Possible planet-forming regions on submillimetre images

Zs. Regály,1⋆ A. Juhász,2 Zs. Sándor3 and C. P. Dullemond4

1Konkoly Observatory of the Hungarian Academy of Sciences, PO Box 67, H-1525 Budapest, Hungary
2Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands
3Max Planck Research Group, Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany
4Institut für Theoretische Astrophysik, Universität Heidelberg, D-69120 Heidelberg, Germany

Accepted 2011 September 14. Received 2011 September 14; in original form 2011 April 25

ABSTRACT
Submillimetre images of transition discs are expected to reflect the distribution of the optically thin dust. Former observation of three transition discs LkHα 330, SR 21N and HD 135344B at submillimetre wavelengths revealed images which cannot be modelled by a simple axisymmetric disc. We show that a large-scale anticyclonic vortex that develops where the viscosity has a large gradient (e.g. at the edge of the disc dead zone), might be accountable for these large-scale asymmetries. We modelled the long-term evolution of vortices being triggered by the Rossby wave instability. We found that a horseshoe-shaped (azimuthal wavenumber \( m = 1 \)) large-scale vortex forms by coalescing of smaller vortices within \( 5 \times 10^4 \) yr, and can survive on the disc lifetime \( (\sim 5 \times 10^6 \) yr), depending on the magnitude of global viscosity and the thickness of the viscosity gradient. The two-dimensional grid-based global disc simulations with local isothermal approximation and compressible-gas model have been done by the GPU version of hydrodynamic code FARGO (GFARGO). To calculate the dust continuum image at submillimetre wavelengths, we combined our hydrodynamical results with a three-dimensional radiative transfer code. By the striking similarities of the calculated and observed submillimetre images, we suggest that the three transition discs can be modelled by a disc possessing a large-scale vortex formed near the disc dead zone edge. Since the larger dust grains (larger than mm in size) are collected in these vortices, the non-axisymmetric submillimetre images of the above transition discs might be interpreted as active planet and planetesimal-forming regions situated far \( (\gtrsim 50 \) au) from the central stars.

Key words: accretion, accretion discs – hydrodynamics – instabilities – methods: numerical – protoplanetary discs.

1 INTRODUCTION

Recently Brown et al. (2009) presented submillimetre images of three transition discs LkHα 330, SR 21N and HD 135344B. Using a two-dimensional (2D) radiative transfer model (Dullemond & Dominik 2004), they showed that the observed inner discs are cleared from dust as the optical through millimetre wave spectral energy distribution (SED) can be fitted with a disc having an inner cavity. The non-axisymmetric intensity distributions (see fig. 1 of Brown et al. 2009), however, require further assumptions such as the presence of an embedded planet at large distances from the central star \( (\gtrsim 50 \) au) that sufficiently perturbs the disc dust distribution (see e.g. Wyatt et al. 1999). A massive enough planet opens a gap, and the disc’s material will be piled up just outside of the giant planet’s orbit (de Val-Borro et al. 2007). Although the giant planet induced non-axisymmetric density distribution hypothesis is challenging, we propose in this paper that the observed asymmetries are due to large observable vortices formed in these discs during their early evolutionary stages.

However, the current most accepted scenarios for giant planet formation, the core accretion (Bodenheimer & Pollack 1986; Pollack et al. 1996) and the gravitational instability (Boss 2001, 2006) face problems forming and keeping giant planets at such a large distances. Within the core-accretion scenario, the optimal places for planet formation are the vicinity of the snow line where the condensation of volatiles into ices increases the surface density of solids by a large factor, increasing the isolation mass above the critical core mass. The location of the snow line determined by the opacity of the dust grains and the mass accretion rate is recently found to be \( \lesssim 16 \) au (Min et al. 2011). For significantly larger distances, the growth rate of the planetary core is strongly...
limited. On the other hand, forming giant planets on wide orbits via gravitational instability does not face many problems (assuming rapid cooling of the gas), except that it is not yet clear how to keep them there. Dodson-Robinson et al. (2009) investigated three possible mechanisms – core accretion (with or without migration), scattering from the inner disc or gravitational instability – which might result in massive planets at wide (≥35 au) orbits discovered by high-contrast direct imaging observations of Fomalhaut (Kalas et al. 2008), HR 8799 (Marois et al. 2008) or CD 352722 (Walhaj et al. 2011) for instance. Dodson-Robinson et al. found that the only possible scenario to form stable wide-orbit planets is the gravitational instability, but planets formed this way were recently shown (by Baruteau, Meru & Paardekooper 2011) to migrate quite rapidly inwards – puts in question its feasibility. Note that Crida, Masset & Morbidelli (2009) have shown that runaway outward migration by planets in mean motion resonance carving a common gap in the protoplanetary disc is an other possible mechanism to place giant planets on wide orbits, if the inner one is significantly more massive than the outer one.

In recent years the formation of vortices in protoplanetary discs have been investigated thoroughly, since they may play an important role in the formation of planets in the core-accretion scenario. Observation of such vortices would have a great importance providing further information and support for planetary formation theories (see e.g. Barge & Sommeria 1995; Klahr & Henning 1997).

