

# Scattered light images of spiral arms in marginally gravitationally unstable discs with an embedded planet

A. Pohl,<sup>1,2★</sup> P. Pinilla,<sup>3★</sup> M. Benisty,<sup>4</sup> S. Ataiee,<sup>5,6</sup> A. Juhász,<sup>7</sup> C. P. Dullemond,<sup>2</sup>  
R. Van Boekel<sup>1</sup> and T. Henning<sup>1</sup>

<sup>1</sup>Max-Planck-Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany

<sup>2</sup>Institute of Theoretical Astrophysics, Heidelberg University, Albert-Ueberle-Str. 2, D-69120 Heidelberg, Germany

<sup>3</sup>Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands

<sup>4</sup>University Grenoble Alpes, IPAG, F-38000 Grenoble, France; CNRS, IPAG, F-38000 Grenoble, France

<sup>5</sup>School of Astronomy, Institute for Research in Fundamental Sciences (IPM), PO Box 19395-5531, 1956836613 Tehran, Iran

<sup>6</sup>Physics Institute, Space Research and Planetary Sciences, Sidlerstrasse 5, CH-3012 Bern, Switzerland

<sup>7</sup>Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

Accepted 2015 July 28. Received 2015 July 17; in original form 2015 June 9

## ABSTRACT

Scattered light images of transition discs in the near-infrared often show non-axisymmetric structures in the form of wide-open spiral arms in addition to their characteristic low-opacity inner gap region. We study self-gravitating discs and investigate the influence of gravitational instability on the shape and contrast of spiral arms induced by planet–disc interactions. Two-dimensional non-isothermal hydrodynamical simulations including viscous heating and a cooling prescription are combined with three-dimensional dust continuum radiative transfer models for direct comparison to observations. We find that the resulting contrast between the spirals and the surrounding disc in scattered light is by far higher for pressure scaleheight variations, i.e. thermal perturbations, than for pure surface density variations. Self-gravity effects suppress any vortex modes and tend to reduce the opening angle of planet-induced spirals, making them more tightly wound. If the disc is only marginally gravitationally stable with a Toomre parameter around unity, an embedded massive planet (planet-to-star mass ratio of  $10^{-2}$ ) can trigger gravitational instability in the outer disc. The spirals created by this instability and the density waves launched by the planet can overlap resulting in large-scale, more open spiral arms in the outer disc. The contrast of these spirals is well above the detection limit of current telescopes.

**Key words:** radiative transfer – scattering – planet–disc interactions – protoplanetary discs – circumstellar matter.

## 1 INTRODUCTION

Young forming protoplanets may leave observational signatures in their parental disc in the form of distinct gaps, vortices, warps or spiral arms. This is thought to happen during the final stages of protoplanetary disc evolution. There is a peculiar group called transition discs, whose spectral energy distribution and sub-millimetre (sub-mm) observations suggest that the inner disc is strongly depleted of dust. Spiral arms have thus far been only observed in this kind of discs, suggesting that the gap formation mechanism is related to spiral arm formation. This needs, however, to be treated carefully since there might be an observational selection effect. Observations in the near-infrared (NIR) and (sub-)mm regime of transition discs

show a variety of asymmetric features. At sub-mm wavelengths, asymmetries are seen for example in the system Oph IRS 48 (van der Marel et al. 2013), SR 21, HD 135344B/SAO 206462 (Pérez et al. 2014) and HD 142527 (Casassus et al. 2013; Fukagawa et al. 2013); for the latter a low-mass stellar companion has been found (Biller et al. 2012; Close et al. 2014). Since discs are optically thick in the optical and NIR, high angular resolution scattered light images trace small hot dust particles in the upper disc layer. Most discs have double spirals with a large opening angle that are almost symmetric, e.g. MWC 758 (Grady et al. 2013; Benisty et al. 2015), LkCa 15 (Thalmann et al. 2014), HD 135344B/SAO 206462 (Muto et al. 2012; Garufi et al. 2013), HD142527 (Casassus et al. 2012; Rameau et al. 2012; Canovas et al. 2013; Avenhaus et al. 2014a) and HD100546 (Boccaletti et al. 2013; Avenhaus et al. 2014b).

The origin of the observed spirals is still debated. For instance, with a simple analytical description based on the spiral density wave

\* E-mail: pohl@mpia.de (AP); pinilla@strw.leidenuniv.nl (PP)

theory, the morphology of spirals triggered by a hypothetical planet has been studied by Muto et al. (2012), Grady et al. (2013) and Benisty et al. (2015). The best spiral fitting model requires a high disc aspect ratio  $h = H/r$  to account for the large pitch angles (angle between the spiral arm and the tangent circle) implying a very warm disc. Recently, Juhász et al. (2015) focused on the contrast of spirals compared to their surrounding background disc. They modelled the spiral waves launched by planets by means of locally isothermal hydrodynamic simulations as well as analytic descriptions. Based on contrast arguments, these authors suggested that the spiral arms observed are the results of pressure scaleheight perturbations rather than of pure surface density perturbations. Dong, Zhu & Whitney (2014) also studied the observational signatures in transition discs by combining two-dimensional two-fluid hydrodynamical calculations with three-dimensional radiative transfer simulations, but rather focused on the observational signatures of gaps opened by one or several planets. They stated that density waves and streamers inside the planetary gap, produced by planet–disc interactions, can be visible in NIR images. Besides processes involving planets, it has been examined whether other mechanisms can lead to spiral features, such as non-ideal magnetohydrodynamical effects (e.g. Flock et al. 2015; Lyra, Turner & McNally 2015) or gravitational instability (e.g. Lodato & Rice 2004, 2005; Rice et al. 2004). Structures induced by the onset of gravitational instabilities tend to produce spirals with higher than  $m = 2$  azimuthal wavenumber (e.g. Cossins, Lodato & Clarke 2009; Forgan et al. 2011). Planets drive, depending on the planet-to-star mass ratio, an  $m = 1$  or sometimes even an  $m = 2$  mode, but the secondary spiral is substantially weaker. Our idea is to combine both spiral formation scenarios in order to see if this strengthens their amplitude and to produce various spiral morphologies.

In general, the precondition for a disc to become unstable is that the amount of gravitational potential energy overcomes pressure and rotational kinetic support. The axisymmetric stability of a thin disc is determined by the so-called Toomre parameter  $Q$  (Toomre 1964). The Toomre criterion for a disc to be unstable is defined as

$$Q = \frac{c_s \kappa}{\pi G \Sigma_g} \lesssim 1, \quad (1)$$

where  $c_s$  is the sound speed,  $\Sigma_g$  is the gas surface density and  $\kappa$  corresponds to the epicyclic frequency, which is equal to the angular velocity  $\Omega = \sqrt{GM_*/r^3}$  in the case of a Keplerian disc. Increasing the disc mass and/or decreasing the disc temperature leads to a lower  $Q$ . Hence, the disc mass has to be a significant fraction of the stellar mass for the disc to be gravitationally unstable.

Current observational methods to estimate the disc mass still have systematic uncertainties. The total disc mass is dominated by the gas, which mostly consists of  $H_2$ . Since its emission is hard to detect, there is only one object (TW Hya) for which hydrogen was directly observed in the form of HD (Bergin et al. 2013). Two prominent proxies, dust continuum emission and CO and its isotopologues, are usually used to estimate disc gas masses. The first method is based on observations of the mm continuum emission from dust grains (e.g. Beckwith et al. 1990; Andrews & Williams 2005), assuming a certain dust opacity and gas-to-dust ratio. The generally assumed value of 100 for the gas-to-dust ratio may not be accurate for all discs, since it strongly depends on the disc evolution (e.g. Brauer et al. 2007; Birnstiel, Dullemond & Brauer 2010). Also, from observations it is known that mm-sized dust grains and the gas do not necessarily have the same spatial distribution in discs (e.g. de Gregorio-Monsalvo et al. 2013; Walsh et al. 2014). Furthermore, in this calculation dust opacity values at the observed

frequency are assumed, which may have quite uncertain values in protoplanetary discs due to the various grain sizes, compositions, internal structures and the presence of ice mantles (e.g. Pollack et al. 1994; Henning & Stognienko 1996; Semenov et al. 2003; Demyk et al. 2013). The second method infers the gas mass independently of the dust content using molecular lines observations. Due to its high abundance and line strength, CO and its isotopologues are the most frequently used tracers of gas in protoplanetary discs (e.g. Dutrey et al. 1996; Williams & Best 2014). Uncertainties concern the  $H_2$ –CO abundance and the topologist ratios because of photodissociation and freeze-out processes. Miotello, Bruderer & van Dishoeck (2014) showed that the disc mass may be underestimated by up to two orders of magnitude if only a single CO isotopologue line is observed and isotope selective effects are not properly taken into account. As mentioned above, the only direct measurement of hydrogen in discs from TW Hya actually suggests that the discs might be more massive than previously thought. Therefore, current disc mass estimations from gas and dust observations only give a very rough approximation and discs around Class II stars might be sufficiently massive to be gravitationally unstable, as is expected for Class 0 and I objects.

