2-11-2020

The Effects of Disc Self-Gravity and Radiative Cooling on the Formation of Gaps and Spirals by Young Planets

Shangjia Zhang
University of Nevada, Las Vegas, shangjia.zhang@unlv.edu

Zhaohuan Zhu
University of Nevada, Las Vegas, zhaohuan.zhu@unlv.edu

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **493,** 2287–2305 (2020)
Advance Access publication 2020 February 11

doi:10.1093/mnras/staa404

The effects of disc self-gravity and radiative cooling on the formation of gaps and spirals by young planets

Shangjia Zhang ★ and Zhaohuan Zhu

Department of Physics and Astronomy, University of Nevada, Las Vegas, 4505 S. Maryland Pkwy, Las Vegas, NV 89154, USA

Accepted 2020 January 31. Received 2020 January 30; in original form 2019 November 4

ABSTRACT

We have carried out 2D hydrodynamical simulations to study the effects of disc self-gravity and radiative cooling on the formation of gaps and spirals. (1) With disc self-gravity included, we find stronger, more tightly wound spirals and deeper gaps in more massive discs. The deeper gaps are due to the larger Angular Momentum Flux (AMF) of the waves excited in more massive discs, as expected from the linear theory. The position of the secondary gap does not change, provided that the disc is not extremely massive ($Q \gtrsim 2$). (2) With radiative cooling included, the excited spirals become monotonically more open (less tightly wound) as the disc's cooling time-scale increases. On the other hand, the amplitude and strength of the spirals decrease when the cooling time increases from a small value to $\sim 1/\Omega$, but then the amplitude starts to increase again when the cooling time continues to increase. This indicates that radiative dissipation becomes important for waves with $T_{\rm cool} \sim 1$. Consequently, the induced primary gap is narrower and the secondary gap becomes significantly shallower when the cooling time becomes $\sim 1/\Omega$. When the secondary gap is present, the position of it moves to the inner disc from the fast cooling cases to the slow cooling cases. The dependence of gap properties on the cooling time-scale (e.g. in AS 209) provides a new way to constrain the disc optical depth and thus disc surface density.

Key words: hydrodynamics – waves – planet–disc interactions – protoplanetary discs.

1 INTRODUCTION

Protoplanetary discs from ALMA observations reveal many substructures (e.g. gaps, rings, and spiral arms) in dust continuum emission (Pérez et al. 2016; Andrews et al. 2018; Long et al. 2018). Among them, the most common features are concentric gaps and rings (Huang et al. 2018). There are many interpretations for these features, such as zonal flows (Flock et al. 2015), aggregate sintering (Okuzumi et al. 2016), secular gravitational instabilities (Takahashi & Inutsuka 2014), self-induced dust pile-ups (Gonzalez, Laibe & Maddison 2017), dust growth at snowlines (Zhang, Blake & Bergin 2015; Pinilla et al. 2017), planet–disc interactions (de Juan Ovelar et al. 2013; Dong, Zhu & Whitney 2015), and so on. Recent high-resolution observations for a relatively large sample of discs (Huang et al. 2018; Long et al. 2018; van der Marel et al. 2019) conclude that snowlines cannot consistently match the positions of these gap features, but it may still apply to individual objects.

Of all the exciting scenarios, this paper focuses on the interpretation of planet–disc interactions. Assuming these features are due to the planets, many authors infer the planet mass from the gap properties by comparing observations with hydrodynamical simulations. These properties include gap width and depth in dust emission (Zhang et al. 2018), and sub/super-Keplerian rotational velocity in CO channel maps (Teague et al. 2018; Zhang et al. 2018; Gyeol Yun et al. 2019). Most past works associate each gap with a planet. Recent simulations show more promising results that one single planet can explain up to five gaps in ALMA observations (Bae, Zhu & Hartmann 2017; Guzmán et al. 2018; Zhang et al. 2018). These multiple gaps that are induced by a single planet is due to the presence of multiple spirals, which results from constructive interference of the density waves at different m modes in the inner disc (Bae & Zhu 2018a,b; Miranda & Rafikov 2019a). The one-armed spiral opens the primary (major) gap at the position of the planet, and the secondary spiral opens the secondary gap in the inner disc and so on. The spacing of the gaps is also useful to constrain the gaseous disc scale height h/r since the ratio of the positions between the secondary gap and the primary gap mainly depends on the disc h/r (Dong et al. 2018; Zhang et al. 2018). If we can further constrain h/r using other observables like gap widths and depths, we can use the position of the secondary gap to even constrain the planet mass (Kanagawa et al. 2015, 2016; Bae et al. 2017; Zhang et al. 2018; Gyeol Yun et al. 2019).

However, most simulations that are used to infer the planet mass from the disc substructures neglect the (1) self-gravity from the disc and (2) radiative cooling. For the latter, they employ locally isothermal Equation of State (EoS) instead, which is equivalent to instant cooling. These two simplifications might not be valid

★ E-mail: shangjia.zhang@unlv.edu

to realistic discs. Consequently, the planet masses inferred from observations might be subject to systematic errors. Self-gravity and radiative cooling can play essential roles in some discs.

Several recent works suggest discs might be more massive than previously thought. Booth et al. (2019) measure the HD 163296 disc mass using $^{13}C^{17}O$ line, and find the gas mass is a factor of 2–6 higher than previously estimated using $C^{18}O$. Using the DSHARP opacity (Birnstiel et al. 2018) and optical–thin assumption, Zhang et al. (2018) find that Toomre Q at the gap edges are $\lesssim$10 for most of the discs (see table 3 therein). Among the same series of papers, Dullemond et al. (2018) also find that most of the prominent rings are marginally gravitationally stable if the gas-to-dust ratio is 100 (see fig. 7 therein). Furthermore, Zhu et al. (2019) suggests that these discs might be very optically thick, even though dust scattering makes these discs look like optically thin. Without the assumption of the dust opacity, Powell et al. (2019) analyse seven discs by measuring the locations of 'dust lines' – the cutoff radii of continuum emission – at different wavelengths, and also find all of those discs have $Q \lesssim 10$, with several approaching unity. For these reasons, disc self-gravity can become quite important for the interactions between planets and discs.

The disc should also have orders of magnitude difference in cooling time at different radii (Zhu et al. 2015; Miranda & Rafikov 2019b). Assuming the Minimum Mass Solar Nebula (MMSN), the cooling is faster at the outer disc and slower at the inner disc – 7 orders of magnitude of the cooling time between 1 au ($10^5/\Omega$) and 100 au ($10^{-2}/\Omega$) (Zhu et al. 2015). The fast cooling resembles the locally isothermal disc, whereas the slow cooling represents the adiabatic disc. However, it is unclear how the disc might look like beyond these two extremes. Radiative cooling with different cooling time-scales at different locations in discs might further change the disc substructures.

In this paper, we study spirals and gaps induced by planets in discs with non-negligible self-gravity and radiative cooling. In Section 2, we lay out the theoretical background of the linear theory and some basic quantities being used throughout the paper. In Section 3, we introduce the simulation setups. In Section 4, we present the results after adding these two physical processes separately, and also explore a situation that includes both self-gravity and radiative cooling. We also compare our simulation results with previous analytical results from Goldreich & Tremaine (1980) and Miranda & Rafikov (2019a). In Section 5, we quantify the change of AMF and discuss the observational implications in linear regime. Then, we take AS 209 disc as a test bed for these two processes, and address some limitations of our simulations. Finally, we conclude this paper in Section 6.

2 THEORETICAL BACKGROUND

Goldreich & Tremaine (1978,1979,1980) develop the linear theory for planet–disc interactions. This theory produces insightful analytical results which can be used to understand numerical simulations on planet migration and gap opening.

In order to study planet–disc interactions in the linear regime, the planet mass should be less than the thermal mass (Goodman & Rafikov 2001)

$$M_{\rm th} = \left(\frac{h}{r}\right)^3_{\rm p} M_* = 1 M_{\rm J} \left(\frac{(h/r)_{\rm p}}{0.1}\right)^3 \frac{M_*}{M_\odot}, \tag{1}$$

where $(h/r)_{\rm p}$ is the disc's aspect ratio at the planet's position. M_* is the mass of the central star. At the thermal mass, the gaseous disc scale height is comparable to both the Hill and Bondi radii of the planet, and the distance between the planet and spiral shock forming region is comparable to the planet's Bondi radius so that spiral shocks begin to affect planet accretion (Béthune & Rafikov 2019). If $M_{\rm p} \gtrsim M_{\rm th}$, the density wave excited by the planet starts from the non-linear regime, so that the spiral waves immediately become spiral shocks after the wave excitation. Below the thermal mass, the spiral shock region can be well separated from the wave launching area. Most of the simulations in this paper have $M_{\rm p} \lesssim M_{\rm th}$, but we also explore several cases in the high mass regime.