The classical approach of accretion discs (Shakura & Sunyaev 1973) postulates a simple scaling relation between the gas viscosity and the level of angular momentum transport parametrized by $\alpha$. An accretion disc coupled dynamically to a weak magnetic field is subject to the so-called magnetorotational instability (MRI) resulting in outward angular momentum transport, which can be described by an effective $\alpha$ viscosity (Balbus & Hawley 1991). The disc becomes MRI active, if the ionization state of the gas is sufficient. Investigating the ionization processes in MRI-active discs, Gammie (1996) proposed the existence of a radially extended region in the disc midplane, where the accretion is strongly reduced. Near the disc inner edge, the midplane gas is collisionally ionized, owing to the high temperatures (Pneuman & Mitchell 1965; Umebayashi 1983). Further out, the temperature is not high enough to collisionally ionize the gas, while the gas is yet dense enough to shield the interstellar cosmic rays or diluted stellar X-rays. The magnetic field of the ionized disc corona might even deflect the cosmic rays. As a result, only a thin surface layer of the disc is ionized at this region, while at even larger distances, the cosmic rays can penetrate the tenuous disc which might ionize the disc midplane again (Glassgold, Najita & Igea 1997; Igea & Glassgold 1999). As a consequence, the turbulent viscosity is low in the poorly ionized region of the disc midplane, resulting in the formation of the so-called dead zone. Within this region the viscous accretion is taking place mostly in a narrow upper layer of the disc, while it is stalled in the disc’s midplane.

Varnière & Tagger (2006) and Terquem (2008) demonstrated that density and pressure enhancements (‘bumps’) appear at the boundaries of a dead zone because of the sharp jump in the viscosity. These pressure bumps are found to be unstable to the Rossby wave instability (RWI; Lovelace et al. 1999; Li et al. 2000), leading to the formation of a large-scale anticyclonic vortices.

Kretke & Lin (2007) have already shown that pressure-gradient inversion formed at the inner edge of the dead zone or at the snow line causes the accumulation of dust. The existence of large-scale vortices can also accelerate the formation of planetesimals and planetary embryos in the core-accretion paradigm. Barge & Sommeria (1995) and Klahr & Henning (1997) have already shown that particles tend to get trapped in anticyclonic vortices. At the radial location of the vortex the type I migration is halted due to radial pressure-gradient inversion, thus not only dust is retained here but also larger bodies too. Particles of about centimetre to metre in size subsequently drift into the pressure maxima, i.e. to the vortices (Lyra et al. 2009; Dzyurkevich et al. 2010; Kato, Fujimoto & Ida 2010) and form gravitationally bound clumps of solids, which can coalesce forming embryos between the masses of the Moon and Mars. The swarm of these embryos evolves further by mutual N-body collisions forming massive cores (~10 $M_\odot$) of giant planets in ≤5 × 10$^3$ yr (Sándor, Lyra & Dullemen 2011). Thus, large-scale anticyclonic vortices formed in discs are extremely important for the planetary formation helping to overcome the metre-size barrier problem (Blum & Wurm 2008), and the time-scale problem of oligarchic growth. From this perspective these vortices can be regarded as ‘planetary factories’.

The micron-sized dust is dynamically coupled to the gas as it is indirectly supported against the settling by gas pressure via the gas drag, thus their spatial distribution is expected to be the same. Assuming that the three transition discs presented in Brown et al. (2009) contain significant amount of gas – supported by Hα (Fernandez et al. 1995), [O i] (van der Plas et al. 2008) and fundamental band CO emissions (Pontoppidan et al. 2008; Salyk et al. 2009) – these vortices can be imaged in submillimetre wavelengths as overdense regions in the dust. The inner holes in the submillimetre images of the three discs may indicate that dust is either consumed here by a planet-forming process or by coagulation to larger sized grains, which are drifted rapidly into the star. Note that coagulation without fragmentation would make the vortex also invisible as the dust opacity is decreasing with grain size. If the dust trapped inside giant vortices (Inaba & Barge 2006) is accountable for the observed features, fragmentation should be happening inside them, in order to replenish the population of small dust grains.

Inspired by these ideas, we explore, whether a large-scale, long-lived (on 10$^4$ yr time-scale) non-axisymmetric gas-dust perturbation formed at the outer edge of the disc dead zone could explain the non-axisymmetric submillimetre images. To calculate the density perturbation, we run long-term numerical hydrodynamical simulations on a GPU aided computer. In Section 2, we present our hydrodynamical disc model, and a study on the robustness of the formation of large-scale vortices. Submillimetre dust continuum images calculated by a three-dimensional (3D) radiative transfer using our hydrodynamical results are presented in Section 3. The paper closes (Section 4) with the discussion and concluding remarks, along with a short discussion on the reasonability of our findings.

2 VORTEX FORMATION

We revisit the development of RWI (Lovelace et al. 1999) by 2D, grid-based, global hydrodynamical disc simulations, assuming the $\alpha$-type viscosity prescription (Shakura & Sunyaev 1973). To model the dead zone, the value of $\alpha$ is smoothly reduced such that $\alpha(R) = \alpha_{\text{mod}}(R)$ and

$$
\delta_{\alpha}(R) = 1 - 0.5 (1 - \alpha_{\text{mod}}) \left[ -\tanh \left( \frac{R - R_{\text{dze}}}{\Delta R_{\text{dze}}} \right) + 1 \right].
$$

[Equation corrected in online version 2011 October 25.] where $\delta_{\alpha}$ is the depth of turbulent viscosity reduction, $R_{\text{dze}}$ and $\Delta R_{\text{dze}}$ are the radial distance and the width of the transition region of turbulent viscosity at the dead zone outer edge, respectively. Regarding the
present study the effect of the inner edge of the dead zone is not relevant, therefore, it is not investigated here. Examples of $\alpha(R)$ used in our models are shown in Fig. 1.

To quantify $R_{\text{dze}}$, we adopt the results of Matsumura & Pudritz (2005). According to their calculations, the outer edge of the dead zone lies between 12 and 36 au, depending on their various model parameters. To induce the vortex formation at distances suggested by submillimetre images of Brown et al. (2009), we set the outer edge of the dead zone to slightly larger distance ($R_{\text{dze}} = 50$ au).