All disc processes including planet–disc interactions and gravitational instability strongly depend on the cooling efficiency, and therefore, the disc temperature. The latter again has an important effect on the shape and contrast of non-axisymmetric disc features. However, many of the numerical simulations so far are limited due to the assumption of a locally isothermal disc. In this paper, a two-dimensional hydrodynamical code which includes an energy equation accounting for disc viscous heating and cooling is used to simulate planet–disc interactions. Gravitational instability is known to produce multi-armed spirals in discs, while planet–disc interactions mostly result in one- or two-armed spirals. In this work, we study the role of heating and cooling, and gravitational instability on the shape and contrast of spiral arms induced by planet–disc interactions. We want to investigate what kind of structures the interplay between gravitational instability and planet–disc interactions can reproduce. This combination has not been investigated so far. The question we ask here is whether the presence of a planet in a marginally gravitationally stable disc can tip it over the limit. The information about the hydrodynamical disc structure is a prerequisite for our radiative transfer modelling in order to link simulation results to observations. The focus is on presenting synthetic scattered light images and on investigating whether they resemble the observations. This paper is divided into five parts. In Section 2 the modifications to the hydrodynamical code and the basic models are introduced. The radiative transfer setup is described in Section 3. Subsequently, the simulation results are presented in Section 4. The associated discussion and the conclusions are summarized in Section 5.

## 2 HYDRODYNAMICAL SETUP

The two-dimensional hydrodynamical grid-based code used for simulating the planet–disc interactions is based on the FARGO-ADSG version (Baruteau & Masset 2008a,b). This modified version of the original FARGO code (Masset 2000) implements an energy equation and the disc self-gravity. The full treatment of disc heating and cooling is important, since in addition to surface density perturbations, temperature perturbations occur as a consequence of shocks along the spirals.

## 2.1 Surface density description

The gas disc is characterized by an initial gas surface density profile  $\Sigma_g(r)$ , a temperature profile  $T(r)$  and a pressure profile  $P(r)$ . The density profile is taken to be a power law combined with an exponential cut-off at long radii, specifically,

$$\Sigma_g(r) = \Sigma_{g,c} \left( \frac{r}{r_c} \right)^{-\delta} \exp \left[ - \left( \frac{r}{r_c} \right)^{2-\delta} \right], \quad (2)$$

where  $r_c$  corresponds to a characteristic scaling radius, which is set to  $\approx 75$  au in accordance with results from high angular resolution disc imaging in the sub-mm regime (Andrews et al. 2010, 2011). In our simulations, this is equal to 0.3 times the outer boundary of the disc.  $\Sigma_{g,c}$  describes the density normalization factor. The surface density index  $\delta$  is taken to be 1 in our simulations. With this surface density profile, it is ensured that most of the disc mass ( $>60$  per cent) is located between  $r_{\text{in}}$  and  $r_c$ , so that artificial reflections from the outer grid boundary can be minimized.

The initial temperature structure is a power law, leading for our case of a non-flaring ( $H/r = \text{const.}$ ) disc to

$$T(r) = \frac{\mu m_p}{\mathcal{R}_g} G M_* h_0^2 r^{-1}, \quad (3)$$

with mean molecular weight  $\mu$ , proton mass  $m_p$ , universal gas constant  $\mathcal{R}_g$ , gravitational constant  $G$  and where the parameter  $h_0$  defines the initial disc aspect ratio.

Such a conical disc defines the boundary between a flaring disc, i.e.  $H/r$  is a monotonic increasing function of radius  $r$ , and a self-shadowed disc, i.e.  $H/r$  becomes smaller with larger radii (Dullemond & Dominik 2004). Considering that we intend to analyse scattered light images, in the case of our non-flaring, constant opening angle geometry, any pressure scaleheight perturbation with a sufficient amplitude can cast a shadow over the remaining outer disc rather easily. In this case of grazing incidence, an object can cast the largest shadow. Furthermore, the background brightness is as dark as possible for grazing infall. Therefore, high-contrast spiral structures are most easily made in non-flaring discs. Furthermore, Juhász et al. (2015) showed that modelling a flaring disc affects the brightness of the disc in the outer regions, but does not improve the visibility of the spirals. From observations, there is also evidence for very low flaring index in MWC 758, one of the sources where a two-armed spiral was detected.

## 2.2 Heating and cooling terms

The development of gravitational instability leads to a self-regulation process due to the competition between heating from the instability and cooling. After a local temperature increase due to a weak shock induced by the spiral density waves, the disc cools back down to a predefined background temperature on a certain time-scale. It takes at least a dynamical time-scale ( $t_{\text{dyn}} \sim \Omega^{-1}$ ) for the disc to hydrodynamically react to any increased temperature. The form of the energy equation implemented in FARGO is

$$\frac{\partial e}{\partial t} + \nabla \cdot (e \mathbf{v}) = -P \nabla \cdot \mathbf{v} + Q_+ - Q_-, \quad (4)$$

where  $e$  denotes the thermal energy per unit area,  $\mathbf{v}$  is the flow velocity,  $P$  is the vertically integrated pressure and  $Q_+/Q_-$  corresponds to the vertically integrated heating/cooling term. The disc heating is assumed to be due to the disc viscosity. For the cooling source term, we assume

$$Q_- = \frac{e}{t_{\text{cool}}} \quad (5)$$

with a cooling time-scale of (cf. Gammie 2001)

$$t_{\text{cool}} = \beta \Omega^{-1}, \quad (6)$$

where the  $\beta$ -factor is taken to be a constant. The power-law background temperature profile to which the disc cools down after each viscous heating event is set by the initial locally isothermal temperature (cf. equation 3). It is assumed to be due to a balance between viscous heating and irradiation heating on the one hand, and radiative cooling on the other hand. With our choice of the background temperature, we expect that local heating by shocks becomes relevant. This shock heating may increase the scaleheight of the disc, which causes local bumps on the disc surface (radial  $\tau = 1$  surface with  $\tau$  being the optical depth). This is different for a flaring irradiated disc ( $T(r) \propto r^{-1/2}$ ) around a Herbig star, where the heating of the disc is dominated by the star apart from the very inner disc. It should be pointed out that in a flaring disc model a different background temperature profile is probably needed.

The question is how fast the disc is able to cool down after a shock. If the cooling is very efficient, i.e.  $t_{\text{cool}} \ll \Omega^{-1}$ , no shock effects will occur and the disc can be assumed to be locally isothermal. For a rough estimation of  $t_{\text{cool}}$  in our models, i.e.  $\beta$ , we do the following calculation. The cooling rate of an accretion disc strongly depends on the optical depth  $\tau$ , but since we are in the range  $\tau \gg 1$ , it is determined by

$$Q_-(r) = 2\sigma T_{\text{eff}}^4(r) = 2\sigma \frac{1}{\tau} T_{\text{mid}}^4(r), \quad (7)$$

where  $\sigma$  is the Stefan–Boltzmann constant and  $T_{\text{mid}}$  corresponds to the mid-plane temperature. The optical depth can be calculated by

$$\tau \simeq \frac{1}{2} \Sigma_d \kappa_d, \quad (8)$$

whereas a dust-to-gas ratio of 0.01 is assumed and for the dust opacity  $\kappa_d$  the Planck mean value is taken. Furthermore, the vertically integrated thermal energy is determined by

$$E = \frac{1}{\gamma - 1} \Sigma_g \frac{k_B T}{\mu m_p}, \quad (9)$$

with Boltzmann constant  $k_B$ , proton mass  $m_p$ , adiabatic index  $\gamma = \frac{7}{5}$  and mean molecular weight  $\mu = 2.3$ . Using equation (5) and dividing  $t_{\text{cool}}$  by the orbital time-scale  $t_{\text{orb}} = 2\pi \Omega^{-1}$  yields the  $\beta$ -factor. Using typical values for the gas density ( $\Sigma_g(50 \text{ au}) \simeq 50 \text{ g cm}^{-2}$ ) and temperature ( $T_{\text{mid}}(50 \text{ au}) \simeq 12 \text{ K}$ ), and setting  $\kappa_d(T_{\text{mid}}) \simeq 10 \text{ cm}^2 \text{ g}^{-1}$  gives  $\beta \simeq 5$ .

There is a critical cooling time-scale,  $t_{\text{cool},c}$ , and a corresponding critical value of  $\beta$ ,  $\beta_c$ , below which the disc can fragment and form gravitationally bound clumps, rather than reaching a steady, gravitoturbulent state. In numerical experiments, Gammie (2001) showed that disc fragmentation can happen for a typical time-scale of  $t_{\text{cool},c} \lesssim 3 \Omega^{-1}$ . Using three-dimensional smoothed particle hydrodynamic (SPH) simulations, Rice et al. (2003) generally confirmed this cooling time for fragmentation, even though the  $\beta_c$ -value is up to a factor of 2 higher for more massive discs. However, Paardekooper (2012) showed that disc fragmentation is a stochastic process and observed it for cooling times up to  $20 \Omega^{-1}$ . In principle, fragmentation is even possible up to  $\beta = 50$ , but it becomes very rare for such high values. A detailed study on the convergence of the critical cooling time-scale with resolution for SPH and grid-based hydrodynamics simulations was done by Meru & Bate (2012). They showed that reducing the dissipation from the numerical viscosity leads to larger values of  $\beta_c$  at a given resolution. In the following, we decided to use  $\beta$ -values of 1 and 10, for which none of our disc models is found to fragment for the selected

**Table 1.** Overview of the models and parameters we studied.

No.	$M_{\text{disc}} [M_{\star}]$	$\alpha$	$M_{\text{p}}/M_{\star}$	$\beta$	$r_c [r_{\text{p}}]$	SG
1	0.15	$10^{-3}$	$10^{-3}$	10	3.0	No
2	0.15	$10^{-3}$	$10^{-3}$	10	3.0	Yes
3	0.15	$10^{-3}$	$10^{-3}$	1	3.0	Yes
4	0.15	$10^{-2}$	$10^{-3}$	10	3.0	Yes
5	0.15	$10^{-3}$	$10^{-2}$	10	3.0	Yes
6	0.2	$10^{-3}$	$10^{-3}$	10	3.0	No
7	0.2	$10^{-3}$	$10^{-3}$	10	3.0	Yes
8	0.2	$10^{-3}$	$10^{-2}$	10	3.0	Yes
9	0.2	$10^{-3}$	$10^{-3}$	10	4.8	Yes

*Notes.* When specified the parameters are given in dimensionless FARGO units, i.e. with respect to the planet’s position and star mass. SG stands for self-gravity. For all simulations  $N_r \times N_{\varphi} = 720 \times 1280$  grid cells, a radial range from  $r_{\text{in}} = 0.2 r_{\text{p}}$  to  $r_{\text{out}} = 10.0 r_{\text{p}}$ , and open boundary conditions at the inner edge are used. Model 1 is taken as the reference model considering a non-self-gravitating disc.

resolution. This allows the disc to settle down into a quasi-steady, self-gravitating state. We note here that, following the argument from Meru & Bate (2012), repeating our simulations with a higher resolution could minimize the artificial viscosity and, therefore, possibly fragment the disc into bound objects.