Angular momentum transport is one of the key aspects in the accretion theory. In the linear theory, the changing rate of the angular momentum in the disc enclosed within a radius r is equal to the planet's torque Γ on that region minus the Angular Momentum Flux[1] (AMF) F_J that flows out of the region, that is

$$\frac{{\rm d}L}{{\rm d}t} = \Gamma(r) - F_J(r). \tag{2}$$

2.1 Torque

The first term on the right-hand side of equation (2) is the planet's torque to the disc region within r,

$$\Gamma = -\int_{\rm disc} \Sigma(\vec{r} \times \vec{F}) {\rm d}f = \int_r \frac{{\rm d}T}{{\rm d}r} {\rm d}r, \tag{3}$$

as in Kley & Nelson (2012), where

$$\frac{{\rm d}T}{{\rm d}r}(r) = r \int_0^{2\pi} \Sigma(r,\phi) \frac{\partial \Phi_p}{\partial \phi} {\rm d}\phi \tag{4}$$

is torque density and $\Phi_{\rm p}$ is the planet's potential in the disc coordinate,

$$\Phi_{\rm p} = -\frac{GM_{\rm p}}{\left\{r_{\rm p}^2 + r^2 - 2r_{\rm p} r \cos(\phi - \phi_{\rm p}) + s^2\right\}^{1/2}}, \tag{5}$$

where $r_{\rm p}$, $\phi_{\rm p}$ are planet's position in radial and azimuthal directions and s is the smoothing length used in simulations to avoid the singularity of the planetary potential. The planet feels the opposite torque, which leads to planet migration if the torque from the inner disc is not balanced by the torque from the outer disc.

Goldreich & Tremaine (1980) derive the one-sided torque as

$$F_{J0} = (M_{\rm p}/M_*)^2 h_{\rm p}^{-3} \Sigma_{\rm p} r_{\rm p}^4 \Omega_{\rm p}^2, \tag{6}$$

where $\Omega_{\rm p}$ is the angular frequency of the planet (all the 'p' as a subscript in the paper refers to that quantity evaluated at the position of the planet). This value is commonly used to normalize the torque and AMF.

2.2 AMF

The second term on the right-hand side of equation (2) is the AMF carried by the disturbance (e.g. spiral waves). The angular momentum carried by the spiral waves will be eventually deposited to the background disc to induce gaps in discs. In order to understand the gap properties, such as depth, width, and secondary gap positions, it is necessary to understand the AMF of the spiral waves. With disc self-gravity included in the analysis, AMF has two components, advective AMF, $F_{\rm A}$, and gravitational AMF, $F_{\rm G}$ (Lynden-Bell & Kalnajs 1972; Goldreich & Tremaine 1979). The

[1]This is coined as angular momentum current in Binney & Tremaine (2008), denoted by capital 'C'.

latter only occurs in self-gravitating discs. The total AMF is, $F_J = F_A + F_G$.

F_A is the advective transport of angular momentum due to the motion of the fluid and leads to a flow of angular momentum through circumference at radius r. It is also related to the Reynolds stress in turbulence studies. It can be calculated as,

$$F_A(r) = r^2\Sigma(r)\oint u_r(r,\phi)u_\phi(r,\phi)\mathrm{d}\phi, \tag{7}$$

where Σ is the gaseous disc surface density and u_r and u_ϕ are the velocity perturbations in the r and ϕ directions.

F_G is due to the non-axisymmetric part of the disc potential. The spiral structure produces a spiral gravitational field, which exerts torque and transfers angular momentum from one part of the disc to another. It is non-zero in self-gravitating discs and becomes important when the following Tommre Q parameter is small,

$$Q = \frac{c_s\Omega}{\pi G\Sigma}, \tag{8}$$

where c_s is the sound speed and $c_s = (h/r)v_\phi$. Given fixed c_s and Ω, a smaller Q means a higher disc surface density. When $Q = 1$, the disc becomes gravitationally unstable. The value of F_G is the torque exerted to the inner disc,

$$F_G(r) = \int_0^r \mathrm{d}r' r' \int_0^{2\pi} \mathrm{d}\phi\Sigma\frac{\partial\Phi_{\rm out}}{\partial\phi}, \tag{9}$$

where $\Phi_{\rm out}$ is the disc potential at r (the formula is the same as equation (4), except that $\Phi_{\rm p}$ is replaced with $\Phi_{\rm out}$). For trailing spiral arms, the angular momentum is transferred from the inner to the outer disc, since the inner disc exerts positive gravitational torque on the outer disc (Lynden-Bell & Kalnajs 1972; Binney & Tremaine 2008).

Note that, in globally isothermal discs, the total AMF is conserved ($\mathrm{d}F_J/\mathrm{d}r = 0$) based on the linear theory (Goldreich & Tremaine 1979). However, as Miranda & Rafikov (2019b) point out, this is not the case in locally isothermal discs due to the torque applied on to the wave by the background shear flow (this is also reported in Lin & Papaloizou 2011; Lin 2015). In locally isothermal discs, $\mathrm{d}(F_J/c_s^2)/\mathrm{d}r = 0$ is conserved.

In the linear theory, all quantities can be further Fourier decomposed into individual m harmonics,

$$\eta(r,\phi) = \sum_{m=-\infty}^{\infty}\eta_m(r)e^{im(\phi-\phi_p)}, \tag{10}$$

where η can be u_r, u_ϕ, and Σ, and η_m is a complex number. Putting $u_r(r)$ and $u_\phi(r)$ in equation (7),

$$\begin{aligned} F_A(r) &= 2\pi r^2\Sigma(r)\sum_{m=-\infty}^{\infty}u_{r,m}(r)u^*_{\phi,m}(r) \\ &= 4\pi r^2\Sigma(r)\sum_{m=0}^{\infty}\Re[u_{r,m}(r)u^*_{\phi,m}(r)]. \end{aligned} \tag{11}$$

The second equality holds because the Fourier transform of real signals is Hermitian. If F_A is expanded with positive integer harmonics (the $m = 0$ term is always zero),

$$F_A(r) = \sum_{m=1}^{\infty}F_{A,m}(r), \tag{12}$$

each m component of the advective AMF can be expressed as,

$$\begin{aligned} F_{A,m}(r) &= 4\pi r^2\Sigma(r)\Re[u_{r,m}(r)u^*_{\phi,m}(r)] \\ &= 4\pi r^2\Sigma(r)\Big(\Re[u_{r,m}(r)]\Re[u_{\phi,m}(r)] + \Im[u_{r,m}(r)]\Im[u_{\phi,m}(r)]\Big), \end{aligned} \tag{13}$$

which is identical to the result using continuous Fourier transform in the shearing sheet geometry (Dong et al. 2011b; Rafikov & Petrovich 2012). Likewise, the total AMF can be expanded as,

$$F_J(r) = \sum_{m=1}^{\infty}F_{J,m}(r). \tag{14}$$

The sum goes to infinity, but the normalized AMF ($F_{J,m}/F_{J0}$) reaches the maximum at $m \sim (1/2)/(h/r)$ and becomes $\ll 1$ at several $1/(h/r)$. Thus, the sum can stop at some m (Goldreich & Tremaine 1980; Artymowicz 1993a,b; Ward 1997), which is known as the 'torque cutoff'.

2.2.1 *Relative importance between* F_A *and* F_G

Here we discuss the relative importance between F_A and F_G. We borrow the WKB solutions in the tightly-wound limit ($k \gg 1/R$) obtained by Goldreich & Tremaine (1979). Their analytical solution for the advective AMF is,

$$F_A = -\frac{\pi m r\Sigma k}{2\pi G\Sigma|k| - k^2c_s^2}\left(1 - \frac{c^2|k|}{2\pi G\Sigma}\right)^2\Phi(r)^2, \tag{15}$$

where $\Phi(r)$ is the amplitude of the perturbed disc gravitational potential in the form of WKB solution. The gravitational AMF is,

$$F_G = \mathrm{sgn}(k)\frac{mr\Phi^2(r)}{4G}. \tag{16}$$

Taking the ratio between the F_A and the F_G, and define the critical wavenumber $k_{\rm crit}$ (Binney & Tremaine 2008) as,

$$k_{\rm crit} = \frac{\Omega^2}{2\pi G\Sigma} = \frac{Q}{2}\frac{\Omega}{c_s} = \frac{Q}{2h}, \tag{17}$$

we get,

$$\frac{F_A}{F_G} = 2\left(\frac{Q^2}{4}\frac{|k|}{k_{\rm crit}} - 1\right) = 2\left(\frac{|k|}{k_{\rm cross}} - 1\right), \tag{18}$$

where $k_{\rm cross} = 4k_{\rm crit}/Q^2 = 2/Qh$. When $|k| = (3/2)k_{\rm cross}$, advective AMF and gravitational AMF have equal contribution, and we denote that $|k|$ as $k_{\rm eq}$. $k_{\rm eq} = (3/2)k_{\rm cross} = 3/Qh$. When $|k| < k_{\rm eq}$, F_G has larger contribution, whereas when $|k| > k_{\rm eq}$, F_A contributes more. Hence, to compare the relative importance of F_A and F_G, we just need to compare the characteristic values of $|k|$ and $k_{\rm eq}$.

