**Astronomy & Astrophysics**

# The subcritical baroclinic instability in local accretion disc models*

G. Lesur and J. C. B. Papaloizou

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK
e-mail: g.lesur@damtp.cam.ac.uk

## ABSTRACT

*Context.* The presence of vortices in accretion discs has been debated for more than a decade. Baroclinic instabilities might be a way to generate these vortices in the presence of a radial entropy gradient. However, the nature of these instabilities is still unclear and 3D parametric instabilities can lead to the rapid destruction of these vortices.
*Aims.* We present new results exhibiting a subcritical baroclinic instability (SBI) in local shearing box models. We describe the 2D and 3D behaviour of this instability using numerical simulations and we present a simple analytical model describing the underlying physical process.
*Methods.* We investigate the SBI in local shearing boxes, using either the incompressible Boussinesq approximation or a fully compressible model. We explore the parameter space varying several local dimensionless parameters and we isolate the regime relevant for the SBI. 3D shearing boxes are also investigated using high resolution spectral methods to resolve both the SBI and 3D parametric instabilities.
*Results.* A subcritical baroclinic instability is observed in flows stable for the Solberg-Høiland criterion using local simulations. This instability is found to be a nonlinear (or subcritical) instability, which cannot be described by ordinary linear approaches. It requires a radial entropy gradient weakly unstable for the Schwartzchild criterion and a strong thermal diffusivity (or equivalently a short cooling time). In compressible simulations, the instability produces density waves which transport angular momentum outward with typically $\alpha \lesssim 3 \times 10^{-3}$, the exact value depending on the background temperature profile. Finally, the instability survives in 3D, vortex cores becoming turbulent due to parametric instabilities.
*Conclusions.* The subcritical baroclinic instability is a robust phenomenon, which can be captured using local simulations. The instability survives in 3D thanks to a balance between the 2D SBI and 3D parametric instabilities. Finally, this instability can lead to a *weak* outward transport of angular momentum, due to the generation of density waves by the vortices.

**Key words.** accretion, accretion disks – instabilities – planets and satellites: formation – turbulence

## 1. Introduction

The existence of long-lived vortices in accretion discs was first proposed by von Weizsäcker (1944) in a model of planet formation. This idea was reintroduced by Barge & Sommeria (1995) to accelerate the formation of planetesimals through a dust trapping process. Moreover, large scale vortices can lead to a significant transport of angular momentum thanks to the production of density waves (Johnson & Gammie 2005b). Many physical processes have been introduced in the literature to justify the existence of these vortices such as Rossby wave instabilities (Lovelace et al. 1999), planetary gap instabilities (de Val-Borro et al. 2006, 2007), 3D circulation models (Barranco & Marcus 2005), MHD turbulence (Fromang & Nelson 2005) and the global baroclinic instability (Klahr & Bodenheimer 2003).

Baroclinic instabilities in accretion discs has regained interest in the past few years. A first version of these instabilities (the global baroclinic instability or GBI) was mentioned by Klahr & Bodenheimer (2003) using a purely numerical approach and considering a disc with a radial entropy gradient. However, the presence of this instability was affected by changing the

numerical scheme used in the simulations, casting strong doubts on this instability as a real physical processes. The linear properties of the GBI were then investigated by Klahr (2004) and Johnson & Gammie (2005a). However, only transient growths were found in this context. Klahr (2004) speculated that these transient growths could lead to a nonlinear instability explaining the result of Klahr & Bodenheimer (2003), but it is still unclear whether transient growths are relevant for nonlinear instabilities (see e.g. Lesur & Longaretti 2005, and reference therein for a complete discussion of this point). The nonlinear problem was then studied in local shearing boxes by Johnson & Gammie (2006) using fully compressible numerical methods. These authors found no instability in the Keplerian disc regime and concluded that the GBI was "either global or nonexistent".

Nevertheless, baroclinic processes were studied again by Petersen et al. (2007a) using anelastic global simulations and including new physical processes. As in Klahr & Bodenheimer (2003), a weak radial entropy gradient was imposed in the simulation. However, a cooling function was also included in their model, in order to force the system to relax to the imposed radial temperature profile. In these simulations, Petersen et al. (2007a) observed spontaneous formation of vortices with radial entropy gradients compatible with accretion disc thermodynamics. In a subsequent paper (Petersen et al. 2007b), these vortices

were found to survive for several hundreds orbits. According to the authors, their disagreement with Johnson & Gammie (2006) was due to a larger Reynolds number and the use of spectral methods, a possibility already mentioned by Johnson & Gammie (2006).

In this paper, we revisit baroclinic instabilities in a local setup (namely the shearing box) both in the incompressible-Boussinesq approximation and in fully compressible simulations. The aim of this paper is to clarify several of the points which have been discussed in the literature for the past 6 years about the existence, the nature and the properties of baroclinic instabilities. We show that a subcritical baroclinic instability (or shortly SBI) *does exist* in shearing boxes. This instability is *nonlinear* (or subcritical) and *strongly linked to thermal diffusion*, a point already mentioned by Petersen et al. (2007b). We also present results concerning the behaviour of the SBI in 3 dimensions, as vortices are known to be unstable because of parametric instabilities (Lesur & Papaloizou 2009). We begin in Sect. 2 by presenting the equations, introducing the dimensionless numbers and describing the numerical methods used in the rest of the paper. Section 3 is dedicated to 2D incompressible-Boussinesq simulations and a qualitative understanding of the instability. We present in Sect. 4 our results in compressible shearing boxes and in Sect. 5 the 3D behaviour of the instability. Conclusions and implications of our findings are discussed in Sect. 6.

## 2. Local model

### 2.1. Equations

In the following, we will assume a local model for the accretion disc, following the shearing sheet approximation. The reader may consult Hawley et al. (1995), Balbus (2003) and Regev & Umurhan (2008) for an extensive discussion of the properties and limitations of this model. As a simplification, we will assume the flow is incompressible, consistently with the small shearing box model (Regev & Umurhan 2008). The shearing-box equations are found by considering a Cartesian box centred at $r = R_0$, rotating with the disc at angular velocity $\Omega = \Omega(R_0)$. We finally introduce in this box a radial stratification using the Boussinesq approximation (Spiegel & Veronis 1960). Defining $r - R_0 \rightarrow x$ and $R_0\phi \rightarrow y$ as in Hawley et al. (1995), one eventually obtains the following set of equations:

$$\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \boldsymbol{\nabla}\boldsymbol{u} = -\boldsymbol{\nabla}\Pi - 2\boldsymbol{\Omega} \times \boldsymbol{u} + 2\Omega S x\boldsymbol{e_x} - \Lambda N^2 \theta\boldsymbol{e_x} + \nu\Delta\boldsymbol{u}, \quad (1)$$

$$\partial_t \theta + \boldsymbol{u} \cdot \boldsymbol{\nabla}\theta = u_x/\Lambda + \mu\Delta\theta \quad (2)$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0, \quad (3)$$

where $\boldsymbol{u} = u_x\boldsymbol{e_x} + u_y\boldsymbol{e_y} + u_z\boldsymbol{e_z}$ is the fluid velocity, $\theta$ the potential temperature deviation, $\nu$ the kinematic viscosity and $\mu$ the thermal diffusivity. In these equations, we have defined the mean shear $S = -r\partial_r\Omega$, which is set to $S = (3/2)\Omega$ for a Keplerian disc. One can check easily that the velocity field $\boldsymbol{U} = -S x\boldsymbol{e_y}$ is a steady solution of these equations. In the following we will consider the evolutions of the perturbations $\boldsymbol{v}$ (not necessarily small) of this profile defined by $\boldsymbol{v} = \boldsymbol{u} - \boldsymbol{U}$.

The generalised pressure $\Pi$ is calculated solving a Poisson equation with the incompressibility condition (3). For homogeneity and consistency with the traditional Boussinesq approach, we have introduced a stratification length $\Lambda$. Note however that $\Lambda$ disappears from the dynamical properties of these equations as one can renormalise the variables defining $\theta' \equiv \Lambda\theta$.

The stratification itself is controlled by the Brunt-Väisälä frequency $N$, defined for a perfect gas by

$$N^2 = -\frac{1}{\gamma\rho}\frac{\partial P}{\partial R}\frac{\partial}{\partial R}\ln\left(\frac{P}{\rho^\gamma}\right), \quad (4)$$

where $P$ and $\rho$ are assumed to be the background equilibrium profile and $\gamma$ is the adiabatic index. With these notations, one can recover the ordinary thermodynamical variables as $\theta \equiv \delta\rho/\rho$, where $\delta\rho$ is the density perturbation and $\rho$ is the background density. The stratification length is then defined by $\Lambda \equiv -\partial_R P/(\rho N^2)$.

### 2.2. Dimensionless numbers

The system described above involves several physical processes. To clarify the regime in which we are working and to make easier comparisons with previous work, we define the following dimensionless numbers:

- The Richardson number $Ri = N^2/S^2$ compares the shear timescale to the buoyancy timescale. This definition is equivalent to the definition of Johnson & Gammie (2005a).
- The Peclet number $Pe = S L^2/\mu$.
- The Reynolds number $Re = S L^2/\nu$.

$L$ is a typical scale of the system, chosen to be the radial box size ($L_x$) in our notations.

The linear stability properties of this flow are quite well understood. The flow is linearly stable for axisymmetric perturbations when the Solberg-Hoïland criterion is satisfied. This criterion may be written locally as:

$$2\Omega(2\Omega - S) + N^2 > 0 \quad \text{Stability}, \quad (5)$$

or equivalently in a Keplerian disc, $Ri > -4/9$. Another linear stability criterion, the Schwarzschild criterion, is often used for convectively stable flows *without rotation nor shear*. This criterion reads

$$N^2 > 0 \quad \text{Stability}, \quad (6)$$

or equivalently $Ri > 0$. As we will see in the following, this criterion is the relevant one for the SBI.