### 2.3 Parameter choice

The basic model consists of a viscous, non-flaring disc with an embedded giant planet. An overview of all the parameters can be found in Table 1. For the complete series of models, the disc is tapered as described in Section 2.1. FARGO uses dimensionless units; therefore, the fixed orbital radius of the embedded planet of  $r_{\text{p}} = 1$  is used as the length-scale. The disc is divided into 1280 azimuthal and 720 radial grid zones, ranging from 0.2 to 10.0. This corresponds to a physical disc extension from 5 to 250 au, assuming that the planet’s orbital radius is 25 au. The high resolution for the hydrodynamical part serves to avoid strong numerical diffusion effects. The outer boundary of the computational grid is therefore far enough away from the regions around the gap on which we focus our study. Thus, together with the choice of a tapered gas surface density, artificial boundary condition effects are minimized. The inner boundary conditions are open, allowing for mass outflow at the inner edge. The disc is considered to be non-flaring with a constant aspect ratio of  $h = 0.05$ . Furthermore, for the disc viscosity the  $\alpha$ -type viscosity (Shakura & Sunyaev 1973) is used with  $\alpha = [10^{-3}, 10^{-2}]$ . The planet’s potential itself is softened over a length that scales with the disc thickness with a factor of  $\epsilon = 0.6$ . The time-scale over which the planet mass is switched on for the potential evaluation at the beginning of each simulation is set to 100 planetary orbits. The results of all hydrodynamical simulations are described in Section 4.1.

## 3 RADIATIVE TRANSFER SETUP

The outputs of the two-dimensional hydrodynamical simulations previously presented in Section 2, i.e. the gas surface density distribution  $\Sigma_{\text{g}}(r, \varphi)$  and the temperature distribution  $T(r, \varphi)$ , are taken for further processing in the context of three-dimensional radiative transfer modelling. The radiative transfer code RADMC-3D<sup>1</sup> is used to calculate synthetic scattered light images in the NIR. For these calculations, the same aspect ratio, flaring index, and radial and

azimuthal grid extensions as in the hydrodynamical simulations are adopted. In order to avoid low photon statistics, the outputs from the hydrodynamical simulations are interpolated to a grid with a lower resolution ( $N_r \times N_{\varphi} = 384 \times 512$  grid cells) for the radiative transfer calculations. This does not affect the shape of any of the structures obtained in the FARGO simulations. In addition,  $N_{\theta} = 128$  cells for the polar direction are taken. Assuming the vertical density profile to be Gaussian and adopting the canonical dust-to-gas ratio for the interstellar medium of 0.01, the dust volume density is given by

$$\rho(R, \varphi, z) = 0.01 \frac{\Sigma_{\text{g}}(R, \varphi)}{\sqrt{2\pi} H(R, \varphi)} \exp\left(-\frac{z^2}{2 H^2(R, \varphi)}\right), \quad (10)$$

where the spherical coordinates  $R$  and  $z$  can be converted into cylindrical formulas via  $R = r \sin(\theta)$  and  $z = r \cos(\theta)$ , where  $\theta$  is the polar angle. The temperature output after a specific orbital evolution from the hydrodynamical simulations is used to determine the pressure scaleheight  $H(R, \varphi)$ ,

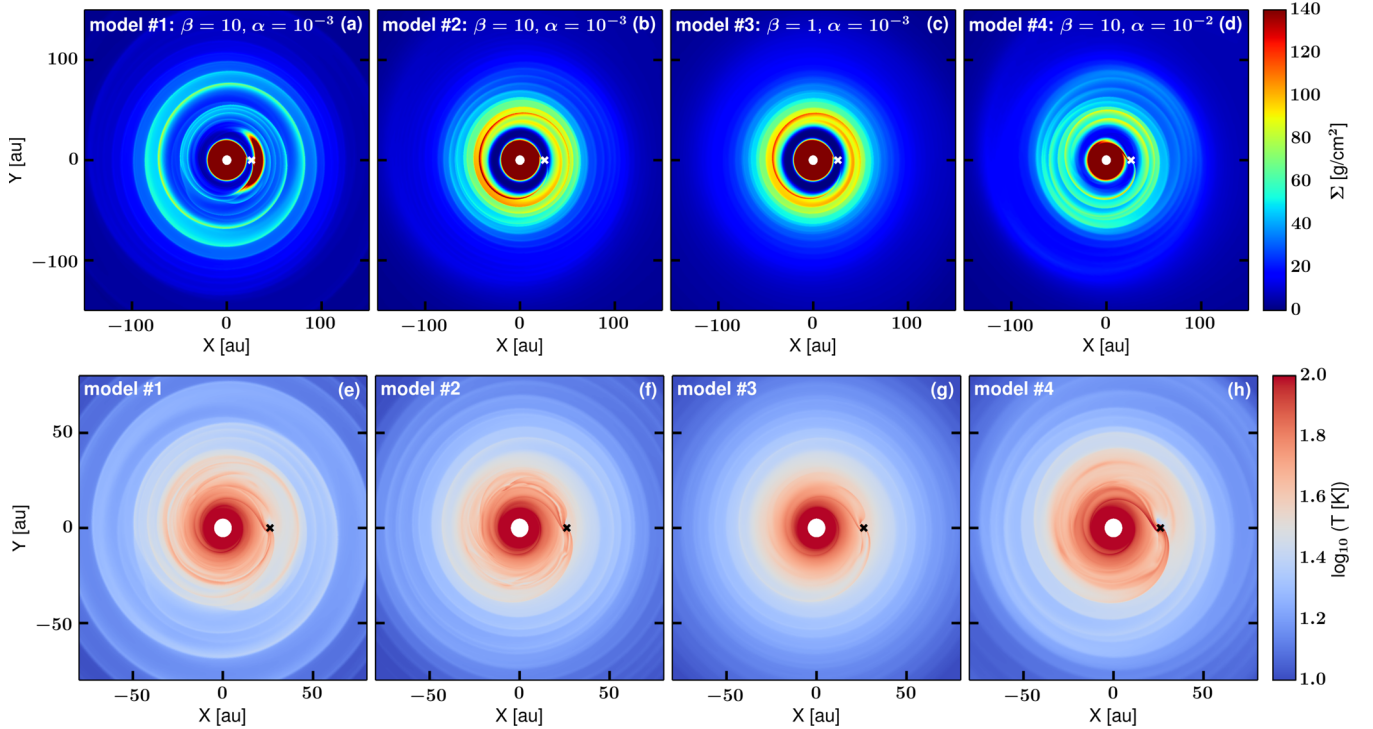
$$H(R, \varphi) = \frac{c_s}{\Omega} = \sqrt{\frac{k_{\text{B}}}{\mu m_{\text{p}} G M_{\star}} T(R, \varphi) R^3} \propto \sqrt{T}. \quad (11)$$

The radiative transfer calculations start with computing the dust temperature structure by means of a thermal Monte Carlo simulation using  $10^7$  photon packages. We want to calculate an equilibrium dust temperature considering the star as the source of luminosity. This temperature is the result of a balance between radiative absorption and re-emission, i.e. a dust grain acquires as much energy as it radiates. This implies that the gas temperature from the hydrodynamical simulations is assumed to only influence the disc scaleheight (cf. equation 11), but is not explicitly related to the dust temperature itself. Hence, a self-enhancement effect, such that the further irradiation of a disc bump actually influences the scaleheight at those positions, is not included in our study. The main inputs for the radiative transfer modelling are the dust density structure from equation (10), dust opacities and the radiation source. The latter is assumed to have typical stellar parameters of a Herbig Ae star ( $T_{\text{eff}} = 9500$  K,  $M_{\star} = 2.0 M_{\odot}$ ,  $R_{\star} = 2.5 R_{\odot}$ ). For simplicity, a blackbody radiation field is assumed. The dust in our models consists of silicate grains with a fixed grain radius of  $0.1 \mu\text{m}$ . The optical constants are obtained from the Jena data base<sup>2</sup> based on the work by Jaeger et al. (1994) and Dorschner et al. (1995). Scattering is considered to be anisotropic with a full treatment of polarization. The scattering phase function depends on the scattering angle and on the polarization state of the input radiation, i.e. the parameters of the full Stokes vector  $[I, Q, U, V]$ . Linearized polarized intensity images can be calculated with  $\text{PI} = \sqrt{Q^2 + U^2}$ . The dust opacity tables are produced with the BHMIE code of Bohren & Huffman (1984) for calculating scattering and absorption by spheres. The images are calculated at  $1.65 \mu\text{m}$  using  $10^8$  photon packages.