Plugging equations (8) and (17) into the dispersion relation,

$$\tilde{\omega}^2 = \Omega^2 - 2\pi G\Sigma|k| + c_s^2k^2, \tag{19}$$

where $\tilde{\omega}(r) = m[\Omega_{\rm p} - \Omega(r)]$ and $\Omega(r)$ is the angular velocity of the disc, we get the absolute value of the wavenumber in unit of $k_{\rm crit}$,

$$\frac{|k|}{k_{\rm crit}} = \frac{2}{Q}\left\{m^2\left(1 - \frac{r^{3/2}}{r_{\rm p}^{3/2}}\right)^2 - \left(1 - \frac{1}{Q^2}\right)\right\}^{1/2} + \frac{2}{Q^2}. \tag{20}$$

Here we choose the positive sign in front of the first term since it corresponds to the short wave that propagates beyond Lindblad resonances in discs (the negative sign solution corresponds to the long wave that is restrained inside Lindblad resonances and outside the forbidden region; see fig. 2 in Lovelace, Jore & Haynes 1997, Fig. 6.14 in Binney & Tremaine 2008 and equation 19 in

Goldreich & Tremaine 1979 for details). Note that this equation reduces to the following (Ogilvie & Lubow 2002; Bae & Zhu 2018a) as $Q \to \infty$,

$$\frac{|k|}{m} = \frac{\Omega}{c_s}\left|\left(1 - \frac{r^{3/2}}{r_p^{3/2}}\right)^2 - \frac{1}{m^2}\right|^{1/2}. \quad (21)$$

With equation (20), we evaluate the ratio of the $|k|$ and $k_{\rm eq}$ at the effective locations of Lindblad resonances. While both sound pressure and self-gravity shift the positions of the Lindblad resonances (Pierens & Huré 2005), we only account for the contribution from the pressure effect for simplicity. The locations of the resonances become,

$$r = r_{\rm p}\left\{1 \pm \frac{\left[1 + m^2(h/r)^2\right]^{1/2}}{m}\right\}^{2/3}. \quad (22)$$

At these radii, the ratio reads,

$$\frac{|k|}{k_{\rm eq}} = \frac{Q}{3}\left\{m^2(h/r)^2 + \frac{1}{Q^2}\right\}^{1/2} + \frac{1}{3}. \quad (23)$$

As $Q \to \infty$, the ratio $|k|/k_{\rm eq} \to \infty$, which means that $F_{\rm G}$ has no contribution to the total AMF. However, when $Q = 5$ and we adopt $m = (1/2)/(h/r)$, $|k| \approx k_{\rm eq}$. Hence, $F_{\rm G}$ contributes comparable amount of AMF as $F_{\rm A}$. This simple calculation shows that as Q becomes smaller, the gravitational AMF becomes increasingly important and will exceed the contribution of the advective AMF at very low Q.

We caution that the detailed analysis on the relative importance of $F_{\rm A}$ and $F_{\rm G}$ requires numerically solving the linearized equations and is beyond the scope of this paper. In this paper, since we focus on discs with $Q \gtrsim 5$, we will only discuss $F_{\rm A}$ for simplicity. As shown in Section 5.4, we find that when $Q \sim 2$, the gap edge quickly becomes unstable even though the disc itself is gravitationally stable.

2.2.2 *Calculation of the total AMF*

Here we summarize the calculation of the total AMF in the linear theory. In regions far from corotation ($|r - r_{\rm p}|/r_{\rm p} \to \infty$), advective AMF dominates ($F_{\rm A} \gg F_{\rm G}$; Goldreich & Tremaine 1980). Since F_J is conserved at different radii, F_J equals $F_{\rm A}$ far from corotation where WKB solutions (of equation 83 therein) are valid. The analytical solution can be obtained (Goldreich & Tremaine 1978) at the following condition (Goldreich & Tremaine 1980),

$$m(h/r) \ll \min\left(\frac{Q^2 - 1}{3Q^2}, 1\right). \quad (24)$$

The resulting AMF is given by,

$$F_J^{WKB}(m) = \frac{4}{3}\frac{m^2\Sigma}{\Omega^2}\left(\frac{GM_p}{R}\right)^2\{2K_0(2/3) + K_1(2/3)\}^2, \quad (25)$$

where the last term is around 6.35, and K_0 and K_1 are zeroth and first-order modified Bessel function of the second kind. The value of $F_J^{WKB}(m)$ is approximately equal to 8.47 $m^2 F_{J0}$. However, the condition (24) is not always satisfied in the disc. The analytical solution deviates from the exact solution when $m \gtrsim 1/(h/r)$, or disc self-gravity is important ($m \gtrsim \frac{Q^2-1}{3Q^2}$). At this regime, Goldreich & Tremaine (1980) numerically integrate the WKB trial solutions for each m harmonic at given Q and obtain $F_J(m, Q)$, expressed in unit of $F_J^{WKB}(m)$ (see figs 2 and 3 therein and Figs 3 and 4 in this paper). Summing up each m mode, the total AMF is given by,

$$F_J = f(Q)\frac{F_J^{WKB}(m)}{m^2} \approx 8.47 f(Q) F_{J0}, \quad (26)$$

where F_{J0} is expressed in equation (6), and $f(Q)$ is a factor, which is a function of Q. From Goldreich & Tremaine (1980),

$$f(Q) = \frac{1}{3}\mu_{\rm max}^3 = \int_0^\infty {\rm d}\mu\mu^2 \frac{F_J(m, Q)}{F_J^{WKB}(m)}, \quad (27)$$

where $\mu = mc_s/\Omega r = m(h/r)$. The coefficient $F_J/F_{J0} \approx 0.93$ if $Q = \infty$ and $F_J/F_{J0} \approx 8.6$ if $Q = 2$, which means that more massive discs (with lower Toomre Q values) have higher AMF. We will compare this analytical solution with our simulation results in 4.1.1.

2.3 Pitch angle

The pitch angle β is defined as the angle between the direction of the line tangent to the spiral arm and the azimuthal direction in the disc. If the spiral arms have m-fold rotational symmetry, and k is the radial wavenumber under the WKB approximation,

$$\beta = \tan^{-1}\left(\frac{m}{kr}\right). \quad (28)$$

The dispersion relation for the spiral wave in the tightly wound limit (equation 19) can also be used to derive the pitch angle. In the absence of disc self-gravity, the middle term on the right-hand side of equation (19) becomes zero. Thus, by manipulating equation (21), the pitch angle for the m mode is,

$$\tan(\beta) = \frac{m}{kr} = \frac{c_{\rm s}}{\left\{|\Omega_{\rm p} - \Omega|^2 - \frac{\Omega^2}{m^2}\right\}^{1/2}}. \quad (29)$$

If $m \gg 1$, different m modes have the same pitch angle and interfere with each other (Ogilvie & Lubow 2002), so that

$$\tan(\beta) = \frac{m}{kr} = \frac{1}{r}\frac{c_{\rm s}}{|\Omega_{\rm p} - \Omega|}. \quad (30)$$

When the disc self-gravity is non-negligible, we can rewrite equation (20) to derive the pitch angle

$$\tan(\beta) = \frac{1}{r}\frac{c_{\rm s}}{\left\{|\Omega_{\rm p} - \Omega|^2 - \frac{\Omega^2}{m^2}\left(1 - \frac{1}{Q^2}\right)\right\}^{1/2} + \frac{\Omega}{mQ}}. \quad (31)$$

Compared to equation (29), two additional terms are $(\Omega/mQ)^2$ within the square root and Ω/mQ outside the square root. When $Q \gg 1$, equation (31) reduces to equation (29). Equation (31) also indicates that the pitch angle becomes smaller for higher disc masses (with smaller Q). However, this effect is not strong unless Q is very close to unity. For example, when $m \gg 1$, equation (31) reduces to equation (30) even in massive discs. Thus, different m modes should still have the same pitch angle and interfere with each other unless Q is close to one. The effect should be stronger in the inner disc than the outer disc since Ω decreases with r. We will confirm this pitch angle calculation using our simulations in Section 4.1.

2.4 Orbital cooling

Miranda & Rafikov (2019b) carry out linear analysis on locally isothermal discs with temperature varying with the radius and run simulations to show that the spiral wave absorbs AMF from the background when it is propagating to hotter regions, resulting in higher AMF and higher density perturbation. The phenomenon is unique to locally isothermal discs. It would not happen even if the adiabatic disc is very close to locally isothermal (γ is very close to 1). They advocate that the inclusion of the energy equation would result in weaker density waves and shallower gaps. Thus, planet

masses inferred from previous locally isothermal simulations might be systematically lower than their actual masses.

