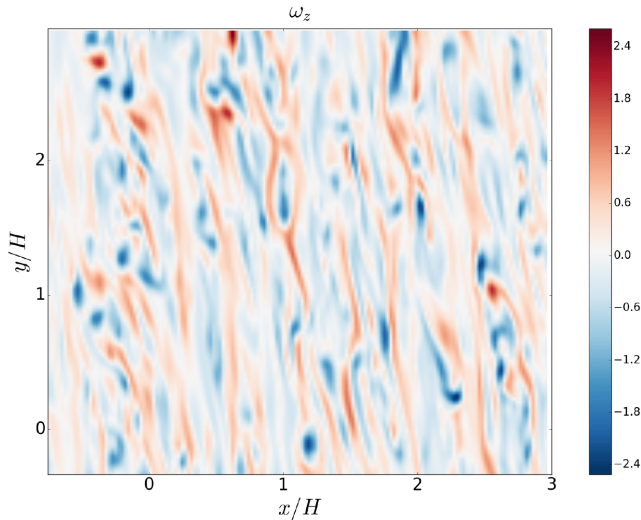
Figure 16. Two-dimensional slice in the *xy*-plane of the *z*-component of the vorticity ω_z taken from a snapshot generated at the start of the cyclical phase (orbit 20) and at $z \approx 0.5H$. For clarity, we have zoomed in on the upper-right quadrant of the *xy*-plane.

convection in a disc leads to inward ($\alpha < 0$) or outward ($\alpha > 0$) angular momentum transport. In both forced and unforced convection, we found $\alpha > 0$ in general in the non-linear phase. These results were confirmed by a separate run using the code ATHENA, but contradict the classical simulations of SB96, who reported inward transport in both cases (forced and unforced) using the code ZEUS.

This discrepancy reveals a set of unfortunate numerical difficulties that complicate the simulation of convection in discs. These, in large part, issue from the fact that the inviscid linear modes of convection grow fastest on the shortest possible scales. Thus, no matter the resolution, the nature of the code's grid dissipation will always impact on the system's evolution, certainly in the linear phase and possibly afterwards. We argue that a more diffusive numerical set-up, such as supplied by ZEUS at low resolution or Riemann solvers such as HLL, imposes an axisymmetry on the flow which leads to generally inward transport of angular momentum. But, on the other hand, we suspect that less diffusive solvers such as Roe and HLLC artificially alias shearing waves in the linear phase of the evolution, leading to spurious non-axisymmetric flow early in a simulation. Though concerning, we believe this is only a problem in the low-amplitude linear phase because physical mode–mode interactions will dominate once the perturbations achieve larger amplitudes. Nonetheless, shearing wave aliasing certainly deserves a separate study.

To properly dispense with these numerical issues, one must add explicit viscosity (and thermal diffusivity), as this regularizes the linear problem. We find that at a Richardson number of Ri ~ 0.05, onset of convection is observed for a critical Rayleigh number $10^5 <$ Ra$_c \leq 10^6$. Just above this value, convection is largely axisymmetric and $\alpha < 0$. At a larger second critical Ra between 10^6 and 10^7, the sign of α switches and the flow becomes more turbulent and non-axisymmetric. This sequence of states mirrors that simulated by Lesur & Ogilvie (2010). At large (resolved) Rayleigh numbers, viscous simulations are initially controlled by the axisymmetric modes; these are then attacked by secondary shear instabilities in both the *xz*- and *xy*-planes, which break the axisymmetry and order of these structures, leading to a more chaotic state. At lower Ra,

viscosity suppresses the non-axisymmetric shear instabilities and axisymmetry is never broken. (At even lower Ra, convection never begins, of course.)

In forced convective runs, rather than maintaining the convection by a fixed heating source at the mid-plane, we instead allowed the vertical equilibrium to relax to its initial, convectively unstable, state. Our thermal relaxation is artificially imposed, but its overall effect is to mimic the heating of the mid-plane and cooling of the corona due to physical mechanisms that maintain the convectively unstable entropy profile, such as the MRI and radiative losses present in the simulations of Hirose et al. (2014). We observed in the non-linear stage now the formation of large-scale convective cells (similar in some respects to elevator flow) that emerge and break down cyclically, in addition to zonal flows and vortices. Although further checks are required, it is tempting to link this cyclical convection to the convective limit cycle observed in the radiative magnetohydrodynamic simulations of Hirose et al. (2014), since both processes rely on the build-up and rapid evacuation of heat.