When viscosity and thermal diffusivity are included, the Solberg-Hoïland criterion is modified, potentially leading to the Goldreich-Schubert-Fricke (GSF) instability (Goldreich & Schubert 1967; Fricke 1968). The stability criterion is in that case

$$\mu 2\Omega(2\Omega - S) + \nu N^2 > 0 \quad \text{Stability}. \quad (7)$$

This criterion is satisfied when both (5) is verified and $\nu/\mu < 1$, which corresponds to the regime studied in this paper.

### 2.3. Numerical methods

We have used two different codes to study this instability. When using the Boussinesq model (Sects. 3 and 5), we have used SNOOPY, a 3D incompressible spectral code. This code uses an implicit scheme for thermal and viscous diffusion, allowing us to study simulations with small $Pe$ without any strong constrain on the CFL condition. The spectral algorithm of SNOOPY has now been used in several hydrodynamic and magnetohydrodynamic studies (e.g. Lesur & Longaretti 2005, 2007) and is freely available on the author's website. For compressible simulations (Sect. 4), we have used NIRVANA (Ziegler & Yorke 1997). NIRVANA has been used frequently in the past to study

various problems involving MHD turbulence in the shearing box (Fromang & Papaloizou 2006; Papaloizou et al. 2004).

In the following, we use the shearing sheet boundary conditions (Hawley et al. 1995) in the radial direction. This is made possible by the use of the Boussinesq approximation in which only the *gradients* of the background profile appear explicitly, through $\Lambda$ and $N$. Therefore, when $\Lambda$ and $N$ are constant through the box, one can assume that the thermodynamic fluctuation $\theta$ is shear-periodic, consistently with the shearing-sheet approximation.

## 3. 2D subcritical baroclinic instability in incompressible flows

To start our investigation, we consider the simplest case of a 2D $(x, y)$ problem in an infinitely thin disc. This setup is the local equivalent of the 2D global anelastic setup of Petersen et al. (2007a), and one expects to find similar properties in both cases if the instability is local. In two dimensions, it is easier to consider the vorticity equation instead of (1). Defining the vertical vorticity of the pertubations by $\omega = \partial_x v_y - \partial_y v_x$, we have:

$$\partial_t \omega + \boldsymbol{v} \cdot \boldsymbol{\nabla} \omega - S x \partial_y \omega = \Lambda N^2 \partial_y \theta + \nu \Delta \omega. \tag{8}$$

In this formulation, the only source of enstrophy $\langle \omega^2/2 \rangle = \int \mathrm{d}x \mathrm{d}y \, \omega^2/2$ is the baroclinic term $\Lambda N^2 \partial_y \theta$ which will be shown to play an important role for the instability.

The simulations presented in this section are computed in a square domain $L_x = L_y = 1$ with a resolution of $512 \times 512$ grid points using our spectral code. When not explicitly mentioned, the time unit is the shear timescale $S^{-1}$. We choose the fiducial parameters to be $Re = 4 \times 10^5$; $Pe = 4 \times 10^3$ and $Ri = -0.01$, in order to have a flow linearly stable for the Solberg-Høiland criterion (5). These parameters are close to the one used by Petersen et al. (2007a) and are compatible with disc thermodynamics.

### 3.1. Influence of initial perturbations

In the first numerical experiment, we choose to test the effect of the initial conditions, keeping constant the dimensionless parameters. In our initial conditions, we excite randomly the largest wavelengths of the vorticity field with an amplitude $Ap$. This initial condition is slightly different from the one used by Petersen et al. (2007a,b) who introduced perturbations in the temperature field. In each experiment we modify the amplitude of the initial perturbation $Ap$, and follow the time evolution of the total enstrophy (Fig. 1)

The numerical results indicate clearly the presence of a *nonlinear* or *subcritical* transition in the flow. Indeed, we find the "instability" for large enough initial perturbations. This also confirms the result of Johnson & Gammie (2006): there is no linear instability in the presence of a weak radial stratification. Moreover, one of the reasons that Johnson & Gammie (2006) did not observe any transition could be because of the weak initial perturbations they used (between $10^{-12}$ and $10^{-4}$). Using similar initial conditions, we did not observe any transition either. The existence of the instability in our simulations was checked by doubling the resolution ($1024 \times 1024$), keeping constant all the dimensionless numbers. We found no significant difference between high and low resolution results, showing that our simulations were converged for this problem.

When the system undergoes a subcritical transition, it develops long-lived *self-sustained* vortices, as shown for $Ap = 1$ in Fig. 2 (top). For such a large Reynolds number, it is known
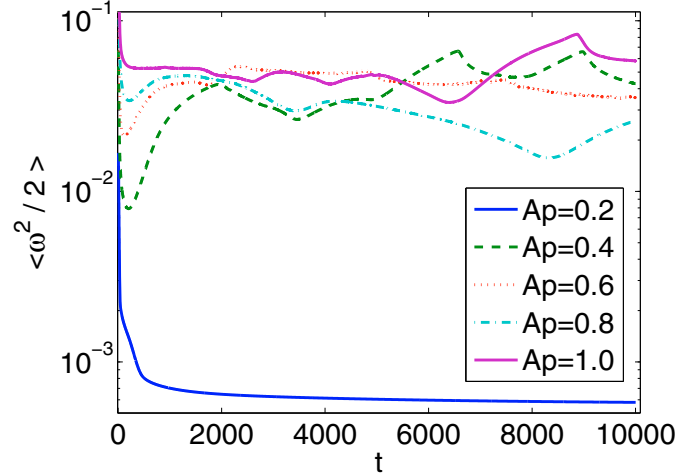


**Fig. 1.** Volume averaged enstrophy for several initial amplitude perturbation ($Ap$) in arbitrary units. A subcritical instability is observed for finite amplitude perturbations between $Ap = 0.2$ and $Ap = 0.4$.

that vortices survive for long times (see e.g. Umurhan & Regev 2004), even without any baroclinic effect. To check that the observed vortices were really due to the baroclinic term, we have carried out the exact same simulation but without stratification (Fig. 2 bottom). This simulation also shows the formation of vortices but these are lately dissipated on a few hundred shear times. At $t = 500 \, S^{-1}$ the difference between the simulation with and without baroclinicity becomes obvious.

The vortices observed in these incompressible simulations lead to a very weak and strongly oscillating turbulent transport of angular momentum. One finds typically an *inward* transport with $\alpha = \langle v_x v_y \rangle / S L^2 \simeq -3 \times 10^{-5}$. This is consistent with the results published by Petersen et al. (2007b) in global simulations.

In several other numerical experiments (not shown here), we have noticed that the amplitude threshold required to get the instability depends on $Re$ and $Pe$, a larger Reynolds number being associated with a weaker initial perturbation. This dependency was pointed out by Petersen et al. (2007a) and it indicates that the amplitude threshold in a realistic disc could be very small (i.e. much smaller than the sound speed). Note however that this threshold might also be scale dependent, a problem which has not been addressed here.

### 3.2. Influence of the Richardson number

To understand how the instability depends on the amplitude of the baroclinic term, we have carried out several runs with $Re = 4 \times 10^5$ and $Pe = 4 \times 10^3$, varying $Ri$ from $-0.02$ to $-0.16$ and from $0.02$ to $0.016$. We compare the resulting shell-integrated enstrophy spectra ($\hat{\omega}_k^2/2$) in Fig. 4. When $Ri < 0$ and $|Ri|$ becomes larger, the instability tends to amplify smaller and smaller scales. In particular for $Ri = -0.02$, the dominant scale is clearly the box scale, as already observed for the fiducial run. This trend is also observed looking directly at vorticity snapshots (Fig. 3). The sign of $Ri$ is also of importance for the onset of the SBI. To demonstrate this effect, we show the time history of the total enstrophy for positive and negative $Ri$ in Fig. 5. In the cases of positive $Ri$, the total enstrophy decays until the flow becomes axisymmetric. Perturbations are then damped very slowly on a viscous timescale. We conclude from these results that a necessary condition for the SBI to appear is a flow which do not satisfy the Schwarzschild criterion (6).
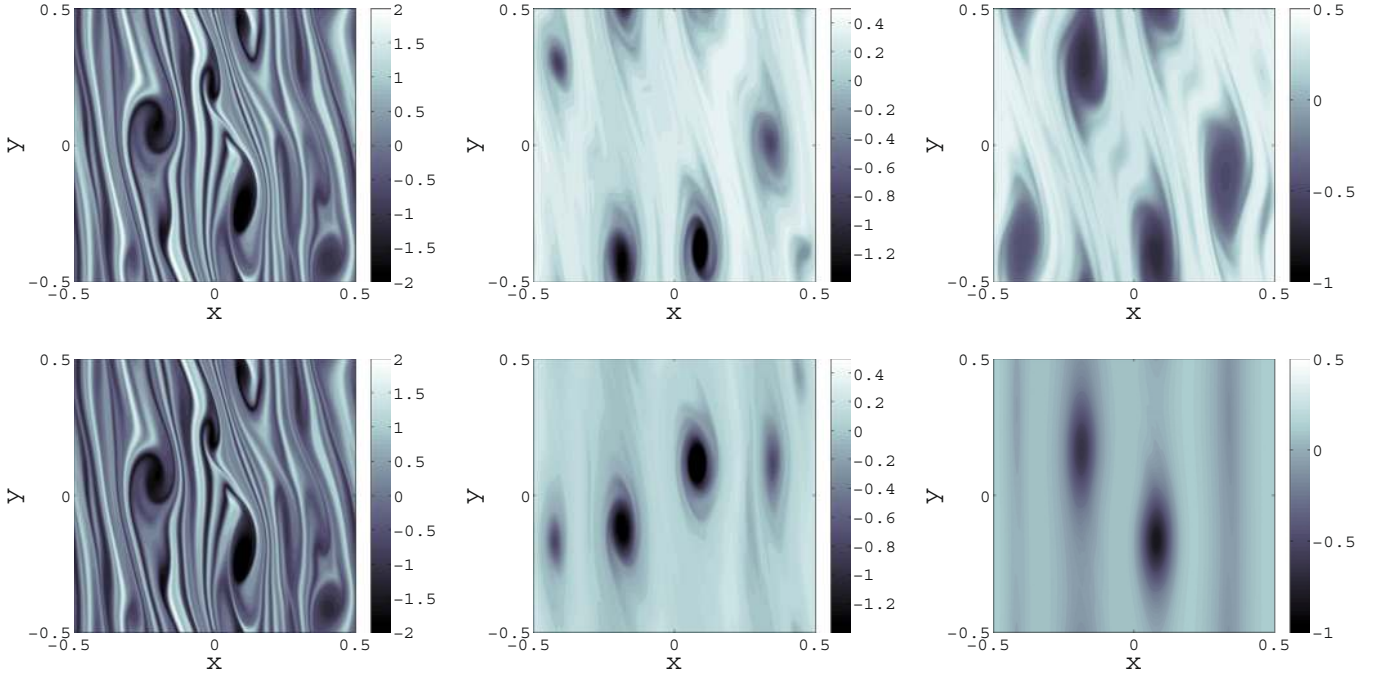
**Fig. 2.** Evolution of the vorticity in the fiducial case. Top row has a baroclinic term with $Ri = -0.01$. Bottom row has no baroclinicity. We show $t = 10$ (*left*), $t = 100$ (*middle*), $t = 500$ (*right*).
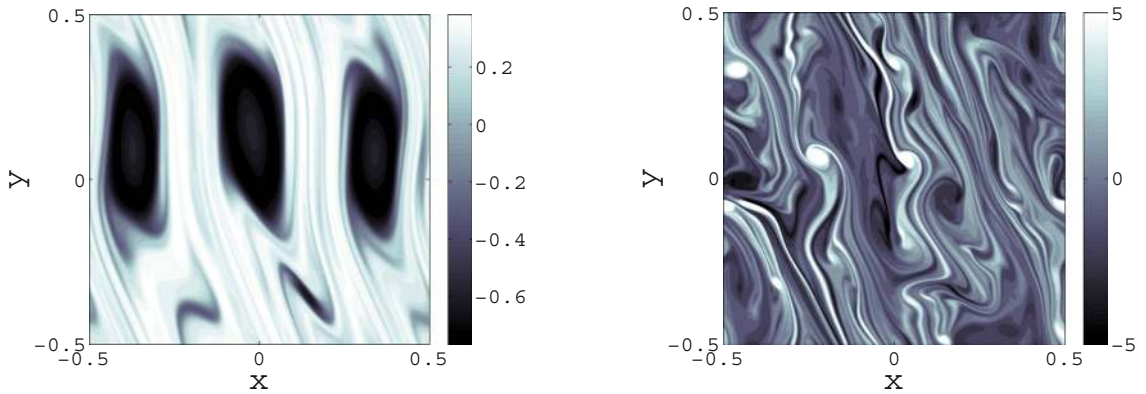