Besides running locally isothermal simulations, we further study this issue by carrying out adiabatic simulations with a simple orbital cooling prescription. The adiabatic EoS is

$$P = (\gamma - 1)E, \tag{32}$$

where E is the internal energy per unit area and γ is the adiabatic index. We adopt $\gamma = 1.4$ in this study. Since the energy equation is solved in simulations with the adiabatic EoS, we can prescribe the effect of radiative cooling using the orbital cooling approach:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\frac{E - c_v \Sigma T_{\rm irr}}{t_{\rm cool}}, \tag{33}$$

where Σ is the disc surface density. $T_{\rm irr}$ is the disc initial temperature, which is determined by the stellar irradiation in realistic discs. $c_v \equiv R/(\mu(\gamma - 1))$ is the heat capacity per unit mass, $R \equiv k/m_{\rm H}$ is the specific gas constant, k is the Boltzmann constant, μ is the mean molecular weight, $m_{\rm H}$ is the atomic unit mass, and $t_{\rm cool}$ is the cooling time. It is useful to define the dimensionless cooling time

$$T_{\rm cool} = t_{\rm cool}\Omega(r). \tag{34}$$

Thus, $T_{\rm cool}$ is $t_{\rm cool}$ in the unit of orbital time over 2π. This prescription is identical to Zhu et al. (2015) and similar to the β-cooling in Gammie (2001). Small $T_{\rm cool}$ means fast cooling. Thus, the adiabatic disc with fast cooling should be closer to the isothermal disc, whereas large $T_{\rm cool}$ means slow cooling, and this disc is closer to the adiabatic disc without cooling. In the rest of the paper, we simply use adiabatic discs to refer to the adiabatic discs without cooling.

To estimate $T_{\rm cool}$ in realistic discs, we follow Zhu et al. (2015) which use the radiative cooling rate of

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\frac{16}{3}\sigma(T_{\rm mid}^4 - T_{\rm irr}^4)\frac{\tau}{1+\tau^2}, \tag{35}$$

where σ is the Stefan-Boltzmann constant, $\tau = (\Sigma/2)\kappa_{\rm R}$ is the optical depth in the vertical direction, $\kappa_{\rm R}$ is the Rosseland mean opacity normalized to the gas surface density assuming dust-to-gas ratio is 1/100, $\kappa_{\rm R} = \kappa_{\rm R}(a_{\rm max}, T)$, and $T_{\rm mid}$ is the mid-plane temperature. Assuming $E = c_v \Sigma T_{\rm mid}$ and using equations (33) and (35), we can derive

$$t_{\rm cool} = \frac{3\Sigma c_v}{16\sigma(T_{\rm mid}^2 + T_{\rm irr}^2)(T_{\rm mid} + T_{\rm irr})}\frac{1+\tau^2}{\tau}. \tag{36}$$

Approximating $T_{\rm mid} = T_{\rm irr} = T_{\rm d}$, where

$$T_{\rm d}(r) = \begin{cases} \left(\frac{\phi L_*}{8\pi r^2 \sigma}\right)^{1/4} & r \le r_{T_{\rm f}} \\ T_{\rm f} & r > r_{T_{\rm f}} \end{cases}, \tag{37}$$

where $\phi = 0.02$, representing flaring angle (Dullemond et al. 2018), and L_* is the stellar luminosity. We assume the disc temperature decreases as r and at a given radius it reaches a minimum floor temperature $T_{\rm f}$, setting by the background heating processes (e.g. cosmic rays). This radius is

$$r_{T_{\rm f}} = 39\ {\rm au}\left(\frac{T_{\rm f}}{20\ {\rm K}}\right)^{-2}\left(\frac{L_*}{L_\odot}\right)^{1/2}\left(\frac{\phi}{0.02}\right)^{1/2}. \tag{38}$$

The dimensionless cooling time becomes,

$$T_{\rm cool} = 0.015\left(\frac{f}{0.01}\right)^{-1}\left(\frac{\kappa_R(a_{\rm max},T_d)}{1\ {\rm cm^2 g^{-1}}}\right)^{-1}\left(\frac{L_*}{L_\odot}\right)^{-3/4}\left(\frac{\phi}{0.02}\right)^{-3/4} \times\left(\frac{M_*}{M_\odot}\right)^{1/2}(1+\tau^2)\begin{cases} 1 & r \le r_{T_{\rm f}} \\ \left(\frac{r}{r_{T_{\rm f}}}\right)^{-3/2} & r > r_{T_{\rm f}} \end{cases}, \tag{39}$$

where f is the dust-to-gas mass ratio. For a given protoplanetary disc interior to $r_{T_{\rm f}}$, and assuming constant f, $T_{\rm cool}$ only depends on the Rosseland mean opacity in the optically thin limit, whereas $T_{\rm cool} \propto \tau^2/\kappa_{\rm R}$ in the optically thick limit. Within $r_{T_{\rm f}}$, it does not explicitly depend on r (the Rosseland mean opacity depends on the temperature which varies with the radius). Note that this optical depth is different from the optical depth in the dust continuum observation, since the Rosseland mean opacity is used for the cooling process. Aside from the optical depth effect, higher opacity, higher stellar radiation, larger flaring angle and smaller stellar mass would make $T_{\rm cool}$ smaller and vice versa. We also ignore energy diffusion in the radial direction (Goodman & Rafikov 2001), which might be important for very optically thick discs.

Here we provide several typical values of $\kappa_{\rm R}$ at different temperatures for references. Using the DSHARP opacity (Birnstiel et al. 2018), the Rosseland mean absorption opacity for the dust population with the maximum dust size $a_{\rm max} = 1$ mm and the size distribution $n(a) \propto a^{-3.5}$ is $\kappa_{\rm R}(20\ {\rm K}) = 0.21\ {\rm cm^2 g^{-1}}$, $\kappa_{\rm R}(40\ {\rm K}) = 0.48\ {\rm cm^2 g^{-1}}$, and $\kappa_{\rm R}(125\ {\rm K}) = 1.03\ {\rm cm^2 g^{-1}}$.

For a solar-type star, suppose $\Sigma = 100\ {\rm g\ cm^{-2}}$, $f = 0.01$ at 10 au and $T = 40$ K, $T_{\rm cool} \sim 20$, which is approaching the adiabatic limit. Suppose $\Sigma = 1\ {\rm g\ cm^{-2}}$, $f = 0.01$ at 100 au, and $T = 20$ K, $T_{\rm cool} \sim 0.02$, which is on the locally isothermal limit. Thus, the $T_{\rm cool}$ estimated from the gap substructures can help constrain the optical depth thus the surface density of the disc. We will use this argument to constrain the surface density at ∼100 au of the AS 209 disc in Section 5.3.

3 METHOD

To test the effects of disc self-gravity and radiative cooling on planet–disc interactions, we run four sets of 2D hydrodynamical simulations. We use FARGO-ADSG (Baruteau & Masset 2008a,b; Baruteau & Zhu 2016) in the first three sets and `Athena++` (Stone et al., in preparation) in the fourth set. For all the simulations, we fix the planet in circular orbits and do not consider planet migration. We apply the evanescent boundary condition to all the simulations. We choose not to include the indirect term for separating its effect from the effects studied in this paper. Our simulation setups are motivated and based on Bae & Zhu (2018a) and Bae et al. (2017). We keep all the parameters the same but add disc self-gravity or/and radiative cooling. We denote them as `B18` (Set 1) and `B17` (Set 2), respectively. We denote simulations with adiabatic EoS and radiative cooling as `AD`, and simulations with self-gravity as `SG`. For example, a simulation that is built on Bae et al. (2017) with both orbital cooling and self-gravity is written as `B17ADSG`. We also run a set of simulations `AS209` (Set 3) including radiative cooling and self-gravity with massive planets by adopting the parameters in the AS 209 α varying model in Zhang et al. (2018) ($\{(h/r)_{\rm p}, \alpha, M_{\rm p}\} = \{0.05, 3\times10^{-4}(r/r_{\rm p})^2, 1\ M_{\rm th}\}$). The `Athena++` (Stone et al. 2020, in preparation) simulations (Set 4) have the same setup as `B17` and `B17AD` to verify the results. All the simulations that appear in this paper are summarized in Table 1, and descriptions are detailed as follows.

(1) `B18` focuses on the linear regime of planet–disc interactions with very low mass planets. This setup is also adopted in Miranda & Rafikov (2019b). The planet mass $M_{\rm p} = 0.01\ M_{\rm th}$. The aspect ratio at $r_{\rm p}$ is $(h/r)_{\rm p} = 0.1$. The gas surface density is $\Sigma \propto r^{-1}$. The disc temperature is $T(r) \propto r^{-0.5}$ ($h/r \propto r^{0.25}$) with the locally isothermal EoS. The disc viscosity α is 0 (inviscid). The inner boundary is 0.05 $r_{\rm p}$, whereas the outer boundary is 5.0 $r_{\rm p}$. The

Table 1. Models.