Note that the exact location of the outer edge of the dead zone is very uncertain, because the physics of the ionization (depending on the ionization rate of the stellar X-rays, interstellar cosmic rays, proton radiation from the high-temperature disc corona and decay of radionuclides such as $^{40}$K or $^{26}$Al) and the dependence of the MRI on this ionization degree is not yet well understood (see e.g. Turner & Drake 2009).

2.1 Numerical simulations

To follow the evolution of the RWI triggered vortices on several $10^6$ yr time-scale, we used GFARGO, the GPU2 version of the FARGO code (Masset 2000). We slightly modified the CUDA kernel of GFARGO code in order to handle the viscosity reduction according to equation (1).

The dimensionless continuity and Navier–Stokes equations, describing the surface mass density evolution in an accretion disc, are numerically solved in 2D cylindrical coordinate system. We use the locally isothermal approximation, in which the vertically integrated gas pressure is $p(R) = c_s(R)^2 \Sigma(R)$, where $c_s(R)$ is the local sound speed, and $\Sigma(R)$ is the vertically integrated surface mass density of gas. The initial surface mass density profile follows the power-law distribution, $\Sigma(R) \sim R^{-0.5}$. We use radially constant disc aspect ratio $h = H(R)/R$, where $H(R)$ is the pressure scale height at the radial distance $R$, for which case $c_s(R) = \Omega H = \sqrt{GM_*/R^3(hR )} \propto R^{-1/2}$.

The disc self-gravity can be neglected as long as the Toomre parameter (Toomre 1964) is high, $Q = hM_*/(\pi R^2 \Sigma) \gg 1$, which holds in our unperturbed discs. Note, however, that $Q$ is relevant only for small scale, i.e. high $m$ modes, for global modes (e.g. $m = 1$) the parameter that might really measure the relevance of self-gravity is $Q_h$ (Papaloizou 2002). Since $Q_h$ might be smaller than unity, particularly inside the vortex owing to the density enhancement, we calculate $Q_h$ for the base model ($h = 0.05$ and $M_{\text{disc}} = 0.01 M_\odot$, and neglecting self-gravity) by the time when the large-scale vortex have fully developed (see Fig. 2). As one can see, $Q_h$ never reaches unity ($Q_h\text{\textsubscript{lim}} = 3.68$ inside the vortex eye), therefore, we neglected the disc self-gravity in our models. Note that the disc mass $M_{\text{disc}} = 0.01 M_\odot$ that we used in our simulations is in the range $(0.006 \leq M_{\text{disc}} \leq 0.026)$ as has been derived by SED modelling for the three transition discs by Andrews et al. (2011). Nevertheless, $Q_h$ might be smaller than unity for a disc which is three times more massive than the one we used, thus for that case one should take into account the self-gravity.

The disc inner and outer boundaries are fixed at 4 and 100 au, respectively. The computational domain contains 256 radial logarithmically and 512 azimuthal uniformly distributed grid cells. The numerical convergence with the above resolution is found to be satisfactory (see Section 2.3). At the inner and outer boundaries the so-called ‘damping boundary condition’ are applied (de Val-Borro et al. 2006), where a solid wall is assumed with wave killing zones next to the boundaries.

2.2 Development of a large-scale vortex

In what follows we give a short description of the formation and evolution of a large-scale vortex in our base disc model ($M_{\text{disc}} = 0.01 M_\odot$, $h = 0.05$, $\alpha = 0.01$, $\alpha_{\text{mod}} = 0.001$, $R_{\text{dze}} = 50$ au and $\Delta R_{\text{dze}} = 3$ au). The major evolutionary stages of the surface mass density distribution are shown in Fig. 3, while the azimuthally averaged surface mass density profiles are shown in Fig. 4.

To excite RWI, a small mass orbiting body (negligible to the local disc mass) was put in the disc at 50 au. The orbiting perturber excites low-$m$ mode density waves emerging from its Lindblad resonances, and these in turn trigger the RWI. Triggering RWI by this method is appropriate, because RWI is a low-$m$ mode instability. We observed, though, that for models where the RWI is strong – e.g. for sharp dead zone edge models ($\Delta R_{\text{dze}} = 1$ au), where the radial gradients of pressure and density are steep – the numerical noise in the radial velocity component can be large enough to trigger the vortex formation. This can be explained by the fact that the magnitude of

1 http://fargo.in2p3.fr/spip.php?rubrique18
2 Graphical Processor Unit constructed with several hundred processors (240 for GTX-280, which we used for the simulations).
3 http://www.nvidia.com/object/what_is_cuda_new.html

© 2011 The Authors, MNRAS 419, 1701–1712
Monthly Notices of the Royal Astronomical Society © 2011 RAS.

Figure 1. Example of smooth viscosity reduction according to equation (1) used in our models. $\delta_s(R)$ for models $\alpha = 0.01$, $\alpha_{\text{mod}} = 0.01$, $R_{\text{dze}} = 50$ au and $\Delta R_{\text{dze}} = 1, 3, 5$ au are shown with black solid, blue dashed and red dotted curves, respectively.

Figure 2. The distribution of $Q_h$ (where $Q$ is the Toomre parameter and $h$ is the disc aspect ratio) calculated for the base model, when a large-scale vortex is fully developed. Since $Q_h$ is well above unity, even inside the vortex eye, it seems reasonable to neglect the self-gravity.
Evolution of the azimuthally averaged surface mass density profile in the base model. The solid black curve represents the initial surface mass density profile. The density profiles shown with coloured dashed curves are calculated from the surface mass density distribution presented in Fig. 3.
\( m = 1 \) mode temporarily drops. When the \( m = 1 \) mode vortex decays, the spectral power of \( m = 4, 5 \) modes abruptly rises, i.e. RWI is re-excited. These newly formed vortices are also subject to coalescing and form again an \( m = 1 \) mode vortex by \( \sim 10^5 \) yr \((P_{50,1} = 300)\). Thus, the typical vortex mode is \( m = 1 \) through \( \sim 1.8 \times 10^6 \) yr \((P_{50,1} = 5200)\) of vortex evolution, which is interrupted by short-lived higher mode \((1 < m < 5)\) vortex configurations. Note that the spectral power in \( m = 0 \) characterizing the average surface mass density drops off immediately as RWI is excited.