Despite our demonstration that hydrodynamic convection can lead to positive stress and outward transport of angular momentum, the fact remains that the stresses are small (typically, we measured $\alpha \sim 10^{-6}$–10^{-5}). Having said that, the magnitude of α is sensitive to the depth of the buoyancy frequency profile, and a deeper profile could increase α by an order of magnitude or more.

Finally, we have not observed self-sustaining hydrodynamic convection in any of our simulations. By self-sustaining convection we mean that (when $\alpha > 0$) energy extracted from the shear by convection might itself cause convective motions, which in turn extract energy from the shear, closing the loop. It is more likely that if convection is to occur in discs it will be as a by-product of other processes, such as heating by density waves, emitted in the presence of a planet, or by dissipation of magnetorotational turbulence. We intend to investigate both mechanisms and their instigation of convection in future work.

REFERENCES

Alexiades V., Amiez G., Gremaud P.-A., 1996, Commun. Numer. Methods Eng., 12, 31
Armitage P. J., 2011, ARA&A, 49, 195
Aurnou J., Calkins M., Cheng J., Julien K., King E., Nieves D., Soderlund K., Stellmach S., 2015, Phys. Earth Planet. Inter., 246, 52
Bell K. R., Lin D. N. C., 1994, ApJ, 427, 987
Bodo G., Cattaneo F., Mignone A., Rossi P., 2013, ApJ, 771, L23
Boley A. C., Durisen R. H., 2006, ApJ, 641, 534
Boyd J. P., 2001, Chebyshev and Fourier Spectral Methods. Dover, New York USA
Cabot W., 1996, ApJ, 465, 874
Cannizzo J. K., 1993, in Wheeler J. C., ed., The Limit Cycle Instability In Dwarf Nova Accretion Disks. World Scientific, Singapore, p. 6
Chiang E., Goldreich P., 1997, ApJ, 490, 368
Coleman M. S., Blaes O., Hirose S., Hauschildt P. H., 2018, ApJ, 857, 52
D'Alessio P., Cantö J., Calvet N., Lizano S., 1998, ApJ, 500, 411
Favier B., Silvers L. J., Proctor M. R. E., 2014, Physics of Fluids, 26, 096605
Goldreich P., Lynden-Bell D., 1965, MNRAS, 130, 125

Guervilly C., Hughes D. W., Jones C. A., 2014, J. Fluid Mech., 758, 407
Hawley J. F., Gammie C. F., Balbus S. A., 1995, ApJ, 440, 742
Hirose S., 2015, MNRAS, 448, 3105
Hirose S., Blaes O., Krolik J. H., Coleman M. S., Sano T., 2014, ApJ, 787, 1
Julien K., Knobloch E., Rubio A. M., Vasil G. M., 2012, PRL, 109, 254503
Klahr H., Henning T., Kley W., 1999, ApJ, 514, 325
Kley W., Papaloizou J., Lin D., 1993, ApJ, 416, 679
Latter H. N., Papaloizou J., 2017, MNRAS, 472, 1432
Lesur G., Kunz M. W., Fromang S., 2014, A&A, 56
Lesur G., Ogilvie G. I., 2010, MNRAS, 404, L64
Lin D., Papaloizou J., 1980, MNRAS, 191, 37
Lin D., Papaloizou J., Kley W., 1993, ApJ, 416, 689
Lyra W., Richert A. J., Boley A., Turner N., Mac Low M.-M., Okuzumi S., Flock M., 2016, ApJ, 817, 102
Mignone A., Bodo G., Massaglia S., Matsakos T., Tesileanu O., Zanni C., Ferrari A., 2007, ApJS, 170, 228
Mignone A., Flock M., Stute M., Kolb S., Muscianisi G., 2012, A&A, 545, A152
Rubio A. M., Julien K., Knobloch E., Weiss J. B., 2014, PRL, 112, 144501
Ruden S. P., Papaloizou J., Lin D., 1988, ApJ, 329, 739
Ryu D., Goodman J., 1992, ApJ, 388, 438
Spruit H. C., Nordlund A., Title A., 1990, ARA&A, 28, 263
Stone J. M., Balbus S. A., 1996, ApJ, 464, 364
Stone J. M., Gardiner T. A., 2010, ApJS, 189, 142
Stone J. M., Gardiner T. A., Teuben P., Hawley J. F., Simon J. B., 2008, ApJS, 178, 137
Zingale M. et al., 2002, ApJS, 143, 539

APPENDIX A: CONVECTIVELY UNSTABLE VERTICAL DISC PROFILES

We describe the two convectively unstable vertical profiles that are used to initialize our simulations. The reader should note that these profiles may or may not correspond to those in *real* astrophysical discs, which will be determined by several sources of heating and cooling, none considered in this paper. Here, we simply present convectively unstable disc profiles that satisfy hydrostatic equilibrium, are convectively unstable, and can conveniently initialize simulations.