FARGO-ADSG									
Run	Q	$T_{\rm cool}$	$M_{\rm p}/M_{\rm th}$	α	$(h/r)_{\rm p}$	q	Domain (R)	Resolution ($R \times \phi$)	$t_{\rm grow}/t_{\rm p}$
B18	∞	0	0.01	0	0.1	0.5	[0.05, 5]	4096 × 5580	1
B18SG	100	0	0.01	0	0.1	0.5	[0.05, 5]	4096 × 5580	1
B18SG	10	0	0.01	0	0.1	0.5	[0.05, 5]	4096 × 5580	1
B18SG	5	0	0.01	0	0.1	0.5	[0.05, 5]	4096 × 5580	1
B18SG	2	0	0.01	0	0.1	0.5	[0.05, 5]	4096 × 5580	1
B18AD	∞	10^{-4}	0.01	0	0.1	0.5	[0.05, 5]	2048 × 2790	1
B18AD	∞	10^{-3}	0.01	0	0.1	0.5	[0.05, 5]	2048 × 2790	1
B18AD	∞	0.01	0.01	0	0.1	0.5	[0.05, 5]	2048 × 2790	1
B17	∞	0	0.1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	20
B17SG	100	0	0.1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	20
B17SG	10	0	0.1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	20
B17SG	5	0	0.1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	20
B17SG	2	0	0.1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	20
B17SG	100	0	0.3	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	10, 60
B17SG	10	0	0.3	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	10, 60
B17SG	5	0	0.3	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	10, 60
B17SG	2	0	0.3	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	10, 60
B17SG	100	0	1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	100
B17SG	10	0	1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	100
B17SG	5	0	1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	100
B17SG	2	0	1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	100
B17SG	100	0	3	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	100
B17SG	10	0	3	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	100
B17SG	5	0	3	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	100
B17SG	2	0	3	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	100
B17AD	∞	0.01	0.1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	20
B17AD	∞	0.1	0.1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	20
B17AD	∞	1	0.1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	20
B17AD	∞	10	0.1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	20
B17AD	∞	100	0.1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	20
B17ADSG	5	0.01	0.1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	20
B17ADSG	5	0.1	0.1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	20
B17ADSG	5	1	0.1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	20
B17ADSG	5	10	0.1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	20
B17ADSG	5	100	0.1	5×10^{-5}	0.07	0	[0.2, 2]	2048 × 5580	20
AS209AD	∞	0.01	1	$3 \times 10^{-4}(r/r_{\rm p})^2$	0.05	0	[0.2, 2]	1024 × 2790	20
AS209AD	∞	0.1	1	$3 \times 10^{-4}(r/r_{\rm p})^2$	0.05	0	[0.2, 2]	1024 × 2790	20
AS209AD	∞	1	1	$3 \times 10^{-4}(r/r_{\rm p})^2$	0.05	0	[0.2, 2]	1024 × 2790	20
AS209AD	∞	10	1	$3 \times 10^{-4}(r/r_{\rm p})^2$	0.05	0	[0.2, 2]	1024 × 2790	20
AS209AD	∞	100	1	$3 \times 10^{-4}(r/r_{\rm p})^2$	0.05	0	[0.2, 2]	1024 × 2790	20
AS209ADSG	5	0.01	1	$3 \times 10^{-4}(r/r_{\rm p})^2$	0.05	0	[0.2, 2]	1024 × 2790	20
AS209ADSG	5	0.1	1	$3 \times 10^{-4}(r/r_{\rm p})^2$	0.05	0	[0.2, 2]	1024 × 2790	20
AS209ADSG	5	1	1	$3 \times 10^{-4}(r/r_{\rm p})^2$	0.05	0	[0.2, 2]	1024 × 2790	20
AS209ADSG	5	10	1	$3 \times 10^{-4}(r/r_{\rm p})^2$	0.05	0	[0.2, 2]	1024 × 2790	20
AS209ADSG	5	100	1	$3 \times 10^{-4}(r/r_{\rm p})^2$	0.05	0	[0.2, 2]	1024 × 2790	20
Athena++									
B17	∞	0	0.1	5×10^{-5}	0.07	0	[0.2, 2]	1024 × 2796	5
B17AD	∞	0.01	0.1	5×10^{-5}	0.07	0	[0.2, 2]	1024 × 2796	5
B17AD	∞	0.1	0.1	5×10^{-5}	0.07	0	[0.2, 2]	1024 × 2796	5
B17AD	∞	1	0.1	5×10^{-5}	0.07	0	[0.2, 2]	1024 × 2796	5
B17AD	∞	10	0.1	5×10^{-5}	0.07	0	[0.2, 2]	1024 × 2796	5
B17AD	∞	100	0.1	5×10^{-5}	0.07	0	[0.2, 2]	1024 × 2796	5

Note. The q is the exponent in the temperature profile, $T(r) \propto r^{-q}$. In AD models, the $(h/r)_{\rm p}$ and the temperature profile are just initial values.

damping regions are inside 0.06 $r_{\rm p}$ and outside 4.6 $r_{\rm p}$. We run five simulations with $Q = \infty$ (non-self-gravitating), 100, 10, 5, and 2, where Q is evaluated at $r_{\rm p}$. The self-gravity smoothing length is 0.3 h. Given the density and temperature profiles, Q is larger at the inner disc and smaller at the outer disc. $Q \propto r^{-4/3}$ and Q at the inner (outer) boundary is 9.5 times (0.30 times) the value at $r_{\rm p}$. The resolution in (logarithmically spaced) r and ϕ directions is 4096 × 5580. There are ∼89 grid cells per scale height in the r direction. The planet potential is in the form of equation (5) and the smoothing length s is 0.6 h. The implementation of the Poisson equation solver for self-gravity can be found in Baruteau & Masset (2008b). To study the AMF in simulations with adiabatic EoS and

temperature varying with the radius, we also run three simulations (B18AD) with lower resolution (2048 × 2790) for $T_{\rm cool} = 10^{-4}$, 10^{-3}, and 0.01 cases.

(2) B17 focuses on gap opening by more massive planets. These simulations start with globally isothermal discs. While one subset keeps the temperature constant, the other subset (AD) allows the disc to cool. The aspect ratio $(h/r)_{\rm p} = 0.07$, with $h/r \propto r^{0.5}$ so that T is a constant. The disc viscosity $\alpha = 5 \times 10^{-5}$, and the gas surface density $\Sigma \propto r^{-1}$. The inner boundary is 0.2 $r_{\rm p}$ and the outer boundary is 2.0 $r_{\rm p}$. The damping regions are inside 0.24 $r_{\rm p}$ and outside 1.6 $r_{\rm p}$. The planet smoothing length s is 0.1 h. If the self-gravity is included, the self-gravity smoothing length is 0.3 h. $Q \propto r^{-1/2}$ and Q at the inner (outer) boundary is 2.2 times (0.71 times) the value at $r_{\rm p}$. The resolution is 2048 (logarithmically spaced) × 5580 in the r and ϕ directions, so that dr:d$\phi \approx$ 1:1 for every grid in the domain. There are ∼ 62 grid cells per scale height in the r direction. The planet mass grows with time as $M_{\rm p} = M_{{\rm p},f} \sin^2(\frac{\pi}{2}\frac{t}{t_{\rm grow}})$, and $t_{\rm grow}$ is 20 $t_{\rm p}$ for 0.1 $M_{\rm th}$, 60 $t_{\rm p}$ for 0.3 $M_{\rm th}$ and 100 $t_{\rm p}$ for 1 and 3 $M_{\rm th}$, where $t_{\rm p} = 2\pi/\Omega$ is the orbital period of the planet. The planet grows to its full mass at $t_{\rm grow}$ and stays constant afterwards. We mainly use $M_{\rm p} = 0.3\,M_{\rm th}$ cases to study the gap opening, but some instability has developed for the $Q = 2$ disc before its planet grows to the full mass, complicating the analysis. Thus, we add four simulations ($M_{\rm p} = 0.3\,M_{\rm th}$, and $Q = 100$, 10, 5, and 2) with $t_{\rm grow} = 10\ t_{\rm p}$. This set of simulations with a shortened planet growing time is almost identical to the previous one at the time-step we choose to analyse, but now the planet grows to its full mass before the instability occurs. Overall, we have $Q = \infty$, 100, 10, 5, and 2 discs, with $M_{\rm p} = 0.1$, 0.3, 1 and 3 $M_{\rm th}$.

This set of simulations also includes dust, represented by 200 000 dust super particles of different sizes. The Stokes numbers (St) of the particles at $r_{\rm p}$ ranges from 1.57×10^{-5} to 1.57. The setup for dust particles is identical to Zhang et al. (2018). The Stokes number St for particles (also called particles' dimensionless stopping time) is

$$St = t_{\rm stop}\Omega = \frac{\pi s \rho_{\rm p}}{2\Sigma_{\rm gas}} = 1.57 \times 10^{-3} \frac{\rho_{\rm p}}{1\,{\rm g\,cm^{-3}}} \frac{s}{1{\rm mm}} \frac{100{\rm g\,cm^{-2}}}{\Sigma_{\rm g}}, \quad (40)$$

where $\rho_{\rm p}$ is the density of the dust particle, s is the radius of the dust particle, and $\Sigma_{\rm g}$ is the gas surface density. We assume $\rho_{\rm p} = 1\,{\rm g\,cm^{-3}}$ in our simulations. The Stokes number mentioned above is the Stokes number at the beginning of the simulations, St_0. As the disc evolves, the dust sizes are fixed, but as they can drift in the disc, the Stokes number can change.