The signal-to-noise ratio of the theoretical images obtained from RADMC-3D is limited only by photon statistics. In order to simulate synthetic observations from these models that are comparable to observational data, the images have to be convolved with the telescope’s point spread function (PSF). The diffraction limited PSF for a perfect optical system based on circular elements would be an Airy pattern, whose central peak can be approximated by a two-dimensional Gauss function. A full width at half-maximum (FWHM) of  $0.04$  arcsec is chosen as the size of the PSF, which is

<sup>1</sup> <http://www.ita.uni-heidelberg.de/~dullemond/software/radmc-3d/>

<sup>2</sup> <http://www.astro.uni-jena.de/Laboratory/Database/databases.html>



**Figure 1.** Surface density (top row) and temperature (bottom row) maps for a planet-to-star mass ratio of  $10^{-3}$  after 1000 planetary orbits for different cooling factors  $\beta$  and viscosity values  $\alpha$ . The disc mass corresponds to  $0.15 M_*$ . Note the difference in the x-axis scaling. Panels (a) and (e) show the only model for which self-gravity is turned off. The planet’s position is marked with the white and black cross, respectively.

representative of *H*-band observations with the Spectro-Polarimetric High-contrast Exoplanet REsearch (SPHERE) instrument at the Very Large Telescope (VLT).

## 4 RESULTS

### 4.1 Hydrodynamical results

The numerical hydrodynamic simulations include a treatment of heating and cooling, providing a more realistic picture of the disc gas density and temperature distribution in self-gravitating discs. Hence, it is possible to study the effect of surface density and pressure scaleheight perturbations on the spiral morphology and contrast in scattered light images.

#### 4.1.1 Energy equation versus locally isothermal structure

For our reference model 1, the simulation considers a non-self-gravitating disc with a mass of  $0.15 M_*$ , a planet-to-star mass ratio of  $10^{-3}$  and an  $\alpha$ -viscosity of  $10^{-3}$ . The  $\beta$ -cooling factor is set to 10. In Fig. 1(a), a snapshot of the gas surface density after 1000 planetary orbits is shown. The embedded planet clears out a gap in less than 100 orbits. The disc viscosity is sufficiently low and the planet-to-star mass ratio is high enough, for the angular momentum flux carried by the density waves to overcome the inflow due to viscous accretion in the disc, leading to the gap opening near the planet’s position. The planet is not able to produce a full cavity as seen in observations of transitional discs. This might be, however, linked to the possibility that dust located in the very inner disc close to the star could not be detected, for instance due to grain growth. Furthermore, distinct spiral density waves are induced by

the planet. The dominant feature of this model is a vortex at the outer edge of the gap that moves with the local Keplerian speed and survives until the end of the simulation. The existence of this vortex is in accordance with Ataiee et al. (2013), who showed that vortices can be long-lived for moderate disc viscosity values and massive planets (cf. also Fu et al. 2014; Zhu & Stone 2014). The vortex influences the surface density perturbations induced by the planet close to its position.

Assuming a locally isothermal equation of state corresponds to an infinitely short cooling time-scale. This means that the temperature always remains constant after a density increase since compressive work is immediately radiated away. Therefore, the decisive advantage of the energy equation simulations over locally isothermal models is that the information of the temperature distribution is tracked. In Figs 1(e)–(h), it can be seen that the spirals themselves are hotter than the background disc. The temperature of the spiral depends on the cooling time-scale. If the cooling time-scale is long, the disc material is unable to cool down to its pre-shock equilibrium temperature before the next heating event occurs with the passing of the next shock wave. This leads to an overall higher temperature and disc scaleheight, also outside of the spiral itself. Another interesting finding is that in all temperature plots, best recognizable in Fig. 1(g), the spiral close to the planet’s position seems to split up into two parts. Outwards of the inner spiral and inwards of the outer spiral wake, an additional high-temperature arm is seen. Those structures are located within the planetary horseshoe orbit leading to a rather circular shape. They are caused by shock fronts going through the U-turn of the horseshoe orbit. We can check whether a planet can generate a strong enough perturbation in the temperature to meet the criteria for spiral detection obtained from analytical modelling by Juhász et al. (2015). They found that a relative change of at least

0.2 in pressure scaleheight ( $\delta H/H$ ) is required for the spirals to create detectable observational signatures. Calculating this relative change for our simulation with  $\beta = 10$  gives 0.2 at the  $\varphi = 270^\circ$  axis. This value increases to a maximum of 0.6 near the planet at a zero azimuth angle. Hence, we expect to see the spirals in scattered light (cf. Section 4.2).

#### 4.1.2 Effect of self-gravity

In Fig. 1 the gas surface density (a–d) and temperature (e–h) maps of four different models for a disc mass of  $0.15 M_*$  are displayed. The leftmost panel represents the reference model with a planet-to-star mass ratio of  $10^{-3}$  after 1000 planetary orbits and a cooling time of  $10 \Omega^{-1}$ . The other panels additionally include self-gravity and only differ in the  $\beta$ -factor and  $\alpha$ -viscosity value. Models 2 and 3 (Figs 1 b and c) are quite different from the reference model without self-gravity. The vortex present in the latter case is smeared out very quickly due to self-gravity effects, and no interaction between the vortex and density waves can occur. Additionally, the gap around the planet’s orbit is much more distinct. In contrast to very low mass planets driving a single one-armed spiral (Ogilvie & Lubow 2002), two arms that spiral outwards of the planet’s position with a mutual azimuthal shift are present. Note that this  $m = 2$  mode is not a result of self-gravity, but caused by the massive planet itself. The primary spiral arm launched at the planet’s position is much stronger than the one azimuthally shifted which has a quite low density contrast. Furthermore, it is noticeable that the self-gravity models show more twisted spirals with lower pitch angles. This can be explained by the fact that the disc self-gravity modifies the positions of the Lindblad resonances, between which density waves can propagate. It shifts the location of the effective resonances closer to the planet (Pierens & Huré 2005).

To study the disc stability, we show the azimuthally averaged Toomre parameter  $Q$  as a function of radial distance in Fig. 4. The middle panel only includes the Toomre profiles for models with a disc mass of  $0.15 M_*$ . We compare the models with an embedded planet to the case of a self-gravitating disc without a planet (solid black versus dotted red line). It can be seen that initially  $Q$  is always larger than 1 for all models; thus, the disc by itself would not become gravitationally unstable. Adding a sufficiently massive planet considerably reduces  $Q$  and puts the disc just at the limit of being gravitationally unstable between 50 and 100 au from the star (blue dotted line). The explanation for this behaviour is the following. When a planet is present in the disc, it will carve a gap around its orbit decreasing the surface density in the gap. Since most of the material from the gap is pushed away due to angular momentum exchange between the planet and the disc, the surface density of the disc will increase outside of the gap edge. Hence, a sharp density jump is created. Since  $Q$  is proportional to  $\Sigma^{-1}$ , it is strongly increased near the planet’s position within the gap, but can be reduced just behind the gap edge, where overdensities grow. Thus, if  $Q$  is low enough at the outer edge of the gap, the local increase of surface density due to gap formation might tip the balance and lower  $Q$  below unity. The strong increase of  $Q$  from  $\sim 100$  au is caused by the exponential mass tapering and the disc mass-loss (Fig. 4, left-hand panel). The issue for a disc mass of  $0.15 M_*$  is, however, that  $Q$  is not low enough at larger radii in order to let self-gravity affect the disc dynamics in the form of additional gravitational instability structures. How a larger disc mass or an increased critical taper radius affects these results is discussed in Section 4.1.5.

#### 4.1.3 Effect of cooling time-scale

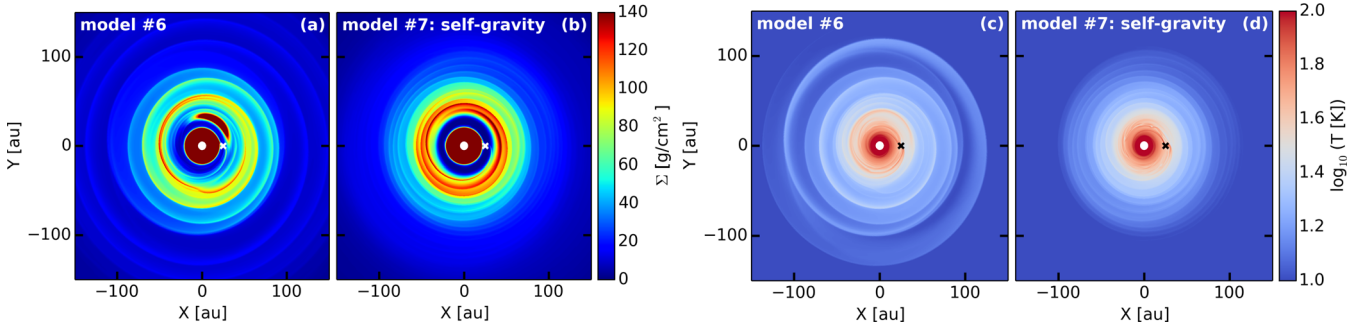
According to equation (6), for constant  $\beta$ ,  $t_{\text{cool}}$  is proportional to  $r^{3/2}$ , so that the cooling is faster in the inner disc. This means that the temperature profile returns back to the locally isothermal background values faster in these inner parts, but in any case not faster than the local dynamical time-scale. In general, we expect to see stronger temperature changes for longer cooling times, i.e. larger  $\beta$ -values. The effect of the cooling time-scale can be seen by comparing models 2 ( $\beta = 10$ ) and 3 ( $\beta = 1$ ). A very fast cooling ( $\beta \leq 1$ ) is very close to the locally isothermal case. In Figs 1(c) and (d), the final surface density and temperature show lower values in absolute strength and contrast along the spirals compared to model 2 (b,f). This slight drop in temperature also decreases the pitch angle, leading to more tightly wound spirals which are smeared out at larger radii. A larger  $\beta$ -value implies less effective cooling, which makes the whole disc, i.e. also the spirals, hotter. The higher the temperature, the faster the waves spread in the disc, since the sound speed is proportional to the square root of the temperature. Hence, more mass flows through the inner disc boundary for  $\beta = 10$  (see Fig. 4, left-hand panel).