A1 Stone and Balbus (1996) profile

The SB96 profile employs a power-law profile in temperature, see equation (9). For $p = 3/2$, the density can be obtained analytically

$$\rho = \rho_0 (1 - s^3)^{-(1+g/3)} (1-s)^g \exp\left\{ 2g \left[s - \frac{1}{\sqrt{3}} \tan^{-1} \left(\frac{\sqrt{3}s}{s+2} \right) \right] \right\},$$

where ρ_0 is the mid-plane density, $s = (z/H)^{1/2} f^{1/3}$, and $g = f^{-4/3}$, in which $f = H^{3/2} A/T_0$, and $H = c_s/\Omega$ is the mid-plane scale height. The

pressure is obtained from the ideal gas equation of state. Note that for $|z| > (T_0/A)^{1/p}$ we must have vacuum, which ties the maximum numerical domain to the ratio T_0/A.

The buoyancy frequency for $p = 3/2$ is given by

$$\frac{N_B^2}{\Omega^2} = \frac{1}{1 - \left(\frac{z}{H}\right)^{3/2} f} \left[\left(1 - \frac{1}{\gamma} \right) \left(\frac{z}{H} \right)^2 - \frac{3}{2} \left(\frac{z}{H} \right)^{3/2} f \right], \quad \text{(A1)}$$

which corresponds to a profile that is negative (convectively unstable) within some region $|L_c| < 0$ about the mid-plane and positive (convectively unstable) outside of this region. The width of the convectively unstable region is given by

$$|L_c| = \left[\frac{3}{2} \left(1 - \frac{1}{\gamma} \right)^{-1} H^2 \frac{A}{T_0} \right]^2. \quad \text{(A2)}$$

Although convenient within a limited choice of parameters, the SB96 profile suffers from the drawback that the width of the convectively unstable region is sensitive to the size of the box through the ratio T_0/A. Increasing the vertical box size necessarily decreases the size of the convectively unstable region.

A2 Gaussian temperature profile

The drawbacks of the SB96 profile motivated us to search for a more convenient unstable profile, one that would leave the size and depth of the convectively unstable region independent of the vertical extent of the box. Setting the temperature to a Gaussian, cf. equation (8), provided such a profile.

The associated density is

$$\rho = \rho_0 e^{z^2/\beta H^2} \exp\left(-\frac{\beta H^2 \Omega^2}{2T_0} e^{z^2/\beta H^2} \right), \quad \text{(A3)}$$

where ρ_0 is mid-plane density and H is the mid-plane scale height. Pressure is obtained from the ideal gas equation of state, as above.

The buoyancy frequency is given by

$$\frac{N_B^2}{\Omega^2} = \frac{2}{\beta H^2} z^2 \left[\left(\left(1 - \frac{1}{\gamma} \right) \frac{\beta H^2 \Omega^2 \mu}{2T_0 \mathcal{R}} e^{z^2/\beta H^2} - 1 \right) \right]. \quad \text{(A4)}$$

The boundary of the convectively unstable region is hence given by

$$|L_c| \equiv \sqrt{\beta H^2 \ln \left[\left(1 - \frac{1}{\gamma} \right)^{-1} \frac{2T_0 \mathcal{R}}{\beta \mu H^2 \Omega^2} \right]}. \quad \text{(A5)}$$

APPENDIX B: TABLES OF SIMULATIONS

Table B1. Simulations of unforced compressible hydrodynamic convection in the shearing box carried out in PLUTO. No explicit viscosity or heating/cooling/thermal relaxation was included. The box size in all simulations is $4H \times 4H \times 4H$. The simulations were initialized either with a Gaussian temperature convectively unstable vertical disc profile (see Appendix A2) or with the convectively unstable vertical disc profile used in Stone & Balbus (1996) (SB96; see Appendix A1). $\langle\langle . \rangle\rangle_t$ denotes a time- and volume-averaged quantity. $E_{\mathrm{kin},\,z}$ is the vertical kinetic energy density, which we have measured both at non-linear saturation (where it is maximum) and at the end of the simulation. $\Delta\rho$ is the percentage change in density over the course of the simulation, R_{xy} is the Reynolds stress, α is the alpha viscosity parameter, and σ is the linear phase growth rate given in units of Ω. All simulations were run for 38 orbits.