For B17AD and B17ADSG discs, $T_{\rm cool}$ is constant everywhere in each simulation, but $t_{\rm cool}$ varies as radius since $\Omega_k \propto r^{-3/2}$. We explore $T_{\rm cool} = 0.01$, 0.1, 1, 10, and 100, covering fast cooling to slow cooling. Note that for these simulations with cooling, the given h/r and the temperature profiles above are just the initial values. Their values are subject to change as the simulations evolve. We also run a set of B17ADSG discs, which includes both radiative cooling and self-gravity. These discs have $Q = 5$ at the planet's position and $T_{\rm cool} = 0.01$, 0.1, 1, 10, and 100.

(3) AS209 is used to test the effects of radiative cooling in the massive planet regime (∼thermal mass). We run a set of varying α ($\alpha = 3 \times 10^{-4}(r/r_{\rm p})^2$) models (AS209AD) with $T_{\rm cool} = 0.01$, 0.1, 1, 10, and 100. The $(h/r)_{\rm p}$ is 0.05 and $M_{\rm p}$ is 1 $M_{\rm th}$ or 0.1 M_J if $M_* = M_\odot$. The resolution in the r and ϕ directions is 1024 × 2790. With lower resolution, we are able to run the simulations longer. We also add a set of simulations that also include self-gravity (AS209ADSG). Q is chosen to be 5 at $r_{\rm p}$. The planet grows to its full mass at 20 $t_{\rm p}$. Other parameters are the same as (2).

(4) Athena++ simulations are used to verify the results. They include an isothermal simulation and a suite of adiabatic simulations, with the same setup as $M_{\rm p} = 0.1\,M_{\rm th}$, B17 and B17AD discs in (2), except for a shorter $t_{\rm grow} = 5\,t_{\rm p}$. To avoid a numerical artefact due to the small smoothing length, we use $s = 0.2\,h$ for $T_{\rm cool} = 100$ (twice as large as what is used in the FARGO simulation).

4 RESULTS

4.1 Self-gravitating discs

To study the effects of disc self-gravity in the linear regime, we first analyse B18SG discs with $M_{\rm p} = 0.01\,M_{\rm th}$. Fig. 1 shows the density perturbations with different disc masses in the polar coordinate. The planet is at $\phi = 180^\circ$ and $r = 1\ r_{\rm p}$. The density perturbation is

$$\delta\Sigma/\Sigma_0 = (\Sigma - \Sigma_0)/\Sigma_0, \quad (41)$$

where Σ_0 is the gaseous surface density at the initial condition, and Σ is the density at the time-step ($t = 10\ t_{\rm p}$ for this plot) that is analysed. From left to right, the disc mass increases. The Toomre Q values at the $r = r_{\rm p}$ are 100, 10, 5, and 2, respectively. As the disc mass increases, the spiral perturbations become stronger. This transition is the most evident from the $Q = 5$ to $Q = 2$ disc. The density perturbation changes slightly until $Q = 5$ and has a noticeable increase from the $Q = 5$ to $Q = 2$ disc. The underlying AMF that influences the density perturbation is shown in Fig. 2. The AMF increases as Q decreases. The AMF at $Q = 100$ (blue solid curve) is almost identical to the one in the non-self-gravitating disc (black curve) and identical to the result in fig. 1(a) in Miranda & Rafikov (2019b). As pointed out by Miranda & Rafikov (2019b), we can clearly see that the AMF increases towards the inner disc where the disc is hotter. However, it cannot be captured even with very small cooling time $T_{\rm cool} = 0.01$ and 0.001, as shown by the short and long dashed curves. Different from the locally isothermal simulation, the spiral waves cannot pick up the AMF from the background disc when they propagate to hotter regions, even with a relatively short $T_{\rm cool}$. Only extremely small cooling time ($T_{\rm cool} = 10^{-4}$, shown by the dotted-dashed curve) can capture this increase of the AMF towards inner disc (the slightly lower values are due to the lower resolution). The jump of AMF is the largest from the $Q = 5$ to $Q = 2$ disc, where the AMF value of the latter almost doubles. The increase of the AMF with a smaller Q is consistent with the fact that the density perturbation becomes stronger with a smaller Q.

The spirals become more tightly wound as the disc mass increases. It is the most evident at the inner disc between the $Q = 5$ and $Q = 2$ disc. This effect is consistent with the prediction in Section 2 around equation (31). Only when Q becomes very small, the two additional terms become important, which makes the pitch angle smaller. Note that Q is larger than $Q_{\rm p}$ in the inner disc, given the surface density profile we choose. The changes in pitch angle would be larger if Q is constant across the disc. This trend is also reported in Pohl et al. (2015), where their simulations lie in the non-linear regime with higher planet masses. Hence, the decrease of pitch angle with the disc mass applies to a large range of planet masses.

Yu, Ho & Zhu (2019) measure the pitch angles from 10 protoplanetary discs in near-infrared scattered-light images and three discs in ALMA millimetre dust continuum images. They find a tight relation between the spiral arm pitch angle and the disc mass – more massive discs have smaller pitch angles. While we find that

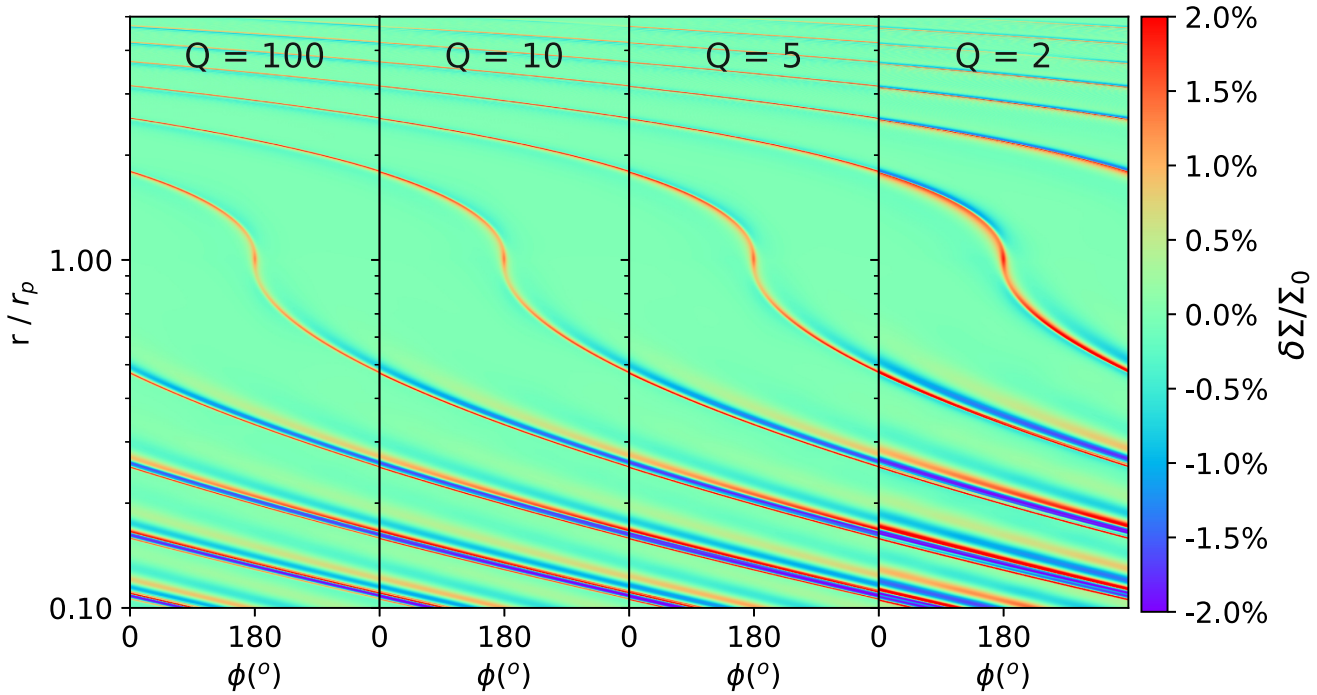

Figure 1. The gaseous density perturbations ($\delta\Sigma/\Sigma_0$) of the inviscid self-gravitating discs (`B18SG`) with a low mass planet $M_{\rm p} = 0.01\,M_{\rm th}$ at 10 orbits. The disc masses increase from left to right. At the position of the planet, the Toomre Q parameters are 100, 10, 5, and 2, respectively. Since the density profile goes as r^{-1}, Q is higher at the inner disc and lower at the outer disc. The spirals become stronger and tighter as the mass increases. This is the most evident in the $Q = 2$ disc.