**Fig. 3.** Vorticity map for $Ri = -0.02$ (*left*) and $Ri = -0.16$ (*right*) at $t = 500$. As already shown by the spectra, small scales are dominant for larger $|Ri|$.

### 3.3. Influence of thermal diffusion

The importance of thermal cooling and thermal diffusion was already pointed out by Petersen et al. (2007b). To check this dependancy in our local model, we have considered several simulations with $Ri = -0.01$, varying $Pe$ from $Pe = 20$ to $Pe = 16\,000$. The resulting enstrophy evolution is presented for several of these runs in Fig. 6. Looking at the snapshots of the vorticity field for these simulations, we find the SBI approximatively for $50 \leq Pe \leq 8000$. Assuming a typical vortex size $l \sim 0.25$ (see e.g. Fig. 2), these Peclet numbers correspond to thermal diffusion times $3\,S^{-1} \leq \tau_{\rm diff} \leq 500\,S^{-1}$ over the vortex size, with an optimum found for $\tau_{\rm diff} \simeq 10\,S^{-1}$ ($Pe = 250$). Note that the cooling time used by Petersen et al. (2007b) lies typically in this range of values.

### 3.4. Instability mechanism

Since the flow is subcritically unstable, no linear analysis can capture this instability entirely. As shown by Johnson & Gammie (2005a), an ensemble of shearing waves in a baroclinic flow is subject to a transient growth with an amplification going asymptotically like[1] $|Ri| t^{1-4Ri}$ for $|Ri| \ll 1$ when $\mu = \nu = 0$, the waves being ultimately decaying when $\mu > 0$ or $\nu > 0$. However, this tells us little to nothing about the nonlinear behaviour of such a flow. Indeed, it is known that barotropic Keplerian flows undergo arbitrarily large transient amplifications (see e.g. Chagelishvili et al. 2003), but subcritical transitions are yet to be found in that case (Hawley et al. 1999; Lesur & Longaretti 2005; Ji et al. 2006). In the SBI case, the subcritical transition happens only for negative $Ri$ as shown above, in flagrant contradiction with the linear amplification described by Johnson & Gammie (2005a).

---

[1] Note that the amplification occurs independently of the sign of $Ri$, as already mentioned by Johnson & Gammie (2005a).
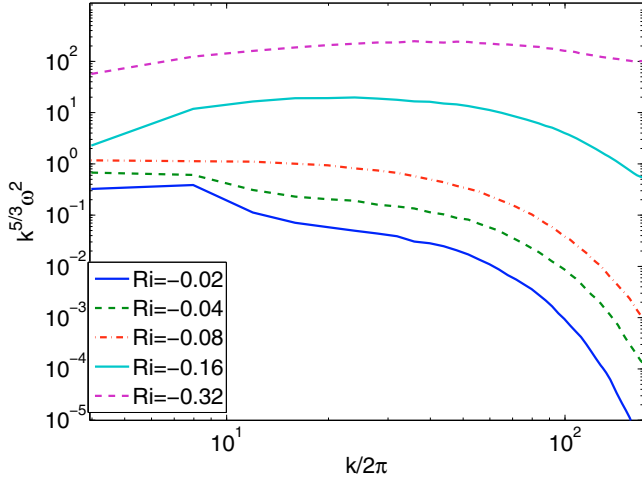
**Fig. 4.** Enstrophy spectrum as a function of the Richardson number $Ri$. As $|Ri|$ becomes larger, the instability moves to smaller scales.
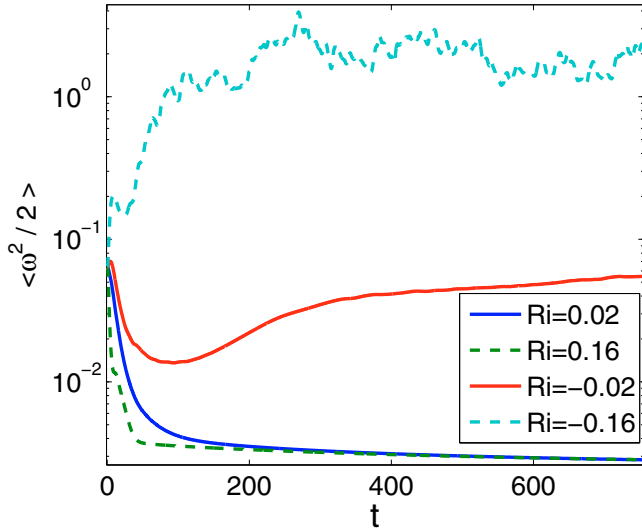


**Fig. 5.** Time history of the total enstrophy for positive and negative Richardson number. The instability is observed only for negative $Ri$.

This is a strong indication that linear theory is of little use for describing this instability.

To isolate a mechanism for a subcritical instability, one should start from a non-linear structure observed in the simulations. For the SBI, these structures are self-sustained vortices. In order to understand how baroclinic effects can feedback on the vortex structure, we initialised our fiducial simulation with a Kida vortex of aspect ratio 4 (Kida 1981). Although this vortex is known to be an exact nonlinear solution of the inviscid equation of motions, it is slowly modified by the explicit viscosity and the baroclinic term, leading to a slow growth of the vortex. We show in Fig. 7 the resulting vortex structure (left) and the associated baroclinic term $\Lambda N^2 \partial_y \theta$ (right) at $t = 100\,S^{-1}$. Note that these structures are quasi steady and evolve on timescales much longer than the shear timescale. It is clear from these snapshots that the baroclinic feedback tends to amplify the vorticity located inside the vortex, leading to the growth of the vortex itself.

To understand the origin of the baroclinic feedback, let's assume the physical quantities are evolving slowly in time, as is
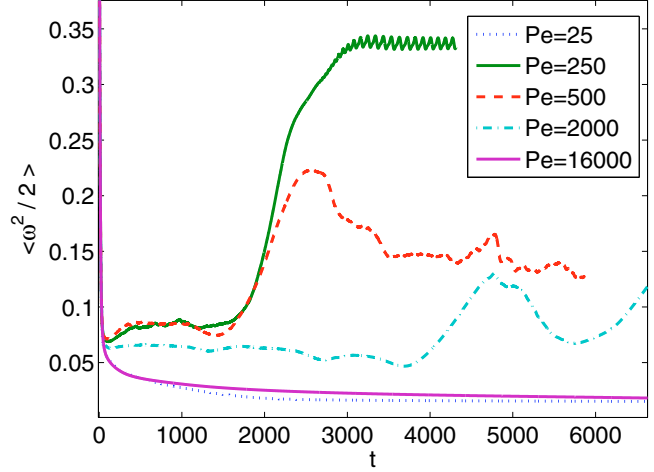


**Fig. 6.** Time history of the total enstrophy for several Peclet number ($Pe$). The instability is observed for $50 \leq Pe \leq 8000$.