As shown in Fig. 4, the density maximum is pushed inward from its original location, i.e. the \( m = 1 \) vortex is slightly shrunk due to the extra pressure caused by the density build-up. As a result, the fully developed \( m = 1 \) mode vortex forms at slightly smaller distance \((R \simeq 40 \text{ au})\) to the central star than the distance where the viscosity reduction \((R_{\text{dze}} = 50 \text{ au})\) begins. It is also appreciable that the density enhancement inside the vortex eye increases with time. Finally, emphasize that the surface mass density of the disc is significantly lowered next to the dead zone edge. This density rarefaction continuously strengthens with time, forming a deep hole in the disc by \( \sim 1.5 \times 10^6 \) yr \((P_{50,1} = 4250)\).

### 2.3 Robustness of large-scale vortex development

To examine the robustness (sensitivity to disc properties) of the development of the large-scale vortex, we performed several hydrodynamic simulations. By harnessing the power of the GPU, each simulation required only a day of computer time. For this parameter study the boundary conditions, the disc mass, and the disc aspect ratio were the same as in the previously shown base model. In what follows we summarize briefly the results of these simulations.

#### 2.3.1 Numerical resolution

To check whether our numerical solution is satisfactory with the original grid \((N_R = 256, N_\phi = 512)\), we recalculated the base model with higher resolution, namely \(N_R = 512, N_\phi = 1024\). We found that the results are practically identical, even by frame-to-frame comparison. Hence, the spatial resolution of the base model \((256 \times 512)\) is sufficient for the following simulations.

#### 2.3.2 Numerical precision

Scientific calculations with conventional GPU cards (e.g. NVIDIA GTX-280 that we used) lack the usual double precision at the moment.\(^4\) To check the validity of our results provided by the single precision GPU we compare them to those of given by the CPU-based FARGO code in the fastest evolving model \((\alpha = 0.1, R_{\text{dze}} = 1 \text{ au}, \alpha_{\text{mod}} = 0.001)\). The two runs give qualitatively the same result: the evolution of the large-scale vortex (growth rate, vortex coalescing, wing interlock and oscillation) is found to be practically the same (see the details in Appendix A).

#### 2.3.3 Dead zone edge width

Li et al. (2000) have found that the excitation of the RWI occurs later in time if the transition region of the viscosity reduction (i.e. the width of the dead zone edge) is wider. We confirmed this by setting

\[^4\] Note that there exist scientific grade graphics cards in the NVIDIA Tesla series which are capable of double precision calculations.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig6.png}
\caption{Fully developed large-scale vortices at \(3.5 \times 10^5 \) yr \((P_{50,1} = 1000)\) in models using \(R_{\text{dze}} = 1, 2 \text{ and } 3 \text{ au dead zone edge width shown in panels A, B and C, respectively. The axisymmetric } m = 0 \text{ configuration for model } R_{\text{dze}} = 5 \text{ au shown in panel D is just the result of the viscosity reduction at the dead zone edge stalling the mass inflow, with subsequent pileup of gas.}\}
\end{figure}
also found to be independent of $\alpha_{\text{mod}}$: for sharp dead zone edge models the vortex mode oscillation do occur, while for thick dead zone edge models the $m = 1$ mode vortex is long lived.

### 2.3.5 Magnitude of global viscosity

We found that the vortex evolution is sensitive to the magnitude of the global viscosity $\alpha$ being in the range of 0.001–0.1. The time required to excite RWI and to form the fully developed large-scale $m = 1$ mode vortex decreases with increasing $\alpha$. Both the radial extension of the large-scale vortex and the density contrast in the vortex increase with $\alpha$. These can be interpreted by the accelerated vortex evolution in models with high global viscosity.

For high-viscosity models ($\alpha = 0.1$), the inner disc becomes ever denser during the later ($t \gg 5 \times 10^6$ yr) epochs, and the density contrast drops between the large-scale vortex and the inner disc. In consequence, the $m = 1$ mode vortex survives only up to $\sim 10^6$ yr in these models. For low-viscosity discs ($\alpha = 0.001$), the vortex evolution slows down considerably, thus the excitation of RWI occurs later and the lifetime of $m = 1$ mode vortex can reach $\sim 5 \times 10^6$ yr.

We found that for low-viscosity discs the vortex mode oscillation is completely missing, even for sharp dead zone edge ($\Delta R_{\text{dze}} = 1 \text{ au}$) models. Contrary to this, for high-viscosity discs, the vortex mode oscillation does occur even for thick dead zones ($\Delta R_{\text{dze}} < 5 \text{ au}$) models. Emphasize that the $m = 1$ mode (horseshoe-shaped) vortex might survive up to the lifetime of the disc for $\alpha = 0.001$, independently of the dead zone edge width being smaller than the twice of the local disc height ($\Delta R_{\text{dze}} \leq 5 \text{ au}$).

### 3 Submillimetre images

In order to show the resemblance of the observations of Brown et al. (2009) and signatures of large-scale Rossby vortices in protoplanetary discs, we calculated submillimetre images using the 3D radiative transfer code RADMC-3D. First, the temperature structure of the disc was determined in a Monte Carlo simulation then the images at 340 GHz (880 $\mu$m) were calculated via ray tracing. Submillimetre Array (SMA) observations were simulated in the Common Astronomy Software Applications (CASA) package.