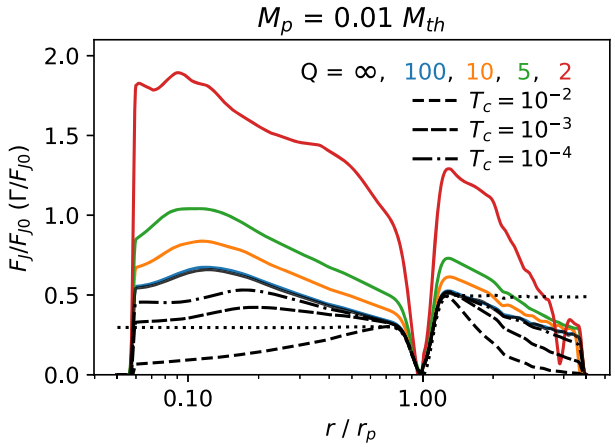

Figure 2. The Angular Momentum Flux (AMF) of the `B18`, `B18SG`, and `B18AD` discs calculated from equation (7) and normalized by equation (6). The black curves represent the AMF (solid curve) and torque (dotted curve) of the non-self-gravitating disc, which is almost identical to fig. 1(a) in Miranda & Rafikov (2019b). The AMF increases as the disc mass increases. The AMF of $Q = 2$ disc is ~ 3 times that of $Q = 100$ disc. This results in a stronger density perturbation. The short dashed, long dashed, and dashed-dotted curves in black represent `B18AD` with $T_{\rm cool} = 0.01$, 10^{-3}, and 10^{-4}, and with lower resolution (2048×2790). The $T_{\rm cool} = 10^{-4}$ curve is similar to the case without cooling at such resolution.

the disc self-gravity cannot lead to such significant changes in pitch angles, spirals in scattered-light images might show more dramatic effects with disc self-gravity which needs to be studied in future. Without the consideration of the self-gravity, the different pitch angles between scattered light and millimetre continuum images might be due to the vertical temperature gradient (Juhász & G. 2018; Rosotti et al. 2019). The near-infrared scattered-light observations probe higher atmosphere with higher temperature, thus the spirals have larger pitch angles than those in the mid-plane probed by ALMA.

4.1.1 *Comparison to the linear theory*

In this subsection, we provide a quantitative comparison with the linear theory. In Section 2, we briefly introduce the AMF in discs with and without self-gravity. The quantity we want to compare is $F_J(m, Q)/F_J^{WKB}(m)$ (ratio of the numerically integrated m harmonics of the AMF and the m mode AMF solved analytically; see equation 27), which has been given in fig. 2 of Goldreich & Tremaine (1980). Similar comparisons have been performed before in non-self-gravitating shearing sheet simulations (Dong et al. 2011b). With our simulations, we measure the AMF at several h away from a low-mass planet ($M_{\rm p} \ll M_{\rm th}$) in the inner disc. Note that the spirals will shock at a certain distance away from the planet. This distance is $l_{\rm sh}$ (Goodman & Rafikov 2001), given by

$$l_{\rm sh} \approx 0.8\left(\frac{\gamma+1}{12/5}\frac{M_{\rm p}}{M_{\rm th}}\right)^{-2/5} h. \tag{42}$$

The linear theory of wave propagation fails for $|r - r_{\rm p}| > l_{\rm sh}$, so the position that we measure the AMF cannot be too far from the planet. The position should also be chosen closer to the planet as the planet mass increases, as $l_{\rm sh}$ becomes shorter with higher mass planets.

We start the comparison with a very low planet mass, $M_{\rm p} = 0.01$ $M_{\rm th}$ in `B18SG` discs. Fig. 3 shows the AMF measured at $5h$, where we find the values match the linear theory the best. The blue curve shows the ratio of the non-self-gravitating disc, and the orange curve shows the ratio of the $Q = 2$ disc. The ratios calculated

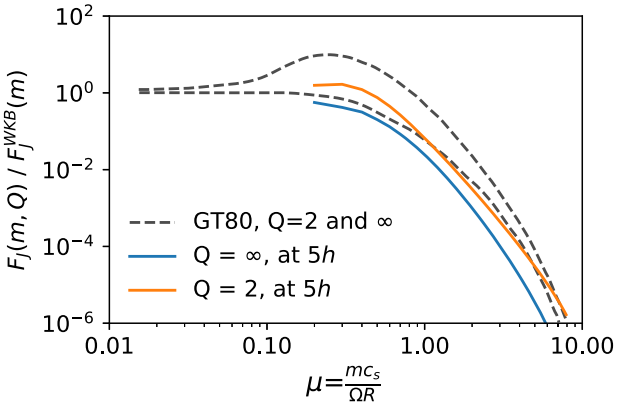

Figure 3. Comparison between $F_J(m, Q)/F_J^{WKB}(m)$ measured from simulations ($M_p = 0.01\ M_{th}$, B18SG) and calculated from the linear theory (Goldreich & Tremaine 1980). This ratio is measured at $5h$ (at $0.5\ r_p$) inside of the planet and is plotted on y-axis. The x-axis is $\mu = \frac{mc_s}{\Omega R}$ (or $m(h/r)$, where $(h/r)_p = 0.1$). The simulation results are shown in blue ($Q = \infty$) and orange ($Q = 2$) solid curves, whereas the analytical results are shown in dashed curves ($Q = 2$ curve is the higher).

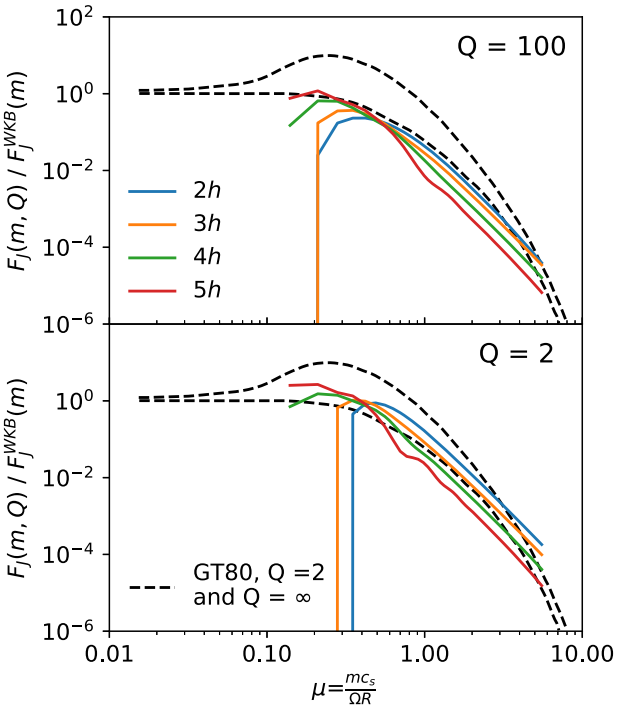

Figure 4. Similar to Fig. 3, but for B17SG, $M_p = 0.1\ M_{th}$. This ratio is measured at $2h$ (blue), $3h$ (orange), $4h$ (green), and $5h$ (red) inside of the planet. The upper panel shows the results in $Q = 100$ disc, whereas the lower panel shows the results in $Q = 2$ disc. The dashed curves are the same as in Fig. 3.

from Goldreich & Tremaine (1980) are shown in dashed curves (the ratio of $Q = 2$, $F_J(m, 2)/F_J^{WKB}(m)$ is shown above and that of $Q = \infty$, $F_J(m, \infty)/F_J^{WKB}(m)$ is shown below). The $Q = \infty$ curve matches the linear calculation relatively well, especially at $\mu \lesssim 1$. As for the differences, one of the reasons is that the temperature is assumed to be constant throughout the disc in the linear calculation, whereas $T \propto r^{-0.5}$ in the simulations. $F_J(m, Q)/F_J^{WKB}(m)$ of the self-gravitating disc is higher than that of $Q = \infty$ disc, which is consistent with the linear theory.

To make a more proper comparison with the analytical theory, we also present B17SG discs that are globally isothermal with $M_p = 0.1\ M_{th}$. $F_J(m, Q)/F_J^{WKB}(m)$ at different distances from the planet are shown in Fig. 4 (Top panel: $Q = 100$ disc, Bottom panel: $Q = 2$ disc). Similar to the results of isothermal shearing sheet simulations in Dong et al. (2011b), we also find a close agreement with the linear theory when $Q \to \infty$. However, the values in $Q = 2$ disc are still different from the linear theory expectation. They are lower by one order of magnitude when $\mu \lesssim 1$ and higher than the linear theory prediction at $\mu \sim 6$. The discrepancy in the high m (or high μ) is possibly due to the limited resolution in the azimuthal direction in simulations. We need very high resolutions to resolve the high m modes properly with a much smaller self-gravity smoothing kernel. As for the low m modes, it is actually not surprising to see discrepancy, since we only measure the advective AMF, and ignore the gravitational AMF (see Section 2.2). In $Q = 2$ discs, F_G should contribute a significant amount of AMF. It should be more important when Q is even smaller (close to 1). To match the analytical result of total AMF, one has to include F_G. On the other hand, we do see the increases of advective AMF as the Q becomes smaller. This indicates that as the disc becomes more massive, both F_A and FG increase to contribute for the increase of total AMF, F_J. Being part of the AMF, F_A alone is enough to explain the change of the density perturbations, i.e. the increases of F_A leads to the increases of the density perturbations. It is not necessary to add F_G when we want to explain the density perturbations qualitatively.