observed in the simulations, so we may write

$$\omega = \omega(\epsilon t) \equiv \omega(\tau), \tag{9}$$
$$\theta = \theta(\epsilon t) \equiv \theta(\tau), \tag{10}$$

where $\epsilon$ is a small parameter and it is assumed that $|\theta_\tau|$ is of the same order as $|\boldsymbol{v} \cdot \nabla\theta|$, and similarly for $\omega$. We therefore have to solve:

$$\epsilon\partial_\tau\omega + (\boldsymbol{v} \cdot \nabla - S\,x\partial_y)\omega = \Lambda N^2\partial_y\theta + \nu\Delta\omega, \tag{11}$$
$$\epsilon\partial_\tau\theta + (\boldsymbol{v} \cdot \nabla - S\,x\partial_y)\theta = v_x/\Lambda + \mu\Delta\theta. \tag{12}$$

Since the Richardson number is assumed to be small, the baroclinic feedback of (12) has to be small. As only this term can lead to an instability, we can assume it scales like $\epsilon$. The role of the viscosity is to prevent the instability from happening, damping vorticity fluctuations. Assuming we are in a regime in which the instability appears, the viscosity has to be of the order of the baroclinic term, scaling like $\epsilon$. At zeroth order in $\epsilon$, we are left with:

$$(\boldsymbol{v} \cdot \nabla - S\,x\partial_y)\omega = 0, \tag{13}$$
$$(\boldsymbol{v} \cdot \nabla - S\,x\partial_y)\theta = v_x/\Lambda + \mu\Delta\theta. \tag{14}$$

This system of equations describes a quite simple physical system: the vorticity field is a steady solution of the inviscid vorticity equation with a constant shear and without any baroclinic feedback, whereas the potential temperature results from advection-diffusion of the background entropy profile due to the flow structure. A family of steady solutions of the inviscid vorticity equation with a constant shear are known to be steady vortices, like the Kida (1981) vortex. More generally, it can be shown that any $\omega$ of the form:

$$\omega = F(\psi) \tag{15}$$

where $\psi$ is the stream function defined by $\boldsymbol{u} = \boldsymbol{e}_z \times \nabla\psi$ is solution of (13).

In the following, we assume (13) is satisfied by $\omega$ and we look for a solution to the entropy Eq. (14). As a further simplification, we neglect the advection of the perturbed entropy $\theta$ by $\boldsymbol{v}$, assuming this effect is small compared to the advection by the mean shear, although this does not affect the argument leading to the possibility of instability (see below). Using a Fourier decomposition $\theta(\boldsymbol{x}) = \int d\boldsymbol{k}\,\hat{\theta}(\boldsymbol{k})\exp(i\boldsymbol{k}\cdot\boldsymbol{x})$, one gets

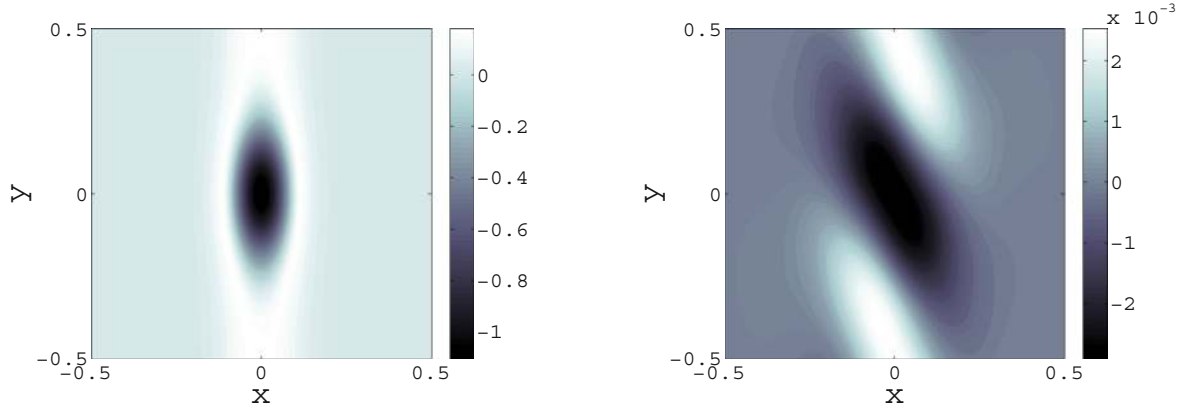$$S\,k_y\frac{d\hat{\theta}}{dk_x} + \mu k^2\hat{\theta} = \hat{v}_x/\Lambda. \tag{16}$$

**Fig. 7.** Vortex structure obtained starting from a Kida vortex. (*Left*) Vorticity map. (*Right*) Baroclinic term ($\Lambda N^2 \partial_y \theta$) map.

A solution to this equation can easily be found using standard techniques for solving first order ordinary differential equations. Assuming $|\hat{v}_x|$ decays to zero when $|\boldsymbol{k}| \to \infty$, one may write the result in the form

$$\hat{\theta}(k_x, k_y) = \frac{1}{\Lambda |S k_y|} \int_{-\infty}^{\infty} dk'_x \hat{v}_x(k'_x, k_y) G(k_x, k'_x, k_y), \qquad (17)$$

where we have defined

$$G(k_x, k'_x, k_y) = \mathcal{H}\left((k_x - k'_x) S k_y / |S k_y|\right)$$
$$\times \exp\left[\frac{\mu}{S k_y}\left(\frac{k'^3_x - k^3_x}{3} + k^2_y[k'_x - k_x]\right)\right], \qquad (18)$$

$\mathcal{H}$ being the Heaviside function. From this expression, it is formally possible to derive the time evolution of the vorticity and look for an instability using the first order term in (12). However, this requires an explicit expression for the vorticity, which is a priori unknown. Instead, we have choosen to compute the evolution of the total enstrophy of the system $\langle \omega^2/2 \rangle = \int dx dy\, \omega^2/2$:

$$\partial_t \langle \omega^2/2 \rangle = \Lambda N^2 \langle \omega \partial_y \theta \rangle - \nu \langle |\boldsymbol{\nabla} \omega|^2 \rangle. \qquad (19)$$

Evidently, an instability can occur only if the first term on the right hand side is positive. Using (17), is possible to derive an analytical expression for this term:

$$\Lambda N^2 \langle \omega \partial_y \theta \rangle = -4\pi^2 \mathfrak{R}\Bigg[N^2 \int dk_x dk_y dk'_x$$
$$\times \frac{|k_y|}{|S||\boldsymbol{k}'|^2} \hat{\omega}^*(\boldsymbol{k}) G(k_x, k'_x, k_y) \hat{\omega}(\boldsymbol{k}')\Bigg], \qquad (20)$$

where $\boldsymbol{k}' = (k'_x, k_y)$. In this expression, the existence of an instability is controlled by the Green's function $G(k_x, k'_x, k_y)$ and the shape of $\omega(\boldsymbol{k})$. To our knowledge, it is not possible to derive any general criterion for the instability, as such a criterion would depend on the background vortex considered. It is however possible to derive a criterion in the limit of a large thermal diffusivity $\mu$. In this limit, $G$ is strongly peaked around $k'_x = k_x$, and one can assume in first approximation that $G$ is proportional to a Dirac distribution. More formally, one can approximate in the limit of large thermal diffusivity:

$$G(k_x, k'_x, k_y) \simeq \mathcal{H}\left((k_x - k'_x) S k_y / |S k_y|\right) \exp\left[\frac{\mu}{S k_y} k^2 (k'_x - k_x)\right], \qquad (21)$$

when $k_x \simeq k'_x$. It is then possible to derive an expression for the baroclinic feedback as a series expansion in the small parameter $\mu^{-1}$:

$$LN^2 \langle \omega \partial_y \theta \rangle \simeq -4\pi^2 \mathfrak{R}\Bigg[N^2 \int d\boldsymbol{k}\left(\frac{k^2_y}{\mu k^4}|\hat{\omega}(\boldsymbol{k})|^2\right.$$
$$\left. - \frac{S k^3_y}{\mu^2 k^4} \hat{\omega}^*(\boldsymbol{k}) \partial_{k_x}\left[\hat{\omega}(\boldsymbol{k})/k^2\right] + \dots\right)\Bigg]. \qquad (22)$$

In this approximation, we find two competing effects. The first term on the right hand side appears when the thermal diffusion completely dominates over the shear in (16). It has a positive feedback on the total enstrophy when $N^2 < 0$ and can be seen as a source term for the SBI. The second term involves a competition between shear and diffusion. It can be seen in first approximation as a phase mixing term, and the resulting can be either positive or negative, depending on the vorticity background one considers. Physically, this term shears out the entropy structure created by the vortex and if it is too strong, it kills the positive feedback of the first term. One should note however that this term will not *reverse* the sign of the baroclinic feedback but will just weaken it by randomising the phase coherence between $\omega$ and $\theta$.

To understand the "growth" described by (22), on can derive an order of magnitude expression for the growth rate $\gamma = d \ln \langle \omega^2 \rangle / dt$ combining (19) with (22):

$$\gamma \sim \frac{(-N^2)\sigma^2}{\mu} \phi_\omega(S\sigma^2/\mu) - \frac{\nu}{\sigma^2}. \qquad (23)$$

In this expression, $\sigma$ is the typical vortex size, with the assumption $\sigma \sim 1/k$. We have included a phase mixing term with the function $\phi_\omega$, as an extension of (22). This function depends on the background vortex solution $\omega(x, y)$ one chooses and cannot be written explicitly in general. As explained above, this phase mixing weaken the baroclinic feedback when $Pe$ is large but it does not change its intrinsic nature. According to these properties, one expect $\phi_\omega(0) = 1$ and $\phi_\omega(x) \to 0$ when $x \to \infty$.