Snapshots of surface mass density distributions shown in Fig. 9 from the hydrodynamical simulations were converted into volume density assuming a vertically Gaussian density distribution with the same constant aspect ratio ($h = 0.05$) as used in the hydrodynamical simulations. The geometrical dimensions of the sources as well as their distances were adapted from Andrews et al. (2011). The hydrodynamical frames had to be rotated and spatially scaled to approximately match the disc shapes and sizes on the measured submillimetre images. The used rotation angles were 330°, 320° and 170° clockwise, while the linear scaling factors were 2, 1.2 and 1.7 for LkHα 330, SR 21N and HD 135344B, respectively. Since there are evidences for a fairly large gap in all discs (Brown et al. 2009), we reduced the dust density within a certain radius according to Andrews et al. (2011). Inside the dead zone, dust grains may grow and drift rapidly inwards leaving a reduced density region, i.e. a gap, behind. In anticyclonic vortices, however, dust grains get trapped and cannot drift inwards, which would qualitatively explain why the outer radius of the gap is close to the location of the RWI. Note that there is also evidence for the presence of an inner disc.

5 http://www.ita.uni-heidelberg.de/~dullemond/software/radmc-3d
6 http://casa.nrao.edu/index.shtml

---

**Figure 7.** Azimuthally averaged surface mass density profiles at $3.5 \times 10^5$ yr ($P_{30.1} = 1000$) in models using $\Delta R = 1, 2, 3$ and 5 au dead zone edge width and $\alpha_{\text{mod}} = 0.001$ (upper panel). Models with lower viscosity reduction $\alpha_{\text{mod}} = 0.01$ (lower panel) show no significant departures to previous simulations.

**Figure 8.** Azimuthal spectral power of the surface mass density distribution against time measured in $P_{30.1}$ for model $\alpha = 0.01$, $\Delta R_{\text{dze}} = 3 \text{ au}$, $\delta \alpha = 0.01$. The Fourier amplitudes are normalized by the total power measured in the $m = 0$–6 modes.
in these sources (Fernandez et al. 1995; Pontoppidan et al. 2008; van der Plas et al. 2008; Salyk et al. 2009). Since these inner discs have an outer radius probably smaller than the inner radius of our hydrodynamic simulations (4 au) we neglected them.

In the radiative transfer simulations we used the hydrodynamic grid in the radial and azimuthal directions while \( N_\theta = 100 \) grid points were used in the \( \theta \) direction. The polar grid was a uniformly distributed, but we split the \( -\pi/2 \leq \theta \leq +\pi/2 \) domain into two regions in order to resolve the disc vertically, while keeping the simulation computationally feasible. We placed 80 grid points in the \( |\theta| \leq 0.35 \) domain and 20 grid points were placed in the \( |\theta| \geq 0.35 \) region. In the thermal Monte Carlo simulations \( 2 \times 10^7 \) photons were used to determine the temperature structure.

The dust in our model consisted of ‘astronomical’ silicate grains (Weingartner & Draine 2001), whose absorption and scattering cross-sections were calculated using Mie theory from the optical constants. Dust grains had an MRN (Mathis, Rumpl & Nordsieck 1977) size distribution, \( n(a) \propto a^{-3.5} \), sampled with 10 logarithmic spaced bins with minimum and maximum grain size of 0.1 and 1000 \( \mu \)m, respectively. Dust particles with different sizes have different temperatures, thus they are not thermally coupled to each other. The dust-to-gas ratio was assumed to be 0.01 and constant within the disc outwards of the outer radius of the gap. Inside the gap the dust-to-gas ratio should be lower than 0.01 if the gap is formed due to the growth and inward drift of dust grains. Finally, the simulation of the observation (Fourier transformation, \( u-v \) sampling, cleaning) was done with CASA. The resulting images are shown in Fig. 10. We also calculated SED from our synthetic disc model for all three sources and compared them to the observations (see Fig. 11).

Our synthetic model images (Fig. 10) resemble the most important features of the submillimetre continuum map of the three transition discs presented in Brown et al. (2009). The disc emission is clearly non-axisymmetric, horseshoe shaped, and it shows single arc-like maximum at position angle around 225° for LkHα 330, 120° for HD 135344B, while two maxima are visible at 120° and 270° position angles for SR 21N. The SEDs (Fig. 11) calculated from our disc models match the observations shortwards of about 1 \( \mu \)m and longwards of about 10 \( \mu \)m. Between 1 and 10 \( \mu \)m, however, our models underpredict the observed fluxes, which is a natural consequence of neglecting the inner disc emission. We concluded that our disc models with large-scale Rossby vortices fit the observations of the mid-infrared to submillimetre emitting regions of all three discs.

Given the fact that our model is vertically optically thin at submillimetre wavelengths the observed asymmetry in the continuum emission can arise from temperature, density perturbations or the combination of both. Since our disc is optically thin at 880 \( \mu \)m the increase in the column density will increase the observed flux. The higher density regions, however, can be shielded from the heating radiation (higher optical depth) and therefore the temperature can decrease. The temperature perturbations in our model are less than a factor of 2, which means the same factor for the flux perturbation caused by temperature effects, since we are in the Rayleigh–Jeans regime. The density perturbations on the other hand are much larger in amplitude (factor of 6–8, see e.g. the red dashed curve in Fig. 3) and are therefore most likely responsible for the observed asymmetries in the image. A detailed study of these effects in non-axisymmetric discs will be the topic of a forthcoming paper.