#### 4.1.4 Effect of viscosity

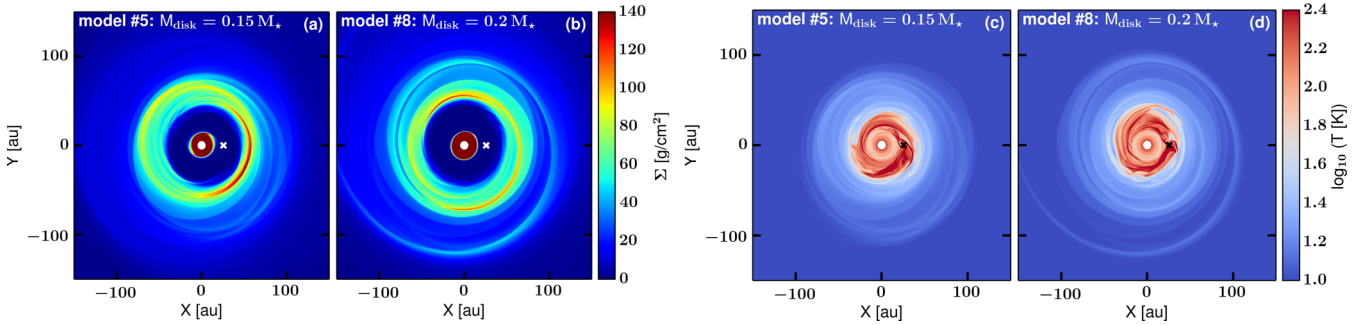
Another important parameter to explore is the viscosity, for which the gap profile dependence is already known (e.g. Crida, Morbidelli & Masset 2006). If the viscosity decreases, the gap becomes gradually deeper. However, the dependence of the gap characteristic on viscosity is less sensitive in numerical simulations compared to theoretical calculations. In simulations there is no simple balance between viscous, pressure and gravitational torque, since parts of the latter are transported away by density waves (Papaloizou & Lin 1984; Rafikov 2002). Nevertheless, by comparing our model 2 ( $\alpha = 10^{-3}$ ) with 4 ( $\alpha = 10^{-2}$ ) in Fig. 1, one can see that as viscosity increases the gap is filled with more gas. Since the primary spiral is launched at the planet’s position within the gap, this also influences the spiral contrast, which is consequently reduced. Furthermore, a higher  $\alpha$ -viscosity smears out the spiral structure and reduces the absolute surface density at the spiral position. Therefore, we use moderate values of viscosity for the rest of our models ( $\alpha = 10^{-3}$ ).

#### 4.1.5 Effects of disc and planet mass

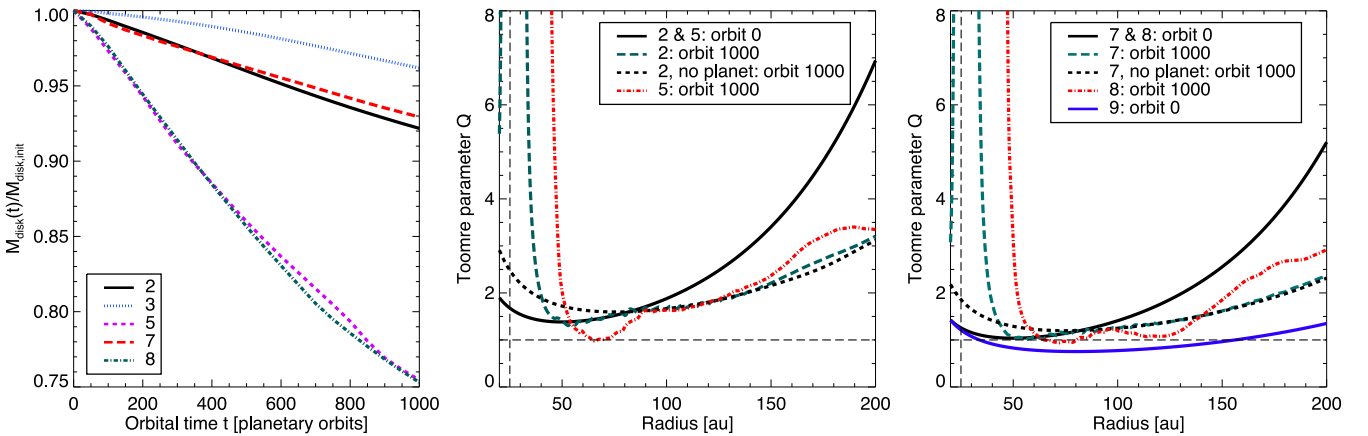
For the simulations 6 and 7 presented in Fig. 2, the disc mass is increased to  $0.2 M_*$ , but the planet-to-star mass ratio of  $10^{-3}$  is kept. The same trends as previously described for the lower disc mass in Section 4.1.2 can be seen. Self-gravity (model 7) destroys the vortex and enhances the strength of the two-armed  $m = 2$  spiral considerably. This is consistent with results from Lin & Papaloizou (2011), who showed that for sufficiently large disc mass and therefore sufficiently strong self-gravity the vortex modes are suppressed. Instead, new global spiral modes develop. Self-gravity effects become even more dominant for higher disc masses. The Toomre parameter profile for this model is plotted in the right-hand panel of Fig. 4 (solid and dashed black lines). The disc is marginally gravitationally unstable from the beginning of the simulation between the planet’s position at 25 and  $\sim 60$  au. This is, however, not sufficient to induce gravitational instability, since the outer disc parts remain stable. Even the perturber with a planet-to-star mass ratio of  $10^{-3}$  is not capable to reduce  $Q$  to a value  $\leq 1$  further out from 60 au (dashed green line). The general effect of increasing the planet-to-star mass ratio is illustrated by comparing



**Figure 2.** Surface density map for a planet-to-star mass ratio of  $10^{-3}$  after 1000 planetary orbits for a model without (a) and with self-gravity (b). Panels (c) and (d) show the corresponding temperature distributions. The disc mass corresponds to  $0.2 M_*$  and  $\beta$  is taken to be 10.



**Figure 3.** Surface density maps (a,b) for a planet-to-star mass ratio of  $10^{-2}$  after 1000 planetary orbits. The disc mass corresponds to  $0.15 M_*$  (a) and  $0.2 M_*$  (b), respectively. The temperature plots are shown in panels (c,d).

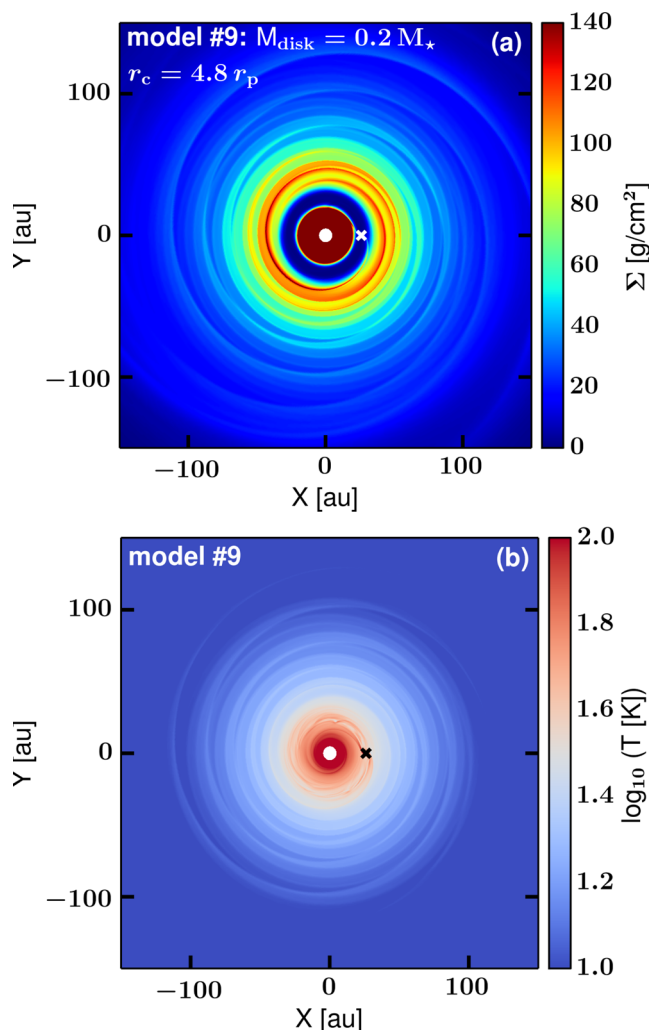


**Figure 4.** Disc mass evolution for models with different disc and embedded planet masses (left). Azimuthally averaged Toomre parameter  $Q$  as a function of radius (middle and right-hand panels). For radii smaller than 20 au,  $Q$  has very large values, which are not relevant for our study and therefore not plotted here. The disc mass in the middle panel corresponds to  $0.15 M_*$ , and for the right-hand panel to  $0.2 M_*$ . The vertical dashed line represents the position of the planet. The horizontal dashed line at  $Q = 1$  shows the critical Toomre value below which the disc is unstable to axisymmetric perturbations. A shearing disc becomes unstable to non-axisymmetric perturbations even at slightly larger values around  $Q \sim 1.5$ – $2$ .