Although $\phi_\omega$ can only be determined for a specific vortex solution, one can still deduce a few general properties for the SBI. For very large thermal diffusivity (very small $Pe$) one will expect the SBI when

$$\frac{(-N^2)\sigma^4}{\nu\mu} > 1, \qquad (24)$$

since $\phi_\omega \sim 1$ in this limit. Although not quantitatively equivalent, this first limit corresponds to the $Pe > 25$ threshold of our
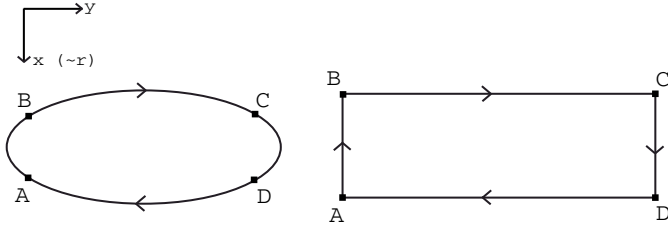
**Fig. 8.** Streamline in a vortex undergoing the SBI. A fluid particle is accelerated by buoyancy effects on A-B and C-D branches. Cooling occurs on B-C and D-A branches (see text).

simulations. It is moreover similar to the criterion for the onset of convection (in the absence of shear) based on the Rayleigh number $Ra = -N^2\sigma^4/\nu\mu$. One finds another instability condition due to phase mixing when $\mu \to 0$. Assuming $\phi_\omega(x)$ decays faster than $1/x$ when $x \to \infty$, one gets a minimum value for $\mu$ solving

$$\frac{\phi_\omega(S\sigma^2/\mu)}{\mu} > \frac{\nu}{(-N^2)\sigma^4} \tag{25}$$

which can be calculated once $\phi_\omega$ is explicitly provided. In our simulations, this second limit corresponds to the $Pe < 16\,000$ threshold.

Although not entirely conclusive, this short analytical analysis tends to explain why a relatively strong thermal diffusion is required to get the SBI. This dependancy was also pointed out by Petersen et al. (2007b) who used both a thermal diffusion and a cooling relaxation time in the entropy equation. Such a cooling time can also be introduced in our analysis in place of the thermal diffusion but this does not change the underlying physical process. We also note that in this analysis, we have assumed that a finite amplitude background vorticity satisfying (13) existed in the first place, from which we have derived the baroclinic feedback on the same vorticity field. In this sense, this analysis describes a non-linear instability and is different from the linear approach of Johnson & Gammie (2005a). In addition, we recall that we made the approximation of only including the background shear in the advection term on the left hand side of Eq. (14). We remark that the dominant driving term in the limit of large $\mu$ in (22) does not depend on this assumption because the heat diffusion term dominates in this limit.

### 3.5. Phenomenological picture

Knowing the basic mechanisms underlying the SBI, one can tentatively draw a phenomenology of this instability. For this exercise, let us consider a single streamline in a vortex subject to the SBI (Fig. 8 left). For simplicity, we reduce the trajectory followed by a fluid particle in the vortex to a rectangle (Fig. 8 right). Let us start with a fluid particle on A, moving radially inward toward B. This fluid particle is initially in thermal equilibrium, and we assume the motion from A to B is fast enough to be considered approximately adiabatic. As the particle moves inward, its temperature and density slowly deviate from the background values. If the background entropy profile is convectively unstable (condition 6), the fluid particle moving inward is cooler and heavier than the surrounding, and is consequently subject to an inward acceleration due to gravity (the particle "falls"). Once the particle reaches B, it drifts azimuthally toward C. On this trajectory, the background density and temperature is constant. Therefore, the fluid particle gets thermalised with the

background and, if the cooling is fast enough, it reaches C in thermal equilibrium. Between C and D, the same buoancy effect as the one between A and B is observed: as the particle moves from C to D, it is always hotter and lighter than the surrounding, creating a buoyancy force directed radially outward on the particle. Finally, a thermalisation episode happens again between D and A, closing the loop.

From this picture, it is evident that fluid particles get accelerated by buoyancy forces on A-B and C-D branches. In the end, considering many particles undergoing this buoyancy cycle, the vortex structure itself gets amplified, explaining our results. Moreover, the role played by cooling or thermal diffusion is evident. It the cooling is too fast, fluid particles tend to get thermalised on the A-B and C-D branches, reducing the buoyancy forces and the efficiency of the cycle. On the other hand, if no cooling is introduced, the particles cannot get thermalised on the B-C and D-A branches. In this case, the work due to buoyancy forces on the B-A trajectory will be exactly opposite to the work done on the C-D trajectory, neutralising the effect on average.

## 4. 2D SBI in compressible flows

In this situation we consider the 2D subcritical baroclinic instability within the framework of a compressible model. We accordingly adopt the compressible counterparts of Eqs. (1)−(3) for an ideal gas with constant ratio of specific heats $\gamma$ in the form

$$\frac{D\boldsymbol{u}}{Dt} = -\frac{1}{\rho}\boldsymbol{\nabla}P - \boldsymbol{\nabla}\Phi - 2\boldsymbol{\Omega}\times\boldsymbol{u} - 2\Omega S x\boldsymbol{e}_x + \nabla\cdot\boldsymbol{T_v}, \tag{26}$$

$$\frac{Dc^2}{Dt} = \frac{(\gamma-1)c^2}{\rho}\frac{D\rho}{Dt} + \frac{1}{\rho}\left(\mathcal{H} + \mathcal{K}\Delta c^2\right) \tag{27}$$

$$\partial_t\rho = -\boldsymbol{\nabla}\cdot\rho\boldsymbol{u}. \tag{28}$$

Here $c$ is the isothermal sound speed which is proportional to the square root of the temperature, $\boldsymbol{T_v}$ is a viscous stress tensor, $\mathcal{H}/(\gamma - 1)$ is the rate of energy production per unit volume, $\Phi$ is a general external potential, and the constant $K = \mu\rho$ where as for the incompressible case $\mu$ is a thermal diffusivity.

For simplicity we adopt a model that may be either regarded as assuming that there is a prescribed energy production rate $\mathcal{H}$ and $\gamma$ is close to unity in (27) so that the compressional heating term $\propto(\gamma - 1)$ can be dropped, or replacing the combination of the compressional heating term and $\mathcal{H}$ by a new specified energy production/cooling rate. In either view, as the energy production rate is specified, the model is simplified as Eq. (27) becomes an equation for $c^2$ alone.

We consider shearing box models of extent $L_x$ in the radial direction and $L_y$ in the azimuthal direction. A characteristic constant sound speed, $c_0$, defines a natural scale height $H = c_0/\Omega$. We present simulations using NIRVANA (see Sect. 2.3) for a small box with $L_x = L_y/2 = 0.6H$ and a large box with $L_x = L_y/2 = 1.2H$. Thus in terms of the characteristic sound speed, when Keplerian shear is the only motion, relative motions in the small box are subsonic, whereas supersonic relative velocities occur in the big box. We have also considered the two resolutions $(N_x, N_y) = (144, 288)$ and $(N_x, N_y) = (288, 576)$ and tests have shown that these give the same results. We have also performed simulations with no applied viscosity and with an applied viscosity corresponding to a Reynolds number $L_x^2\Omega/\nu = 12\,500$. This was applied as a constant kinematic Navier Stokes viscosity but with stress tensor acting on the velocity $\boldsymbol{v}$ being the deviation from the background shear rather than $\boldsymbol{u}$. This is not significant for an incompressible simulation but ensures that unwanted phenomena such as viscous overstabilty produced by perturbations
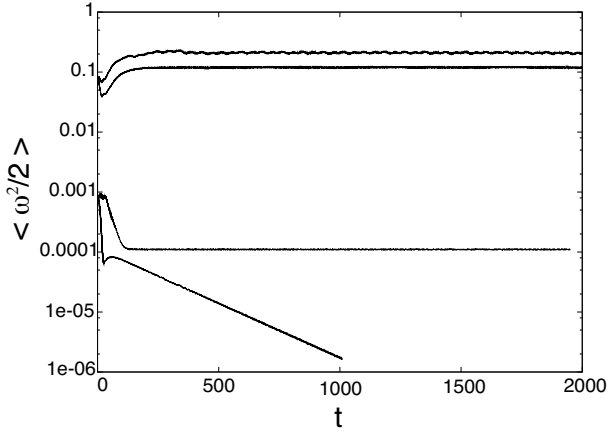
**Fig. 9.** Time history of the evolution of the enstrophy corresponding to the perturbations for the small box. The uppermost curve is for $\mathcal{A} = 0.25$ with no applied viscosity. The next uppermost is the corresponding case with an imposed viscosity. The lowermost curve is for $\mathcal{A} = 0.025$ with applied viscosity and the next lowermost curve is for the corresponding case with no viscosity.

of the background viscous stress (see Kley et al. 1993) are absent. The diffusivity corresponding to the viscosity we used is the same magnitude as the magnetic diffusivity used in Fromang et al. (2007) and as in their case we find that it is adequately resolved.

In order to perform simulations corresponding to the incompressible ones we need to impose gradients in the state variables that produce a non zero Richardson number (appropriate for $\gamma = 1$)

$$Ri = -\frac{1}{S^2\rho}\frac{\partial P}{\partial R}\frac{\partial}{\partial R}\ln\left(\frac{P}{\rho}\right). \tag{29}$$

Because of the shearing box boundary conditions, imposed background gradients have to be periodic. Thus we choose an initially uniform density $\rho = \rho_0$ and initial profiles for $P$ and $c^2$ of the form $P = \rho_0 c_0^2(1 - C_p\sin(2\pi x/L_x))$ and $c^2 = c_0^2(1 - C_p\sin(2\pi x/L_x))$, where $C_p$ is a constant. For a given value $H/L_x$, $Ri$ depends only on $C_p$ which was chosen to give a maximum value for this quantity of $-0.07$. Note that in our case, as $C_p$ is small, we have approximately $Ri \propto \cos^2(2\pi x/L_x)$ in contrast to the incompressible simulations where it is constant. Once the initial value of $\mu$ has been chosen to correspond to a specified uniform Peclet number, the external potential $\Phi$ and the energy source term $\mathcal{H}$ were chosen such that the choice of state variables with Keplerian shear resulted in an equilibrium state and then held fixed. For comparison with the incompressible simulations we adopted $Pe = 379$ for the simulations reported here. At this point we stress that in view of the rather artificial set up, this model is clearly far from definitive. However, it can be used to show that the phenomena found in the incompressible case are also seen in a compressible model but with the addition of the effects of density waves in the case of the large box.