4 DISCUSSION AND CONCLUSION

In this paper we propose an interesting physical phenomenon that can be responsible for the large non-axisymmetric density perturbations revealed by submillimetre images of three transition discs of Brown et al. (2009). Based on the results of our long-term, two-dimensional, grid-based, global hydrodynamic simulations (using compressible-gas model, neglecting the thermodynamical effects and the disc self-gravity), we suggest that the observed asymmetries are caused by a large-scale vortex developed where the turbulent viscosity has a steep gradient e.g. at the outer edge of the disc dead zone. Vortices are triggered by the RWI, which later are subject to a coalescing process and form a large-scale anticyclonic vortex. Note that Wolf & Klahr (2002) have previously shown that such large-scale anticyclonic vortices formed in the inner disc (\(< 10 \) au) can cause detectable intensity asymmetry in millimetre wavelengths (345 and 900 GHz) which are proposed to be observed by the fully operable Atacama Large Millimeter/submillimeter Array (ALMA) with the longest available baseline.

After the initial excitation of RWI, a large-scale anticyclonic vortex develops by the coalescing of small-scale vortices at \( R \approx 50 \) au, near the outer edge of disc dead zone (Figs 3 and 4), which...
and the magnitude of global viscosity (the larger the viscosity, the faster the vortex evolution). The $m = 1$ mode vortex temporarily disappears and reforms through RWI re-excitation and subsequent vortex coalescing processes (Fig. 5) with a cadence of $\sim 10^5$ yr ($P_{50,1} = 300$) for sharp dead zone edge models ($\Delta R_{dze} < 2$ au) and large global viscosity ($\alpha \geq 0.01$). However, the $m = 1$ mode vortex remains stabilized for $\sim 5 \times 10^6$ yr ($P_{50,1} = 14,000$) for thick dead zone edge models ($2 \leq \Delta R_{dze} < 5$ au) or small global viscosity ($\alpha \leq 0.01$). For even thicker dead zone edge models ($\Delta R_{dze} \geq 5$ au), $m = 1$ mode vortex does not form, only $m = 0$ density perturbation develops. In short, we conclude that the horseshoe-shaped vortex ($m = 1$ mode) formed due to the viscosity jump near the disc dead zone edge being at $\sim 50$ au can survive up to the disc lifetime ($t \geq 5 \times 10^6$ yr), if the turbulent viscosity is small ($\alpha = 0.001$), and the width of the dead zone edge is smaller than twice of the local disc height ($\Delta R_{dze} \ll 2H$).

To demonstrate the observability of the large vortices, we calculated synthetic submillimetre images with a 3D radiative transfer code using the surface mass density distribution provided by our hydrodynamical simulations. The distribution of the dust, emitting in the submillimetre wavelengths, is assumed to follow that of the gas, owing to the dust collecting effect of the pressure bump. The submillimetre intensity distributions given by our ray-tracing models – fitting the observed SEDs in the stellar and submillimetre emitting domain (Fig. 11) – show significant azimuthal asymmetries (Fig. 10). Comparing our submillimetre predictions to the observations of three transition discs LkHα 333, SR 21N and HD 135344B (Brown et al. 2009), we found astonishing similarities. Note that there exist further examples of discs showing non-axisymmetric submillimetre images, such as LkCa15 (Piétu et al. 2006), GM Aur (Hughes et al. 2009), SR 24S (Andrews et al. 2010) and J1604−2130 (Williams & Cieza 2011) for which the same phenomenon might be responsible. Since the formation process of planetesimals and planetary embryos could be accelerated by the dust collecting and embryo trapping effect of the anticyclonic vortices caused by the vortical motions and the pressure bump (Barge & Sommeria 1995; Klahr & Henning 1997), the non-axisymmetric submillimetre images presented by Brown et al. (2009) might be interpreted as snapshots of active planet-forming regions.

Note that there are other possible routes to generate vortices in protoplanetary discs, which also might result in $m = 1$ vortex formation. For the presence of gap-carving giant planets the gas is expelled from the planet orbit and settles at the gap edges resulting in density maximum which is RWI unstable and develops into vortices as shown by de Val-Borro et al. (2007). Another way to generate vortices is the baroclinic instability, proposed by Klahr & Bodenheimer (2003), and recently confirmed by several works (e.g. Petersen, Julien & Stewart 2007a; Petersen, Stewart & Julien 2007b; Lesur & Papaloizou 2010; Lyra & Klahr 2011).

As it was mentioned before, not only enhanced dust coagulation, fragmentation should also happen inside the vortices if the observed non-axisymmetric features of submillimetre images are indeed due to dust trapped inside them. Otherwise, the submillimetre intensity would be weak due to the decreasing dust opacity of large grains. Thus, this would constitute indirect evidence of elliptic instability (Lesur & Papaloizou 2009) inside vortices leading to 3D turbulence in the vortex core, which Lyra & Klahr (2011) quantified in compressible simulations to be at the level of 10 per cent of the sound speed.

Here we have to mention some of the caveats of our calculations. 2D and 3D turbulence behave quite differently regarding the energy cascade, which is form small to large $m$ (direct cascader) in 3D but grows both in size and density contrast. The time required to form a $m = 1$ mode large-scale vortex is found to be $\sim 1.5−4 \times 10^4$ yr ($P_{50,1} = 40−115$), depending on the dead zone edge width $\Delta R_{dze}$ (the sharper the dead zone edge, the faster the vortex evolution).
can be form large to small $m$ (inverse cascade) in 2D. 3D global disc simulations of Meheut et al. (2010) revealed that although the RWI results in significant vertical streaming, the growth rate and evolution of vortices in 3D is very similar to that of in 2D. This may indicate that the vorticity generation by the RWI is strong enough to counter the 3D vortex stretching (the mechanism responsible for the direct cascade), so that the vortices survive and cascade inversely via viscous merging.