Fig. 3 to Figs 1 and 2. For both disc masses, the higher planet mass forces the gap to become eccentric and a distinct two-armed spiral is not seen anymore. For a  $10^{-2} M_*$  planet mass,  $Q$  is below or close to 1 even for the outer disc region between 100 and 150 au (red dashed line in Fig. 4). The consequence can be recognized in Fig. 3(b). The massive planet is able to initiate gravitational instability, visible as a quite open spiral arm in the outer disc. Particular attention has to be paid to the temperature structure. The massive planet produces significant shock heating within the inner 50 au of the disc. These temperature changes cause perturbations in the vertical structure of the disc, and following equation (11) this strongly

influences the characteristics of our NIR scattered light images (cf. Section 4.2).

Apart from changing the planet mass, the other possibility to trigger gravitational instability in the outer disc is to change the distribution of mass as a function of radius by increasing the mass in the outer disc for a given disc mass. This is done in our simulation with  $0.2 M_*$  by shifting the critical mass taper radius outwards to  $4.8 r_p$  ( $=120$  au). The corresponding surface density and temperature plots are illustrated in Figs 5(a) and (b). A more open, global spiral pattern is seen. These large-scale spirals reflect the onset of gravitational instability, but close to the planet they overlap with



**Figure 5.** Surface density map (a) for a planet-to-star mass ratio of  $10^{-3}$  after 1000 planetary orbits for the model with a critical taper radius of  $4.8 r_p$ . The disc mass corresponds to  $0.2 M_*$ . The temperature structure is shown in panel (b).

the planetary density waves. This instability is supported by the Toomre parameter profile in Fig. 4 (right-hand panel; blue dash-dotted line). The disc is initially gravitationally unstable in most locations.

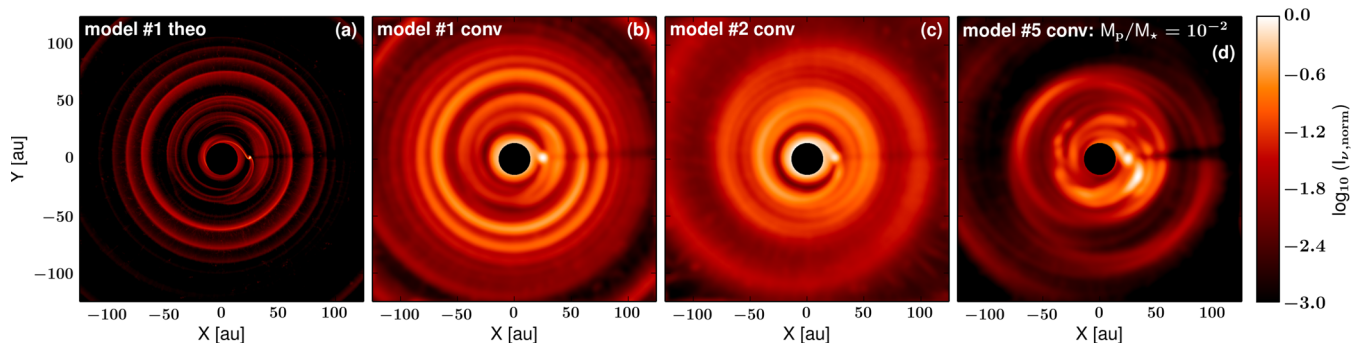
When the disc is at the limit of being gravitationally unstable, only a slight increase in disc mass changes the non-axisymmetric structures considerably. This effect can be seen in Fig. 3. For a disc mass of  $0.2 M_*$  (Fig. 3b), a spiral arm in the outer disc induced by gravitational instability is present. Its pitch angle is significantly higher than that for the planet-induced spirals from models with a lower disc mass (Fig. 3a).

## 4.2 Synthetic radiative transfer images

We aim to investigate the characteristics of the main observational features in transition discs, i.e. gaps and spiral structures, for which full three-dimensional radiative transfer simulations are required. We are interested in the number of spiral arms, their pitch angle, and the brightness contrast between the spiral features and the surrounding background disc to investigate the observability of spirals. We expect that the shock heating along the spiral may increase the scaleheight of the disc locally. This causes bumps on the disc surface, which are irradiated by the star, and thus produce a significant brightness contrast. Based on the model setup described in Section 3, scattered light images in polarized intensity at  $1.65 \mu\text{m}$  are calculated. The disc is nearly face-on with an inclination angle of  $10^\circ$ . The wavelength at which the following images are made is chosen such that the images are directly comparable to  $H$ -band observations made with SPHERE/VLT and the High Contrast Instrument for the Subaru Next Generation Adaptive Optics (HiCIAO). For a realistic comparison to observations, the radiative transfer images are first convolved with a  $0.04$  arcsec Gaussian beam, assuming the source to be at  $140$  pc. Then, each pixel intensity value is multiplied with  $r^2$ , where  $r$  is its distance from the star, in order to compensate for the falloff of the stellar irradiation. This way the outer disc structures become much better visible. A coronagraph is mimicked by masking the inner  $0.1$  arcsec of the disc ( $14$  au at  $140$  pc distance). The images are normalized to the highest disc surface brightness while using a dynamic range of 1000 for plotting the images in logarithmic scale.

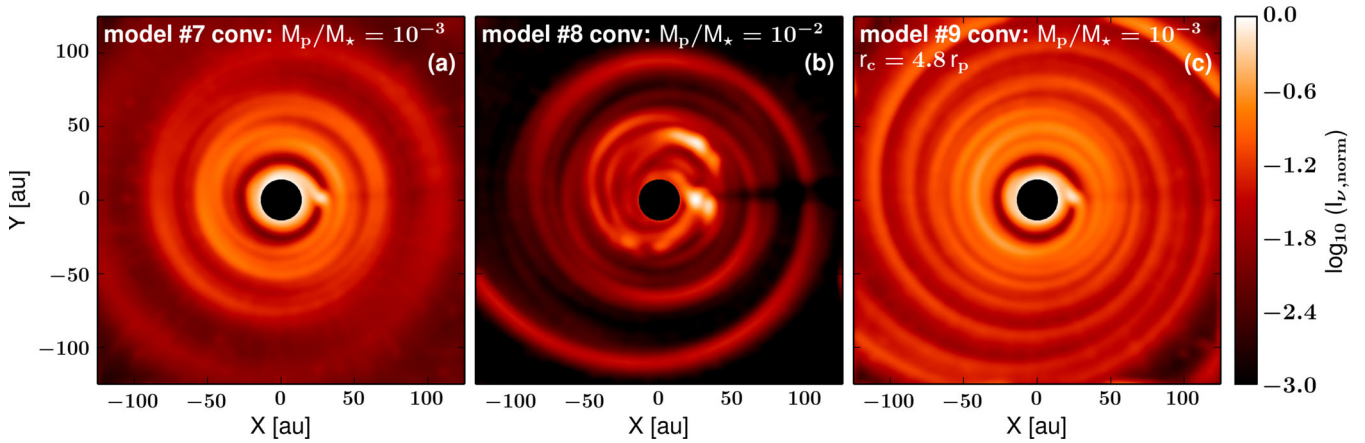
### 4.2.1 Images from models without self-gravity

The planetary  $m = 1$  spiral is clearly visible in the scattered light image at infinite resolution based on the reference model 1 without self-gravity (see Fig. 6a). The brightness contrast between the spiral and the contiguous disc ranges from 12 to 30. The large vortex present in the hydrodynamical simulations is seen in the scattered



**Figure 6.** Simulated NIR scattered light images in  $H$ -band polarized intensity ( $\lambda = 1.65 \mu\text{m}$ ). All models consider a disc mass of  $0.15 M_*$ . Panel (a) corresponds to the reference model 1 without self-gravity and shows the image at original resolution as calculated with the radiative transfer code RADMC-3D. All other images (b–d) are convolved with a Gaussian beam using an FWHM of  $0.04$  arcsec (at  $140$  pc distance), which is representative of observations with SPHERE/VLT in the  $H$  band. The central  $0.1$  arcsec of the image was masked to mimic the effect of a coronagraph similar to real observations.