As for the incompressible case, solutions with sustained vortices were only found for sufficiently large initial velocity perturbations of the equilibrium state. We adopted an incompressible velocity perturbation of the form

$$v_x = (3L_y\Omega/(8\pi))\mathcal{A}\sin(2\pi y/L_y)\cos(2\pi x/L_x) \tag{30}$$

and

$$v_y = -(3L_y^2\Omega/(8\pi L_x))\mathcal{A}\cos(2\pi y/L_y)\sin(2\pi x/L_x), \tag{31}$$

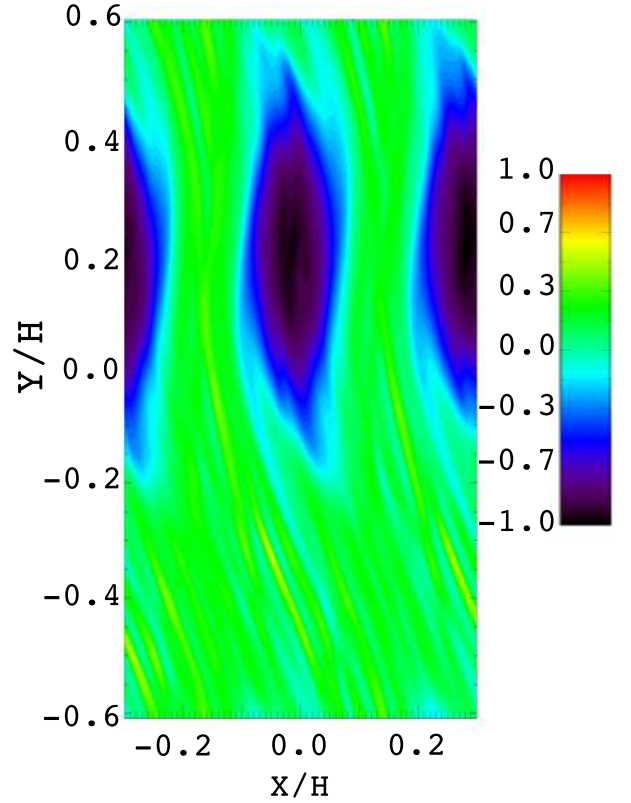where $\mathcal{A}$ is specified amplitude factor.

**Fig. 10.** Vorticity map for the small box simulation with applied viscosity and $\mathcal{A} = 0.25$ at $t = 2000$.
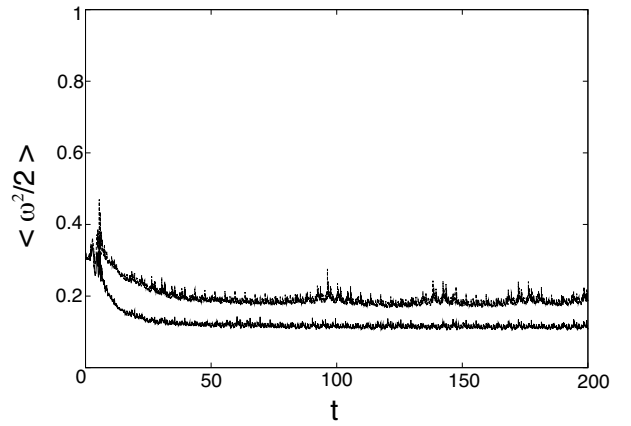


**Fig. 11.** Time history of the evolution of the enstrophy corresponding to the perturbations for the large box. The uppermost curve is for $\mathcal{A} = 0.5$ with no applied viscosity. The lower curve is for the corresponding case with an imposed viscosity.

### 4.1. Simulation results

In Fig. 9 we show the the evolution of the enstrophy associated with the perturbations for the small box. Cases with $\mathcal{A} = 0.25$ with and without applied viscosity are shown. These lead to solutions for which anticyclonic vortices are sustained for long times. The case with no viscosity is more active as expected but otherwise looks similar to the case with applied viscosity. By contrast when the amplitude of the initial perturbation is reduced to $\mathcal{A} = 0.025$ no sustained vortices are seen. This demonstrates that our numerical setup is not subject to any linear instability, such as Rossby wave instabilities (Lovelace et al. 1999). The enstrophy attains a low level in the case with no applied
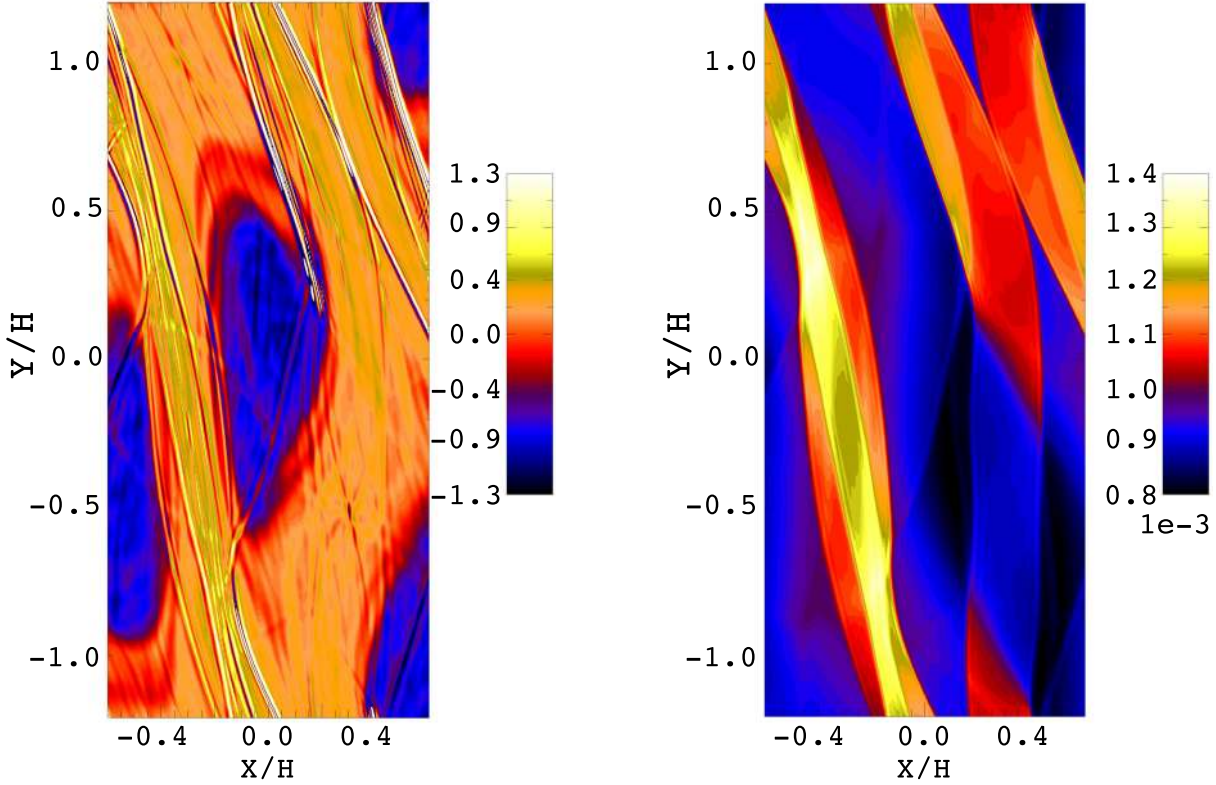
**Fig. 12.** Vorticity map (*left*) and surface density map (*right*) for the large box simulation with applied viscosity and $\mathcal{A} = 0.5$ at $t = 190$.
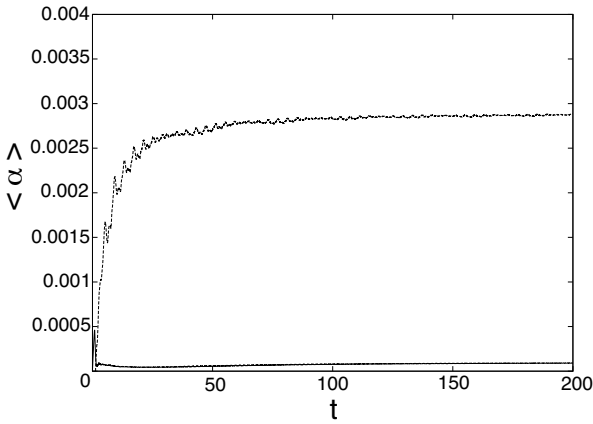


**Fig. 13.** Time history of the running means of the spatially averaged value of $\alpha$ for the small box (lower curve) and large box (upper curve) simulations with applied viscosity and with $\mathcal{A} = 0.5$ and $\mathcal{A} = 0.25$ respectively.

viscosity which is a consequence of a long wavelength linear axisymmetric disturbance that shows only very weak decay provided by numerical viscosity in this run. Figure 10 shows a vorticity map for the small box simulation with applied viscosity and $\mathcal{A} = 0.25$. Anticylconic vortices are clearly seen in this case supporting the finding from the incompressible runs that a finite amplitude initial kick is required to generate them.

The time history of the evolution of the enstrophy for large box simulations with $\mathcal{A} = 0.5$ with and without applied viscosity is shown in Fig. 11. Corresponding vorticity and surface density maps for the case with applied viscosity are shown in Fig. 12. Again the inviscid case is more active but nonetheless the corresponding maps look very similar. In these cases the anticyclonic vortices are present as in the small box case but there is increased activity from density waves as seen in the surface density maps. These waves could be generated by a process similar to the swing amplifier with vorticity source described by Heinemann & Papaloizou (2009a,b), although the structures we observe do not strictly correspond to a small scale turbulent flow. The density waves are associated with some outward angular momentum transport. However, the value of $\alpha$ measured from the volume average of the Reynolds stress is always highly fluctuating. Accordingly we plot running means as a function of time for the small box and large box with applied viscosity in Fig. 13. In the small box case there is a small residual time average $\sim 10^{-4}$ but in the large box this increases to $\sim 3 \times 10^{-3}$. This is clearly a consequence of the fact that the small box is close to the incompressible regime, whereas the large box being effectively larger than a scale height in radial width allows the vortices to become large enough to become significantly more effective at exciting density waves.