Our simulations cover several million years of the disc evolution, yet, the disc photoevaporation acting on the same time-scale was neglected for simplicity. Since photoevaporation process disperse the gas content of the disc, this might cause inward migration of the dead zone edge. Thus, photoevaporation process might play an important role in vortex evolution, which will be addressed in a forthcoming paper.

Lin & Papaloizou (2011) have investigated the role of self-gravity on vortex instabilities generated at the outer edge of the gap opened by a giant planet in locally isothermal disc models. They found that the unstable modes shifts to higher azimuthal wavenumbers assuming self-gravity: the wavenumber of the most unstable mode increases as $Q$ decreases. For weak self-gravity ($M_{\text{disc}} \lesssim 0.024 M_\odot$), a single vortex forms through coalescing, which process is delayed as $Q$ decreases. However, for higher disc mass ($M_{\text{disc}} = 0.031 M_\odot$) the final configuration is a vortex pair, $m = 2$ mode. Since we used $M_{\text{disc}} = 0.01 M_\odot$, our models are in the weak self-gravity regime, where the coalescing process leads to the formation of $m = 1$ vortex.

Self-gravitating disc simulations require the inclusion of thermodynamics (the solution of energy equation) too (Pickett et al. 1998, 2000; Gammie 2001), otherwise the gas concentration would be overpredicted. Lyra et al. (2009) found that assuming adiabatic equation of state for the gas, the pressure bump (that gives rise to RWI) sharpens considerably compared to their locally isothermal models due to the high temperatures associated with compression. As a result, RWI might be triggered easier in adiabatic than in isothermal model. Petersen et al. (2007b) investigated the baroclinic feedback, in which a vortex enhances azimuthal temperature gradients to reinforce the vortex itself, in global disc simulations with an elastic-gas model. They concluded that this baroclinic feedback mechanism producing stronger local temperature gradients results in long-lived vortices if the radial temperature gradient is sufficiently large and either the radiative cooling time is small, or the thermal dissipation is high. On the other hand, vortices in the course of vortex coalescing may grow massive enough and reach the Jeans length, which provides a natural barrier to growth as shown by Mamatsashvili & Rice (2009). In their shearing sheet simulations, assuming a simple cooling law with a constant cooling time and including self-gravity, the vortex starts to shock and radiate away the excess vorticity around the Jeans length. As a result, the vortices cannot grow above the Jeans length and short lived, if there is no continuous vorticity production like for RWI, for which case the shear is continuously being converted into vorticity (Lyra et al. 2009). In our simulations, however, the vorticity is not generated by the local advective heat transport enhancing the azimuthal temperature gradients (see e.g. Petersen et al. 2007b), but the steep pressure bump and azimuthal density perturbations formed near the dead zone edge. In short, investigating the vortex evolution and longevity in a self-gravitating global disc model, considering non-isothermal and compressible-gas model, would be important, and will be addressed in a forthcoming paper.

Finally, worth noting that multiple viscosity jumps might exist in the disc e.g. at the inner edge of the dead zone or at the snow line (Kretke & Lin 2007), which could also result in the formation of large-scale vortices. In a non-magnetized disc its self-gravity causes outward angular momentum transport via quasi-steady self-gravitating turbulence, which results in the generation of effective viscosity (Pringle 1981). Since only the outer disc ($R \gtrsim 100$ au) is expected to become gravitationally unstable in a standard protoplanetary discs, the self-gravity caused effective viscosity has a gradient there. In such systems the large-scale vortex formed beyond $\sim 100$ au might result in vortex aided planet formation process at extremely high distances, such as observed around Formalhaut (Kalas et al. 2008). As a conclusion, planets might be assembled in diverse fashion by these ‘planetary factories’ located at different distances to their host star. In protoplanetary discs where the RWI is excited at smaller radii than we investigated, the development of large-scale vortices might lead to the formation of a planetary system similar to the Solar system.

ACKNOWLEDGMENTS

This research has been supported in part by DAAD-PPP mobility grant P-MOB/841/ and ‘Lendület’ Young Researcher Program of the HAS. We are grateful to F. Masset who helped us making some necessary modification in his GPARGO code. We thank M. R. Hogerheijde and Ch. Brinch for the help and discussion on CASA data processing. We also thank the referee Wladimir Lyra for thoughtful comments that helped to significantly improve the quality of the paper.

REFERENCES

Andrews S. M., Wilner D. J., Hughes A. M., Qi C., Dullemond C. P., 2010, ApJ, 723, 1241
Andrews S. M., Wilner D. J., Espaillat C., Hughes A. M., Dullemond C. P., McClure M. K., Qi C., Brown J. M., 2011, ApJ, 732, 42
Balbus S. A., Hawley J. F., 1991, ApJ, 376, 214
Barge P., Sommeria J., 1995, A&A, 295, L1
APPENDIX A: ON THE NUMERICAL PRECISION

As mentioned in Section 2.3, the GPU version of FARGO code (GFARGO) is only capable of single precision calculation at the

Figure A1. Surface mass density evolution in a fast evolving model (i.e. model with $\alpha = 0.1$, $\alpha_{mod} = 0.001$, $R_{dze} = 50$ au and $\Delta R_{dze} = 1$ au). The simulations were done by single precision GFARGO code (left-hand column) and double precision FARGO code (right-hand column) with exactly the same boundary conditions and numerical resolution. The surface mass density (shown with vertical colour bars for each frame) is measured in g cm$^{-2}$. The frames were saved at exactly the same time-steps for both simulations. $\mathcal{P}_{50\,,\,1}$ means the number of orbits at 50 au assuming a solar mass central star.
moment, which obviously results in the degradation of precision. In order to be convinced that the results of GFARGO (single precision) are appropriate, we ran additional simulations by FARGO (double precision) for the model with $\alpha = 0.1$, $\alpha_{\text{mod}} = 0.001$, $R_{\text{dze}} = 50$ au and $\Delta R_{\text{dze}} = 1$ au. In these simulations we applied the same boundary conditions, numerical resolutions ($N_\phi = 256$, $N_\theta = 512$) and disc masses for GFARGO and FARGO. The surface mass density distributions plotted at identical time-steps for both simulations are shown in Fig. A1.