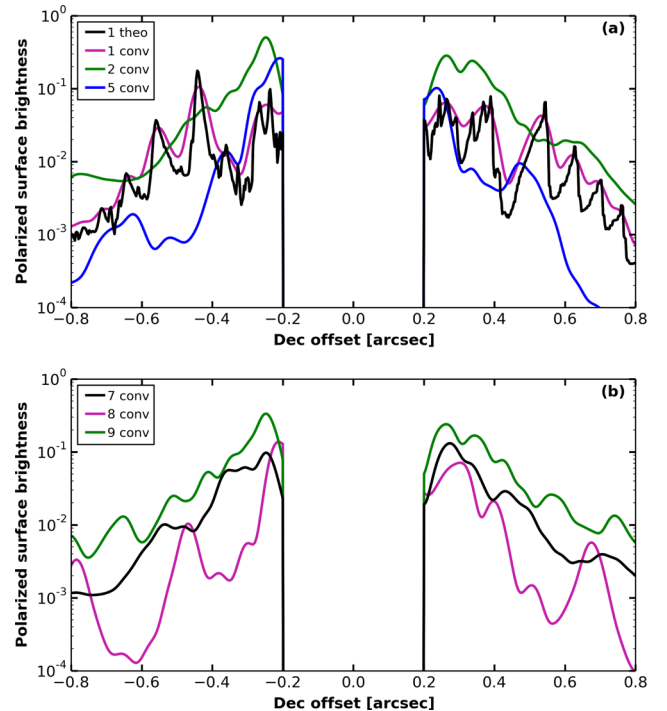
**Figure 7.** Simulated NIR scattered light images in  $H$ -band polarized intensity ( $\lambda = 1.65 \mu\text{m}$ ). All images are based on a  $0.2 M_*$  disc. They are convolved with a Gaussian PSF with FWHM of 0.04 arcsec (at 140 pc distance). For panel (c), the critical taper radius is set to  $4.8 r_p$ . The central black region (0.1 arcsec) mimics the effect of a coronagraph similar to real observations.

light as well. This has, however, to be interpreted with caution, since it is related to the two-dimensional character of our hydrodynamical simulations. Vortices can extend through the whole disc from the mid-plane to the atmosphere (cf. Zhu & Stone 2014), but can be only accurately simulated with three-dimensional hydrodynamical models. If the vertical elongation of the vortex is small, it is expected to appear less extended in scattered light, which traces the disc surface layer. After convolving this theoretical image with the Gaussian 0.04 arcsec PSF, the vortex as well as the planet-induced spirals are significantly smeared out, but still clearly visible. As can be seen in Fig. 8(a), the contrast of the spirals with the surrounding disc is reduced by a factor of about 3, giving contrast values between 3 and 10 (magenta line) for a resolution of 0.04 arcsec. We note that the contrast is expected to be further reduced in observations due to instrument-related and additional noise effects.

#### 4.2.2 Images from models including self-gravity

In all simulations shown in Figs 6(c), (d) and 7(a)–(c), self-gravitating discs are considered. First, we study the effect of self-gravity by comparing the results of models 1 and 2. As already explained in Section 4.1.2, the inclusion of self-gravity destroys the vortex immediately. This leads to a more distinct primary spiral launched close to the planet’s position (see Fig. 6c). Furthermore, the radial brightness cross-sections in Fig. 8(a) illustrate that the clear wiggle structure from the models without self-gravity (magenta line) is smeared out (green line). Self-gravity forces the spiral close to the planet to be more tightly wound. This in turn reduces the total number of individual spiral windings seen in the scattered light, since they cannot be distinguished anymore. The contrast of the spiral with the surrounding disc varies between 1.5 and 10 for a resolution of 0.04 arcsec.

The effect of increasing the planet-to-star mass ratio is displayed in Figs 6(d) and 7(b). The absolute brightness near the planet is increased, leading to several bright spots. Moreover, the gap is not fully detectable anymore in the scattered light. The irregular structure close to the planet’s position is caused by the perturbation in temperature (cf. Figs 3c and d). Thus, this heating strongly increases the disc scaleheight within the gap. This elevated region is now irradiated by the star hiding the gap in the scattered light image. At the same time, structures characteristic of gravitational instability are visible. They occur as quite open spirals reaching also outer



**Figure 8.** Radial profiles of the normalized polarized surface brightness belonging to the calculated images in Figs 6 and 7, respectively. Since the focus is on measuring the contrast of the outer spirals, the profiles are only plotted beyond an inner circle with a radius of 0.2 arcsec.

disc parts. The spiral contrast in the outer disc is extremely high with values of up to 40. A second possibility to create gravitational instability structures is a model with a disc that is gravitationally unstable from the beginning, i.e. no massive planet has to work as a trigger. The resulting scattered light image in Fig. 7(c) shows a quite symmetric spiral for the whole disc. The planet-induced spiral is only visible close to the planetary orbit and the outer disc is dominated by the spirals triggered by gravitational instability. For this model, the polarized surface brightness in Fig. 8(b) has a clear wiggle structure (green line) with a brightness contrast of approximately 3. This contrast is similar to the images from models without self-gravity, but a factor of 10 smaller than for a self-gravitating disc

model with a massive planet ( $10^{-2} M_*$ ). This shows that the brightness contrast between the spirals and the contiguous disc is strongly influenced by the planet-to-star mass ratio in combination with the disc mass distribution. As soon as gravitational instability spirals triggered by a planet develop, the contrast increases tremendously, the spiral shape is strongly influenced and its pitch angle slightly increases.

Another feature to mention, which is present in all images, is the shadow of the planet along a position angle of  $PA = -90$  deg. This is just an artefact of our simulation methods, since we do two-dimensional hydrodynamical, but three-dimensional radiative transfer simulations. This means that a surface density enhancement, e.g. a circumplanetary ring, will increase the total density at all disc heights in the radiative transfer models. It is unlikely that even a circumplanetary disc around a massive planet with several tens of  $M_{\text{jup}}$  can cast a shadow over the whole outer disc in its full vertical extent. The shadow is expected to be seen in the mid-plane, but the disc atmosphere should remain nearly unaffected.

## 5 DISCUSSION AND SUMMARY

We present hydrodynamical simulations of planet–disc interactions considering viscous heating, cooling and additionally self-gravity effects. Subsequently, the density and temperature structure resulting from these simulations are used for the follow-up radiative transfer modelling, in order to predict synthetic scattered light images from a protoplanetary disc around a typical Herbig Ae star. The focus is set on analysing the morphology and detectability of the main observational features of transition discs in scattered light observations, distinct gaps and non-axisymmetric spiral arms.

The conclusions of this paper are as follows.

(1) A clear  $m = 2$  spiral structure, and in some cases even higher azimuthal wavenumber modes, can be excited by a single planet in a self-gravitating disc. The shape of the planet-induced spirals depends on the planet mass and disc parameters, such as the viscosity and heating/cooling time-scales.

(2) When self-gravity of the disc is taken into account, the spirals are found to have a smaller pitch angle, making them more tightly wound. For sufficiently large disc mass, and therefore sufficiently strong self-gravity, the vortex modes are suppressed, and the strength of the spiral is enhanced considerably.

(3) All our calculated polarized intensity images show a planetary gap and spiral arms with a variety of morphologies depending on planet mass, disc mass distribution and the influence of the disc's self-gravity. Convolving the images with a circular Gaussian PSF with an FWHM of 0.04 arcsec (typical for current 8–10 m class telescopes) lowers the contrast by a factor of  $\sim 3$ , but the spiral features are still visible.

(4) The scaleheight along the spiral increases due to the temperature changes, forming local bumps on the disc surface, and thus creates a signal in the reflected light. Therefore, the contrast in the images also depends on the cooling time-scale via the  $\beta$ -factor. In a disc where the cooling is very efficient (i.e.  $\beta \lesssim 1$ ), which approximates the case of a locally isothermal disc, any temperature increase is immediately suppressed. The remaining density perturbation alone is not capable of producing a sufficiently strong spiral contrast in scattered light to be observable with current state-of-the-art telescopes (cf. Juhász et al. 2015).

(5) Our result of a non-self-gravitating disc with an embedded planet shows that the combination of planet-induced density and scaleheight perturbation along the spiral due to shock heating is

sufficient to achieve the contrast between the spiral and the contiguous disc necessary to be detectable in observations. The contrast values estimated for a resolution of 0.04 arcsec are in a range from 3 to 10, which is consistent with the brightness contrast of spirals seen in recent observations.

(6) There are two models in our series for which the Toomre parameter profiles allow the creation of spirals due to gravitational instability. However, their initiation mechanism is different in both cases. For model 9, the disc itself is massive enough, and the mass distribution allows it to be gravitationally unstable from the beginning of the simulation. Spirals are visible in the whole disc; a difference is, however, seen in the pitch angle between the spirals close to the planet and in the outer disc. The spiral is more tightly wound near the planet's position as it is driven by the planet, while the spirals in the outer disc are created by gravitational instability showing a larger pitch angle. A long-scale spiral with quite regular windings is visible in the corresponding scattered light image with a contrast of about 3 for a resolution of 0.04 arcsec.

(7) The second possibility is a disc that is initially at the edge of being gravitationally unstable. As soon as a sufficiently massive planet is included, gravitational instability can be triggered by this planet (model 8). A spiral arm with a huge brightness contrast of  $\sim 40$  and with a higher pitch angle than the spirals driven by a planet in a gravitationally stable disc is generated. The strong temperature perturbations near the high-mass planet partly hide the low-density gap and dominate the irregular structures in this disc region.