## 5. The SBI in 3D

The results presented previously were obtained in a 2D setup. It is however known that vortices like the one observed in these simulations are linearly unstable to 3D perturbations due to parametric instabilities (Lesur & Papaloizou 2009). The question of the survival of these vortices in 3D is therefore of great importance, as their destruction would lead to the disappearance of the SBI. As shown by Lesur & Papaloizou (2009), 3D instabilities involve relatively small scales compared to the vortex size when the aspect ratio[2] $\chi$ of the vortex is large. This leads to strong numerical constraints on the resolution one has to use to resolve

---

[2]  The aspect ratio of a vortex is defined by the ratio of the azimuthal size to the radial size of the vorticity patch generated by the vortex.
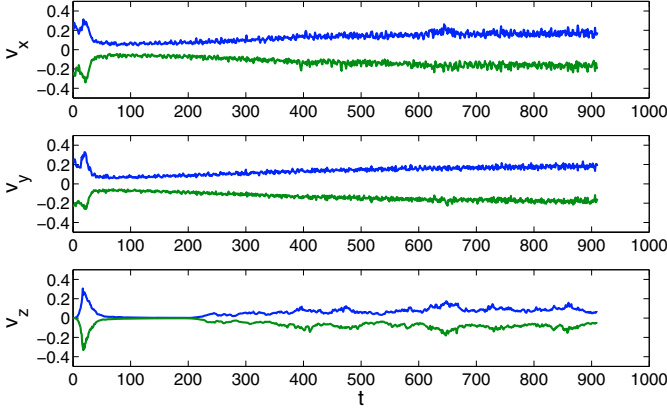
**Fig. 14.** Time history of the maxima of the 3 components of the velocity field when starting from 2D+3D noise. Once the SBI has formed vortices, 3D motions due to parametric instabilities appears and balance the 2D instability.

both the 2D SBI and 3D instabilities. To optimise the computational power, we have carried out all the 3D simulations using our spectral code in the Boussinesq approximation, with a setup similar to Sect. 3 and constant stratification.

### 5.1. Fiducial simulation

We first consider a box extended horizontally to mimic a thin disc with $L \equiv L_x = L_y = 8$ and $L_z = 1$. We set $Re = 1.02 \times 10^6$, $Pe = 6400$ and $Ri = -0.01$ with a numerical resolution $N_x \times N_y \times N_z = 1024 \times 512 \times 128$ in order to resolve the small radial structures due to the 3D instabilities. The horizontal boundary conditions are identical to the one used in 2D and periodic in the vertical direction. Note that we don't include any stratification effect in the vertical direction to reduce computational costs. Initial conditions are to be chosen with care as 3D random perturbations will generally decay rapidly due to the generation of 3D turbulence everywhere in the box. To avoid this effect, we choose to start with a large amplitude 2D white noise ($\langle \sqrt{v^2} \rangle \sim 0.1$) to which we add small 3D perturbations with an amplitude set to 1% of the 2D perturbation amplitude.

We show in Fig. 14 the time history of the extrema of the velocity field in such a simulation. As expected, vertical motions triggered by 3D instabilities appear at $t = 200$. However, these "secondary" instabilities *do not have a destructive effect on the SBI*. Instead, an almost steady state is reached at $t \simeq 600$.

Looking at the snapshots in the "saturated state" at $t = 800$ (Fig. 15) we observe the production of large scale anticyclonic vortices, as in the 2D case. However, in the core of these vortices (regions of negative vorticity), we also observe the appearance of vertical structures. Looking at the vertical component of the velocity field, one finds clearly that these vertical structures and motions are localised *inside* the vortex core. Although it is likely that the 3D instability observed in this simulation is related to the elliptical instability described by Lesur & Papaloizou (2009), we can't formally prove this point as the vortices found in the simulation are different from the one studied by Lesur & Papaloizou (2009).

Despite the presence of 3D turbulence in these vortices, the turbulent transport measured in this simulation is directed inward, with $\alpha \simeq -3 \times 10^{-5}$. This is to be expected as the 3D instabilities extract primarily their energy from the vortex structure and *not* from the mean shear. However, allowing for

compressibility will certainly change the turbulent transport as vortices will then also produce density waves (see Sect. 4).

### 5.2. Evolution of a single vortex

To isolate the interplay between the SBI and 3D elliptical instabilities, we have considered the case of an isolated 2D vortex in a 3D setup. We take the same parameters as in the fiducial case, but we consider this time a Kida vortex of aspect ratio $\chi = 6$ as an initial condition. Thanks to these initial conditions, it is possible to follow the evolution of this single vortex, and in particular measure its aspect ratio as a function of time. This is done with a post-processing script, assuming that the vortex core boundary is located at $\omega_b = 0.5[\min(\omega) + \max(\omega)]$. We then measure the vortex aspect ratio as being the ratio $l_y/l_x$ of the boundary defined above. One should note however that this procedure does not check that the vortex is still an ellipse. Moreover, it gives wrong values for $\chi$ if 3D perturbations create strong fluctuations of $\omega$.

We show in Fig. 16 the result of such a procedure as well as the extrema on the vertical velocity for comparison. Starting from $\chi = 6$, the first effect of the SBI is to reduce the vortex aspect ratio. This is consistent with the Kida vortex solution for which

$$\frac{\omega}{S} = -\frac{1}{\chi}\left(\frac{\chi+1}{\chi-1}\right). \tag{32}$$

Therefore, the vorticity amplification in the vortex core due to the baroclinic feedback leads to a reduction of the aspect ratio. When $\chi \simeq 4$, we note the appearance of strong vertical motions. During these "bursts" of 3D turbulence, the aspect ratio measurement is not reliable anymore. However, the vertical motions become ultimately weak, and a weaker vortex appears with $\chi \sim 10$. The baroclinic feedback then amplifies the vortex and the cycle starts again. Interestingly, the $\chi < 4$ Kida vortices described by Lesur & Papaloizou (2009) were shown to be strongly unstable due to a "horizontal" instability. The behaviour observed in these simulations tends to indicate that the parametric modes unstable for $\chi > 4$ are not fast enough to take over the SBI. However, as $\chi$ passes through the critical value $\chi = 4$, the horizontal modes become unstable and strongly damp the vortex in a few orbits. This interpretation is somewhat confirmed by the snapshots of the velocity field (Fig. 17), where growing perturbations are localised inside vortex cores as observed by Lesur & Papaloizou (2009).

The "burst" of turbulence observed in the Kida vortex case might be a specific property of this vortex solution. Indeed, on longer timescales, the spatial distribution of vorticity might be modified, leading to a more progressive appearance of 3D instabilities and a state more similar to the one observed in 5.1 could be achieved. However, this extreme example clearly demonstrate that 3D parametric instabilities do not lead to a total destruction of the vortices produced by the SBI.

## 6. Conclusions

In this paper, we have shown that a subcritical baroclinic instability was found in shearing boxes. Our simulations produce vortices with a dynamic similar to the description of Petersen et al. (2007a) and Petersen et al. (2007b). We find that the conditions required for the SBI are (1) a negative Richardson number (or equivalently a disc unstable for the radial Schwarzschild criterion) (2) a non negligible thermal diffusion or a cooling
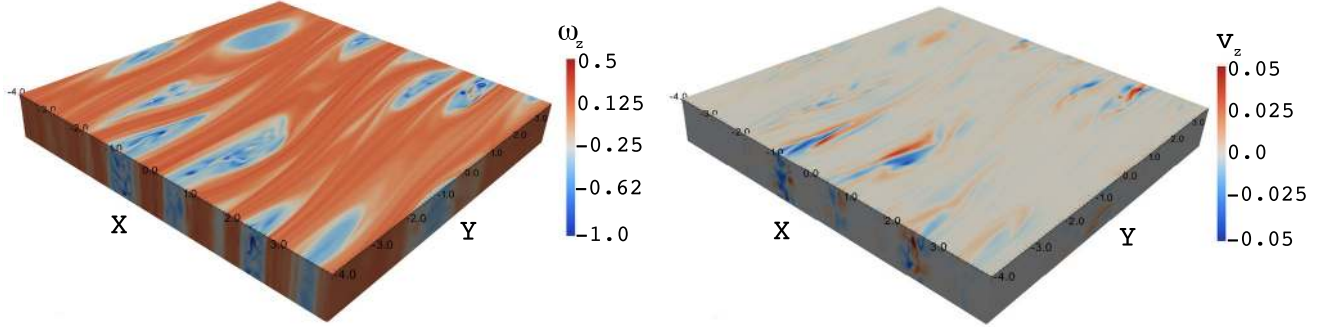
**Fig. 15.** 3D structure obtained from our fudicial 3D simulation at $t = 800$. *Left*: vertical component of the vorticity. *Right*: vertical component of the velocity field. We find that 3D instabilities involving vertical motions and vertical structures are localised inside vortex cores.
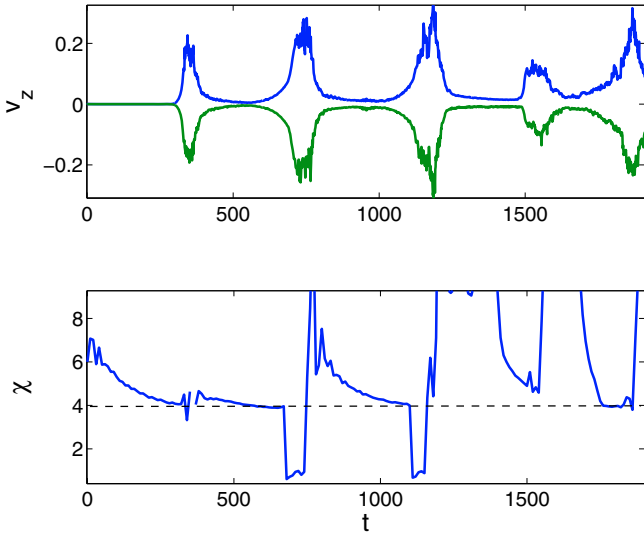


**Fig. 16.** Time history of $v_z$ extrema (*top*) and of the vortex aspect ratio $\chi$ (*bottom*). Once the SBI has formed vortices, 3D motions due to parametric instabilities appears and balance the 2D instability.

function and (3) a finite amplitude initial perturbation, as the instability is subcritical[3].