During both simulations, the RWI is excited with $m = 6$ azimuthal wavenumber at $3 \times 10^3$ yr ($P_{50.1} = 8$). The subsequent vortex coalescing and vortex growth are qualitatively equivalent. However, the azimuthal angle of the $m = 1$ mode vortices at $10^5$ yr ($P_{50.1} = 300$) and $3 \times 10^5$ yr ($P_{50.1} = 900$) differs for the single and double precision simulations. This is because of the fact that $m = 1$ mode vortex slowly drifts inward (see e.g. Fig. 4) due to the extra pressure caused by the density enhancement at the dead zone edge, causing prograde precession of the vortex. Since the single precision simulation results in slightly higher density peak than the double precision simulation (by $\sim 5$ per cent higher in average), the vortex drifts faster for single precision simulation. Consequently, the azimuthal angles of the vortices are different.

In order to investigate the vortex mode oscillation presented in Section 2.2, we calculated the azimuthal spectral power of the surface mass density distribution against time (Fig. A2). As one can see, the evolution of the spectral power in the $0 \leq m \leq 6$ modes is very similar for the two simulations. Note, however, that although the first and second vortex decay occur at the same time for both simulations ($P_{50.1} \approx 600$ and $\approx 1000$, respectively), the time required to reform the stable $m = 1$ mode is slightly longer for the double precision simulation.

In short, we conclude that the lack of double precision capability of GFARGO is not dangerous on a million year time-scale in our simulations.

Figure A2. Azimuthal spectral power of the surface mass density distribution against time measured in $P_{50.1}$ for simulations shown in Fig. A1. Results given by GFARGO and FARGO are shown in the upper and lower panels, respectively. The Fourier amplitudes are normalized by the total power measured in the $m = 0$–6 modes.
Appendix B: On the Excitation of RWI

In order to make sure that no important feature is being missed during the excitation of the RWI, we ran several models where the RWI was seeded by conventional methods. To provide seeds for the RWI, we added low-level perturbations to the surface mass density when the threshold for the RWI had been reached (i.e. when significant density bump had been developed at the dead zone edge).

In the first series of simulations, the perturbation was calculated as a superposition of sines, similarly to equation (10) presented in Meheut et al. (2010):

\[
\Sigma(R, \phi) = \Sigma(R, \phi) + |\Sigma(R, \phi)| \epsilon \sin(2\pi R) \\
\times [\sin(\phi) + \sin(2\phi) + \sin(3\phi) + \sin(5\phi)],
\]

(B1)

where \(\Sigma(R, \phi)\) and \(\Sigma(R, \phi)\) are the unperturbed and perturbed surface mass density, respectively. The magnitude of the perturbation is set by \(\epsilon\). In the second series, the surface mass density was perturbed by normal distribution white noise. While the former type of perturbation has a drawback that it excites only the low-\(m\) modes as does a small-mass orbiting planet, large part of the spectrum is being excited by the latter type of perturbation. We also tested the RWI excitation by small-mass planet orbiting at 50 au.

The RWI excitation was investigated in models where parameters of \(\alpha = 0.01, \alpha_{\text{mod}} = 0.01, R_{\text{dz}} = 50\) au and \(\Delta R = 1\) au were chosen, while the grid resolution was set to \(N_R = 256\) and \(N_\phi = 512\). The magnitude of perturbations were in the order of \(\epsilon = 10^{-4}\). The simulations were done by \textsc{fargo}, because it has smaller numerical noise in the radial velocity component than in \textsc{gefargo} simulations.

First, the threshold for the RWI was let to develop, i.e. a significant radial density gradient with magnitude of \(\Delta \Sigma/\Sigma \sim 2\) at the dead zone edge. Then one of the above-described perturbations was applied. To ascertain whether the RWI is not excited by the numerical noise of \textsc{fargo}, we ran models without perturbation. In this unperturbed model, we found no RWI excitation.

Fig. B1 shows the evolution of azimuthal spectral power of the surface mass density followed as long as \(m = 1\) mode vortex had been fully fledged (by \(P_{50,1} = 150\)) for models using different type of perturbations: superposition of sines in surface mass density with magnitude of \(\epsilon = 10^{-4}\) (upper panel); normal distribution white noise in surface mass density with magnitude of \(\epsilon = 10^{-4}\) (middle panel) and small mass (\(M_{pl} = 10^{-7}M_*\)) planet orbiting at 50 au (lower panel). As one can see the evolutions of spectral power are very similar for any type of perturbations. Independent of the perturbation type applied, we found that (i) the RWI is always triggered after \(P_{50,1} \approx 5–10\) orbits at 50 au; (ii) the most unstable mode of RWI is \(m = 5\) or \(6\) at the onset of vortex formation; (iii) following the evolution of the vortices, \(m = 1\) mode (horseshoe-shaped) vortex always forms. Consequently, the mode of the excitation of RWI has no importance regarding our main findings that long-lived \(m = 1\) mode vortex appears near the disc dead zone edge.

This paper has been typeset from a \TeX/\LaTeX file prepared by the author.