To summarize, perturbations in the vertical structure of the disc, i.e. scaleheight perturbations caused by temperature variations due to planet-induced accretion heating or local heating by gravitational instability, are able to create the necessary spiral-background contrast seen in scattered light observations. The presence of a sufficiently massive planet embedded in a marginally gravitationally stable disc can lead to a variety of spiral morphologies depending on the planet and disc mass. It is possible to create a tightly wound spiral close to the planet and a more open spiral arm in the outer disc. However, the explanation of the origin of symmetric, open double-armed spirals seen e.g. in SAO 206462 remains challenging. Recently, three-dimensional global hydrodynamical simulations (isothermal and adiabatic, assuming a non-self-gravitating disc) combined with radiative transfer calculations after a short time of evolution (10–20 local orbits) have shown that inner spirals caused by massive planets ( $6 M_{\text{jup}}$  at  $\sim 100$  au) can be visible at scattered light with similar pitch angle, extension and symmetry as observational results (Dong et al. 2015). However, at longer times of evolution when the disc reaches a quasi-steady state, it is expected that such a planet opens a visible gap in the outer disc (100–200 au) contrary to observational results at NIR and (sub-)mm. With our model setup, there is either one dominant primary spiral in the outer disc when the planet is working as a trigger for gravitational instability or higher mode spirals with similar contrasts are seen when the disc is already initially unstable. A second planet would be required in order to create two primary spirals of an  $m = 1$  or  $m = 2$  structure with nearly the same contrast. However, for this scenario the two planets would need to have just the right mass ratios and be located exactly at the right radial locations and azimuthal angles in the disc to cause symmetric spirals. In addition, other mechanisms apart from self-gravity may play a role for non-axisymmetric structure formation as well. We note that a caveat of our approach is that we do not perform fully consistent radiation hydrodynamics simulations, e.g. including the consistent release of accretion energy from the planet via radiation, since this is quite computational expensive for a parameter study.

Long-term simulations in three dimensions that include these effects are needed to investigate the differences with our current results and for the direct comparison with observations. This will be addressed in future work. Additionally, we are aware that the cooling law from Gammie (2001), used also in most of the previous work about self-gravitating discs, is a simplified description. It has to be elaborated in order to calculate the disc temperature structure more precisely, which influences the local scaleheight changes, and thus also the brightness contrast in scattered light.

## ACKNOWLEDGEMENTS

We are grateful to C. Dominik for useful discussions and remarks on the manuscript. We thank C. Baruteau for his support in solving some numerical problems with the FARGO code. AP is a member of the International Max Planck Research School for Astronomy and Cosmic Physics at the University of Heidelberg, IMPRS-HD, Germany. She acknowledges the CPU time for running simulations on bwGRiD, member of the German D-Grid initiative, funded by the Bundesministerium für Bildung und Forschung and the Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg. PP is supported by Koninklijke Nederlandse Akademie van Wetenschappen (KNAW) professor prize to Ewine van Dishoeck. AJ acknowledges support of the DISCSIM project, grant agreement 341137 funded by the European Research Council under ERC-2013-ADG.

## REFERENCES

- Andrews S. M., Williams J. P., 2005, *ApJ*, 631, 1134  
 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  
 Ataiee S., Pinilla P., Zsom A., Dullemond C. P., Dominik C., Ghanbari J., 2013, *A&A*, 553, L3  
 Avenhaus H., Quanz S. P., Schmid H. M., Meyer M. R., Garufi A., Wolf S., Dominik C., 2014a, *ApJ*, 781, 87  
 Avenhaus H., Quanz S. P., Meyer M. R., Brittain S. D., Carr J. S., Najita J. R., 2014b, *ApJ*, 790, 56  
 Baruteau C., Masset F., 2008a, *ApJ*, 672, 1054  
 Baruteau C., Masset F., 2008b, *ApJ*, 678, 483  
 Beckwith S. V. W., Sargent A. I., Chini R. S., Guesten R., 1990, *AJ*, 99, 924  
 Benisty M. et al., 2015, *A&A*, 578, L6  
 Bergin E. A. et al., 2013, *Nature*, 493, 644  
 Biller B. et al., 2012, *ApJ*, 753, L38  
 Birmstiel T., Dullemond C. P., Brauer F., 2010, *A&A*, 513, A79  
 Boccaletti A., Pantin E., Lagrange A.-M., Augereau J.-C., Meheut H., Quanz S. P., 2013, *A&A*, 560, A20  
 Bohren C. F., Huffman D. R., 1984, *Nature*, 307, 755  
 Brauer F., Dullemond C. P., Johansen A., Henning T., Klahr H., Natta A., 2007, *A&A*, 469, 1169  
 Canovas H., Ménard F., Hales A., Jordán A., Schreiber M. R., Casassus S., Gledhill T. M., Pinte C., 2013, *A&A*, 556, A123  
 Casassus S., Perez M. S., Jordán A., Ménard F., Cuadra J., Schreiber M. R., Hales A. S., Ercolano B., 2012, *ApJ*, 754, L31  
 Casassus S. et al., 2013, *Nature*, 493, 191  
 Close L. M. et al., 2014, *ApJ*, 781, L30  
 Cossins P., Lodato G., Clarke C. J., 2009, *MNRAS*, 393, 1157  
 Crida A., Morbidelli A., Masset F., 2006, *Icarus*, 181, 587  
 de Gregorio-Monsalvo I. et al., 2013, *A&A*, 557, A133  
 Demyk K. et al., 2013, in Andersen A., Baes M., Gomez H., Kemper C., Watson D., eds, *Proc. The Life Cycle of Dust in the Universe: Observations, Theory, and Laboratory Experiments (LCDU2013)*. SISSA, Trieste, p. 44  
 Dong R., Zhu Z., Whitney B., 2014, *ApJ*, preprint (arXiv:1411.6063)  
 Dong R., Zhu Z., Rafikov R., Stone J., 2015, *ApJ*, preprint (arXiv:1507.03596)  
 Dorschner J., Begemann B., Henning T., Jaeger C., Mutschke H., 1995, *A&A*, 300, 503  
 Dullemond C. P., Dominik C., 2004, *A&A*, 417, 159  
 Dutrey A., Guilloteau S., Duvert G., Prato L., Simon M., Schuster K., Menard F., 1996, *A&A*, 309, 493  
 Flock M., Ruge J. P., Dzyurkevich N., Henning T., Klahr H., Wolf S., 2015, *A&A*, 574, A68  
 Forgan D., Rice K., Cossins P., Lodato G., 2011, *MNRAS*, 410, 994  
 Fu W., Li H., Lubow S., Li S., Liang E., 2014, *ApJ*, 795, L39  
 Fukagawa M. et al., 2013, *PASJ*, 65, L14  
 Gammie C. F., 2001, *ApJ*, 553, 174  
 Garufi A. et al., 2013, *A&A*, 560, A105  
 Grady C. A. et al., 2013, *ApJ*, 762, 48  
 Henning T., Stognienko R., 1996, *A&A*, 311, 291  
 Jaeger C., Mutschke H., Begemann B., Dorschner J., Henning T., 1994, *A&A*, 292, 641  
 Juhász A., Benisty M., Pohl A., Dullemond C. P., Dominik C., Paardekooper S.-J., 2015, *MNRAS*, 451, 1147  
 Lin M.-K., Papaloizou J. C. B., 2011, *MNRAS*, 415, 1426  
 Lodato G., Rice W. K. M., 2004, *MNRAS*, 351, 630  
 Lodato G., Rice W. K. M., 2005, *MNRAS*, 358, 1489  
 Lyra W., Turner N. J., McNally C. P., 2015, *A&A*, 574, A10  
 Masset F., 2000, *A&AS*, 141, 165  
 Meru F., Bate M. R., 2012, *MNRAS*, 427, 2022  
 Miotello A., Bruderer S., van Dishoeck E. F., 2014, *A&A*, 572, A96  
 Muto T. et al., 2012, *ApJ*, 748, L22  
 Ogilvie G. I., Lubow S. H., 2002, *MNRAS*, 330, 950  
 Paardekooper S.-J., 2012, *MNRAS*, 421, 3286  
 Papaloizou J., Lin D. N. C., 1984, *ApJ*, 285, 818  
 Pérez L. M., Isella A., Carpenter J. M., Chandler C. J., 2014, *ApJ*, 783, L13  
 Pierens A., Huré J.-M., 2005, *A&A*, 433, L37  
 Pollack J. B., Hollenbach D., Beckwith S., Simonelli D. P., Roush T., Fong W., 1994, *ApJ*, 421, 615  
 Rafikov R. R., 2002, *ApJ*, 572, 566  
 Rameau J., Chauvin G., Lagrange A.-M., Thébault P., Milli J., Girard J. H., Bonnefoy M., 2012, *A&A*, 546, A24  
 Rice W. K. M., Armitage P. J., Bate M. R., Bonnell I. A., 2003, *MNRAS*, 339, 1025  
 Rice W. K. M., Lodato G., Pringle J. E., Armitage P. J., Bonnell I. A., 2004, *MNRAS*, 355, 543  
 Semenov D., Henning T., Helling C., Ilgner M., Sedlmayr E., 2003, *A&A*, 410, 611  
 Shakura N. I., Sunyaev R. A., 1973, *A&A*, 24, 337  
 Thalmann C. et al., 2014, *A&A*, 566, A51  
 Toomre A., 1964, *ApJ*, 139, 1217  
 van der Marel N. et al., 2013, *Science*, 340, 1199  
 Walsh C. et al., 2014, *ApJ*, 791, L6  
 Williams J. P., Best W. M. J., 2014, *ApJ*, 788, 59  
 Zhu Z., Stone J. M., 2014, *ApJ*, 795, 53

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