Run	Resolution	Profile	$\max(\langle E_{\mathrm{kin},\,z}\rangle)$	$\langle E_{\mathrm{kin},\,z}\rangle\vert_{\mathrm{end}}$	$\Delta(\langle\rho\rangle)$ (%)	$\langle\langle R_{xy}\rangle\rangle_t$	$\langle\langle\alpha\rangle\rangle_t$	$\sigma(\Omega)$
NSTR22c01	64^3	Gaussian Temp.	1.7×10^{-4}	2.0×10^{-6}	-1.4	$+4.8 \times 10^{-6}$	$+2.3 \times 10^{-5}$	0.2404
NSTR22c02	128^3	Gaussian Temp.	1.1×10^{-4}	3.7×10^{-6}	-3.4	$+2.7 \times 10^{-6}$	$+1.5 \times 10^{-5}$	0.2707
NSTR22c03	256^3	Gaussian Temp.	7.3×10^{-5}	9.8×10^{-6}	-8.8	$+1.7 \times 10^{-6}$	$+9.6 \times 10^{-6}$	0.2853
NSTR22c04	512^3	Gaussian Temp.	6.4×10^{-5}	1.9×10^{-5}	-12.7	$+1.3 \times 10^{-6}$	$+7.1 \times 10^{-6}$	0.2811
NSTR21c01	64^3	SB96	9.4×10^{-4}	9.0×10^{-6}	-2.1	$+1.9 \times 10^{-5}$	$+2.4 \times 10^{-5}$	0.2059
NSTR21c03	256^3	SB96	3.6×10^{-6}	3.9×10^{-5}	-10.3	$+6.0 \times 10^{-6}$	$+7.7 \times 10^{-6}$	0.2424

Table B2. Simulations of unforced compressible hydrodynamic convection in the shearing box with different numerical schemes in PLUTO. No explicit viscosity or heating/cooling/thermal relaxation was included. All simulations were run in a box of size $4H \times 4H \times 4H$, and initialized with the Gaussian temperature profile discussed in Appendix A2. The time-averages of the volume-averaged Reynolds stress $\langle R_{xy}\rangle$ and of the volume-averaged angular momentum transport parameter $\langle\alpha\rangle$ have been taken over an interval spanning non-linear saturation. All simulations were run for 38 orbits.

Run	Resolution	Solver	Interpolation	Time-stepping	$\langle\langle R_{xy}\rangle\rangle_t$	$\langle\langle\alpha\rangle\rangle_t$	Comments
NSTR22e01a	64^3	HLLC	WENO3	RK3	$+4.5 \times 10^{-6}$	$+2.5 \times 10^{-5}$	–
NSTR22e01c	256^3	HLLC	WENO3	RK3	$+1.8 \times 10^{-6}$	$+1.0 \times 10^{-5}$	–
NSTR22e02a	64^3	HLL	WENO3	RK3	-1.3×10^{-7}	-7.4×10^{-7}	$\vert\delta\mathbf{u}\vert = \mathcal{O}(10^{-3})$
NSTR22e02c	256^3	HLL	WENO3	RK3	-1.1×10^{-5}	-6.1×10^{-5}	–
NSTR22e03a	64^3	TVDLF	WENO3	RK3	–	–	No instability
NSTR22e11a	64^3	HLLC	Linear TVD	Hancock	$+3.4 \times 10^{-6}$	$+2.2 \times 10^{-5}$	–
NSTR22e11i	256^3	HLLC	Linear TVD	Hancock	$+1.5 \times 10^{-6}$	$+9.2 \times 10^{-6}$	UMIST limiter
NSTR22e12a	64^3	HLL	Linear TVD	Hancock	-2.2×10^{-5}	-1.4×10^{-4}	–
NSTR22e12b	128^3	HLL	Linear TVD	Hancock	-1.6×10^{-5}	-1.0×10^{-4}	–
NSTR22e12c	256^3	HLL	Linear TVD	Hancock	-9.5×10^{-6}	-6.2×10^{-5}	–
NSTR22e12i	256^3	HLL	Linear TVD	Hancock	-1.2×10^{-5}	-8.2×10^{-5}	UMIST limiter
NSTR22e13a	64^3	TVDLF	Linear TVD	Hancock	-3.5×10^{-5}	-2.3×10^{-4}	–
NSTR22e13b	128^3	TVDLF	Linear TVD	Hancock	-1.4×10^{-5}	-9.4×10^{-5}	–
NSTR22e21a	64^3	HLLC	Linear TVD	RK2	$+4.0 \times 10^{-6}$	$+2.6 \times 10^{-5}$	–
NSTR22e22a	64^3	HLL	Linear TVD	RK2	-2.2×10^{-5}	-1.3×10^{-4}	–
NSTR22e22b	128^3	HLL	Linear TVD	RK2	-2.0×10^{-5}	-1.2×10^{-4}	–
NSTR22e23a	64^3	TVDLF	Linear TVD	RK2	-1.9×10^{-5}	-1.2×10^{-4}	–
NSTR22e23b	128^3	TVDLF	Linear TVD	RK2	-1.3×10^{-5}	-7.6×10^{-4}	–