When computed in compressible shearing boxes, the vortices sustained by the SBI produce density waves which could lead to a *weak* outward angular momentum transport ($\alpha \sim 10^{-3}$), a process suggested by Johnson & Gammie (2005b). However, we would like to stress that this transport cannot be approximated by a turbulent viscosity, as the transport due to these waves is certainly a non local process. Several uncertainties remain in the compressible stratified shearing box model. In particular, the imposed temperature profile that is subsequently maintained by a heating source is somewhat artificial. This occurs while angular momentum transport causes a significant density redistribution causing the local model to deviate from the original form. Clearly one should therefore consider global compressible simulations with a realistic thermodynamic treatment to draw any firm conclusion on the long term evolution and the angular momentum transport aspect of the SBI.

Although the vortices produced by the SBI are anticyclones, the precise pressure structure inside these vortices is not as simple as the description given by Barge & Sommeria (1995). In particular, since the vorticity is of the order of the local rotation rate, vortices are not necessarily in geostrophic

---

[3] Note that point (1) and (2) were already mentioned by Petersen et al. (2007b), although in a somewhat different form.

equilibrium and pressure minima can be found in the centre of strong enough vortices. Moreover, these vortices interact with each other, which leads to complex streamline structures and vortex merging episodes. This leads us to conclude that the particle concentration mechanism proposed by Barge & Sommeria (1995) might not work in its present form for SBI vortices.

In 3D, vortices are found to be unstable to parametric instabilities. These instabilities generates random gas motions ("turbulence") in vortex cores, but they generally don't lead to a proper destruction of the vortex structure itself. More importantly, the presence of a weak hydrodynamic turbulence in vortex cores will lead to a diffusion of dust and boulders, balancing the concentration effect described above. In the end, dust concentration inside these vortices might not be large enough to trigger a gravitational instability and collapse to form planetesimals.

Given the instability criterion detailed above, one can tentatively explain why Johnson & Gammie (2006) failed to find the SBI in their simulation. First, we would like to stress that our compressible simulations are not more resolved than theirs. Since these simulations use similar numerical schemes, the argument of a too small Reynolds number proposed by Petersen et al. (2007b) seems dubious. We note however that Johnson & Gammie (2006) did not include two other important ingredients: a thermal diffusivity and finite amplitude initial perturbations. As shown in this paper, if one of these ingredient is missing, the flow never exhibits the SBI and it gets back to a laminar state rapidly. Note that even if this argument explains the negative result of Johnson & Gammie (2006), it also indicates that the SBI should be absent from the simulations presented by Klahr & Bodenheimer (2003), as no thermal diffusion nor cooling function was used (the initial conditions were not explicitly mentioned). This suggests that either the global baroclinic instability described by Klahr & Bodenheimer (2003) and the SBI are two separate instabilities, or strong numerical artefacts have produced an artificially large thermal diffusion in Klahr & Bodenheimer (2003) simulations.

In the limit of a very small kinematic viscosity, the scale at which the instability appears has to be related to the diffusion length scale $(\mu/S)^{1/2}$. In a disc, one would therefore expect the instability around that scale, provided that the entropy profile has the right slope (condition 1). However, as the instability produces enstrophy, vortices are expected to grow, until they reach a size of the order of the disc thickness where compressible effects balance the SBI source term. We conclude from this argument that even if the SBI itself is found at small scale (e.g. because of a small thermal diffusivity), the end products of this instability are large scale vortices, with a size of the order of a few disc scale heights.
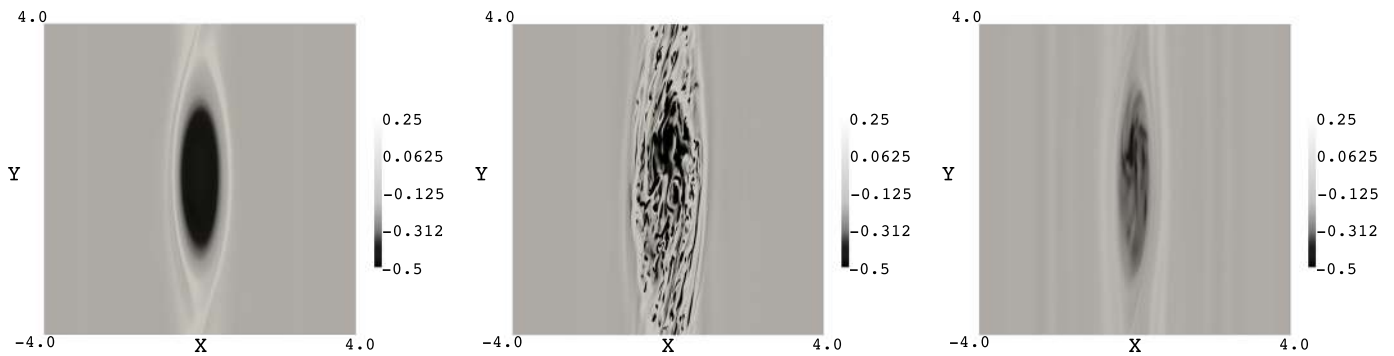
**Fig. 17.** Vorticity map from a simulation starting with a 2D Kida vortex plus a weak 3D noise. We show a snapshot at $t = 660$ (*left*), $t = 750$ (*middle*) and $t = 830$. After the 3D instability burst, a weak vortex survives and is amplified by the SBI.

We would like to stress that the present work constitutes only a *proof of concept* for the subcritical baroclinic instability. To claim that this instability actually exists in discs, one has to study the disc thermodynamic in a self consistent manner, a task which is well beyond the scope of this paper. Moreover, it is not possible at this stage to state that the SBI is a solution to the angular momentum transport problem in discs. Although this instability creates some transport through the generation of waves, this transport is weak and large uncertainties remain on its exact value. Finally, the coexistence of the SBI and the MRI (Balbus & Hawley 1991) is for the moment unclear. In the presence of magnetic fields, vortices produced by the SBI might become strongly unstable because of magnetoelliptic instabilities (Mizerski & Bajer 2009). Although the nonlinear outcome of these instabilities is unknown, one can already suspect that vortices produced by the SBI will be weaken in the presence of magnetic fields.

# References

Balbus, S. A. 2003, ARA&A, 41, 555
Balbus, S. A., & Hawley, J. F. 1991, ApJ, 376, 214
Barge, P., & Sommeria, J. 1995, A&A, 295, L1
Barranco, J. A., & Marcus, P. S. 2005, ApJ, 623, 1157
Chagelishvili, G. D., Zahn, J.-P., Tevzadze, A. G., & Lominadze, J. G. 2003, A&A, 402, 401
de Val-Borro, M., Edgar, R. G., Artymowicz, P., et al. 2006, MNRAS, 370, 529
de Val-Borro, M., Artymowicz, P., D'Angelo, G., & Peplinski, A. 2007, A&A, 471, 1043
Fricke, K. 1968, ZAp, 68, 317
Fromang, S., & Nelson, R. P. 2005, MNRAS, 364, L81
Fromang, S., & Papaloizou, J. 2006, A&A, 452, 751
Fromang, S., Papaloizou, J., Lesur, G., & Heinemann, T. 2007, A&A, 476, 1123
Goldreich, P., & Schubert, G. 1967, ApJ, 150, 571
Hawley, J. F., Gammie, C. F., & Balbus, S. A. 1995, ApJ, 440, 742
Hawley, J. F., Balbus, S. A., & Winters, W. F. 1999, ApJ, 518, 394
Heinemann, T., & Papaloizou, J. C. B. 2009a, MNRAS, 397, 52
Heinemann, T., & Papaloizou, J. C. B. 2009b, MNRAS, 397, 64
Ji, H., Burin, M., Schartman, E., & Goodman, J. 2006, Nature, 444, 343
Johnson, B. M., & Gammie, C. F. 2005a, ApJ, 626, 978
Johnson, B. M., & Gammie, C. F. 2005b, ApJ, 635, 149
Johnson, B. M., & Gammie, C. F. 2006, ApJ, 636, 63
Kida, S. 1981, Phys. Soc. Japan, 50, 3517
Klahr, H. 2004, ApJ, 606, 1070
Klahr, H. H., & Bodenheimer, P. 2003, ApJ, 582, 869
Kley, W., Papaloizou, J. C. B., & Lin, D. N. C. 1993, ApJ, 409, 739
Lesur, G., & Longaretti, P.-Y. 2005, A&A, 444, 25
Lesur, G., & Longaretti, P.-Y. 2007, MNRAS, 378, 1471
Lesur, G., & Papaloizou, J. C. B. 2009, A&A, 498, 1
Lovelace, R. V. E., Li, H., Colgate, S. A., & Nelson, A. F. 1999, ApJ, 513, 805
Mizerski, K. A., & Bajer, K. 2009, J. Fluid Mechanics, 632, 401
Papaloizou, J. C. B., Nelson, R. P., & Snellgrove, M. D. 2004, MNRAS, 350, 829
Petersen, M. R., Julien, K., & Stewart, G. R. 2007a, ApJ, 658, 1236
Petersen, M. R., Stewart, G. R., & Julien, K. 2007b, ApJ, 658, 1252
Regev, O., & Umurhan, O. M. 2008, A&A, 481, 21
Spiegel, E. A., & Veronis, G. 1960, ApJ, 131, 442
Umurhan, O. M., & Regev, O. 2004, A&A, 427, 855
von Weizsäcker, C. F. 1944, Z. Astrophys., 22, 319
Ziegler, U., & Yorke, H. W. 1997, Computer Phys. Commun., 101, 54