Table B3. Simulations of unforced compressible hydrodynamic convection in the shearing box with explicit kinematic viscosity ν and thermal diffusivity χ included. No thermal relaxation was included. All simulations were run in a box of size $4H \times 4H \times 4H$. $\langle\langle\alpha\rangle\rangle_t\vert_{\mathrm{linear}}$ is the volume-averaged alpha viscosity parameter time-averaged over the linear phase, $\langle\langle\alpha\rangle\rangle_t\vert_{\mathrm{NL}}$ is the same quantity time-averaged over the non-linear phase, and $\min(\langle\$R_{xy}\rangle)$ is the minimum value of the xy-component of the volume-averaged Reynolds stress.

Run	Resolution	Instability?	ν	χ	Ra	Ri	Pr	$\langle\langle\alpha\rangle\rangle_t\vert_{\mathrm{linear}}$	$\langle\langle\alpha\rangle\rangle_t\vert_{\mathrm{NL}}$	$\min(\langle R_{xy}\rangle)$
NSTR22Ra1	256^3	N	1.075×10^{-3}	4.300×10^{-4}	10^5	0.05	2.5	–	–	–
NSTR22Ra2	256^3	Y	3.400×10^{-4}	1.360×10^{-4}	10^6	0.05	2.5	-5.1×10^{-6}	-3.9×10^{-5}	-2.1×10^{-5}
NSTR22Ra3	256^3	Y	1.075×10^{-4}	4.300×10^{-5}	10^7	0.05	2.5	-3.5×10^{-6}	$+3.9 \times 10^{-6}$	-2.7×10^{-5}
NSTR22Ra4a	256^3	Y	3.400×10^{-5}	1.360×10^{-5}	10^8	0.05	2.5	-1.0×10^{-6}	$+1.4 \times 10^{-5}$	-1.1×10^{-5}
NSTR22Ra4b	512^3	Y	3.400×10^{-5}	1.360×10^{-5}	10^8	0.05	2.5	-2.8×10^{-6}	$+7.1 \times 10^{-6}$	-1.4×10^{-5}
NSTR22Ra5a	256^3	Y	1.075×10^{-5}	4.300×10^{-6}	10^9	0.05	2.5	-3.5×10^{-8}	$+1.5 \times 10^{-5}$	-9.9×10^{-8}
NSTR22Ra5b	512^3	Y	1.075×10^{-5}	4.300×10^{-6}	10^9	0.05	2.5	-1.6×10^{-7}	$+1.3 \times 10^{-5}$	-3.4×10^{-6}
NSTR22Ra6a	256^3	Y	3.400×10^{-6}	1.360×10^{-6}	10^{10}	0.05	2.5	$+7.6 \times 10^{-8}$	$+1.5 \times 10^{-5}$	$+2.2 \times 10^{-13}$
NSTR22Ra6b	512^3	Y	3.400×10^{-6}	1.360×10^{-6}	10^{10}	0.05	2.5	-1.1×10^{-8}	$+1.6 \times 10^{-5}$	-4.9×10^{-8}

This paper has been typeset from a TEX/LATEX file prepared by the author.