4.1.2 *Gap opening*

In the linear theory, if the planet mass M_p is large enough (close to or higher than M_{th}), the planet can open gaps in the disc. However, the gap opening mass can be much lower, since the non-linear effects such as shock can occur even if the system starts in a linear regime (Goodman & Rafikov 2001; Rafikov 2002). The study of gap opening is very important since gaps are observable in both gaseous and dusty discs. They have been found ubiquitously in recent high-resolution ALMA observations (Huang et al. 2018; Long et al. 2018). Assuming these gaps are due to the planet, one might infer the planet mass from the gap width and depth (e.g. Zhang et al. 2018). Now we use higher mass planets to demonstrate the effect of the self-gravity on gap opening.

We first show the results of B17SG discs with $M_p = 0.3\ M_{th}$, where the evolution starts in linear regime, marginally opening gap for non-self-gravitating disc at longer time evolution. Figs 5 and 6 show the density perturbations in 2D and 1D at $t = 20\,t_p$ (we use the models whose $t_{grow} = 10\,t_p$ here). As the Q decreases, the gap becomes deeper. Pohl et al. (2015), Li & Li (2016) also find this effect for higher mass planets. The transition is the most evident from $Q = 5$ to $Q = 2$ discs, where the gap depth becomes a factor of 3 stronger at the major gap where the planet is located. The major gap is separated by the horseshoe region, thus resembles 'w' shape. The gap width changes only slightly comparing to the gap depth (see Section 5.2 for details). The secondary gap is around 0.5 to 0.6 r_p, the location of which does not change as the Q decreases.

To understand the density perturbations, we plot the AMF of each simulation in Fig. 7. The left-hand panel shows $M_p = 0.1\ M_{th}$ discs and right-hand panel shows $M_p = 0.3\ M_{th}$ discs. Given the same Q, the normalized AMF decreases as the planet mass increases, which is due to the increasing non-linear shock dissipation. This is

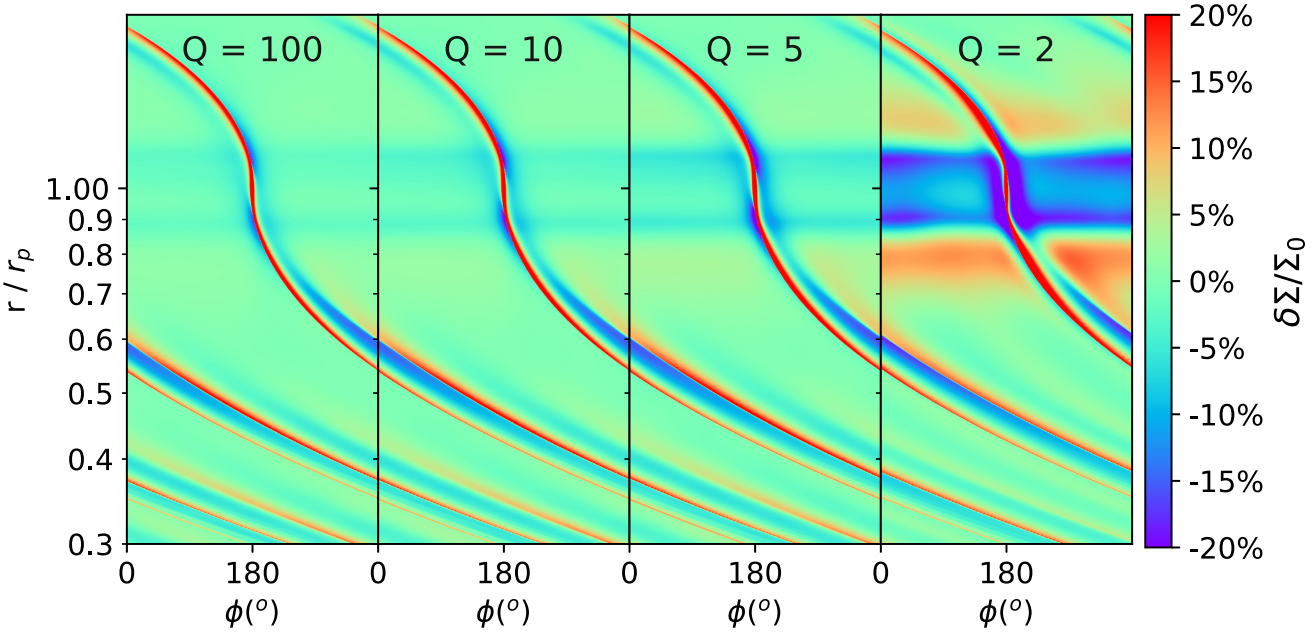

Figure 5. The gas density perturbations ($\delta\Sigma/\Sigma_0$) of the isothermal self-gravitating discs `B17SG` with a low mass planet $M_p = 0.3\ M_{th}$ at 20 orbits (with $t_{grow} = 10\,t_p$). Compared to Fig. 1, higher planet mass can open deeper gap and even the secondary gap (see Fig. 6). The depth of the primary gap becomes deeper with the decreases of Q. This is the most evident in $Q = 2$ disc.

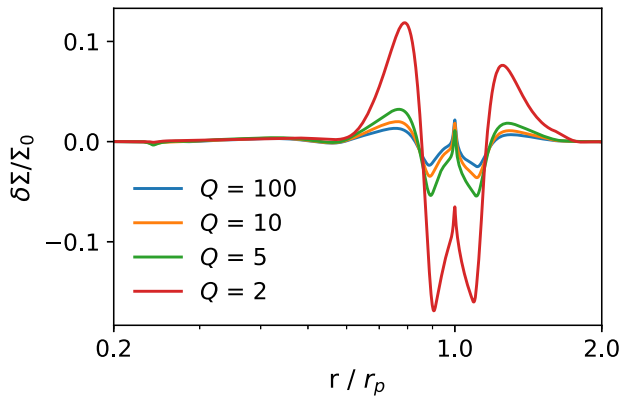

Figure 6. The azimuthally averaged density perturbations of Fig. 5. The blue, orange, green, and red curves represent $Q = 100$, 10, 5, and 2 discs. As the disc mass increases, the perturbations become stronger and gaps become deeper, but the position of the secondary gap does not change significantly.

consistent with the result of non-self-gravitating discs in Miranda & Rafikov (2019b). The changes of the AMF at each Q has the same trend as the changes in density perturbations. From $Q = 5$ to $Q = 2$, the AMF doubles. We will quantify this trend in Section 5.1. Since the normalization factor $F_{J0} \propto Mp^2$, the absolute value of AMF is still higher in the discs with higher planet mass.

To study how disc self-gravity affects the dust continuum observations, we evolve simulations for a longer time-scale (500 t_p) and plot the density perturbations of both gas and dust with different Stokes numbers in Fig. 8. From top to bottom, it shows the $\delta\Sigma/\Sigma_0$ of gas, and dust with $St = 0.001$ (represented by $St = [0.0005, 0.002]$), 0.01 (represented by $St = [0.005, 0.02]$) and 0.1 (represented by $St = [0.05, 0.2]$). Notice that the St_0 is the Stokes number at the beginning of the simulations. As the simulations evolve, the particle sizes are fixed, so the dust grain might have different St after it drifts. From left to right, the planet masses are $M_p = 0.1$, 0.3, and 1.0 M_{th}. The blue, orange, and green curves represent $Q = 100$, 10, and 5 discs. As expected, the perturbations in the dust are stronger than those in gas, and dust with higher Stokes number has larger perturbations. This is because dust particles with higher St drift faster when $St \lesssim 1$. For $M_p = 0.1\ M_{th}$, $St = 0.1$ disc, the dust particles have almost drifted to the inner disc. Similar to the trend in the gas, the gap becomes deeper, and the ring becomes higher as the Q becomes smaller, which are consistent with Li & Li (2016). The positions of the secondary gaps do not change. The overall shapes of the profiles do not change, except the gaps outside the planet become weaker for $M_p = 0.3$ and $1.0\ M_{th}$ discs at $St_0 = 0.001$. The positions of them also shift outside. The region beyond $r = 1.6\,r_p$ should not be trusted as it is in the wave damping zone. For small Q and large M_p, some curves for discs with $Q = 5$ are missing since the discs have already become unstable at $500\,t_p$. We will discuss this phenomenon in Section 5.4.

4.2 Adiabatic discs with radiative cooling

In this subsection, we study the effects of radiative cooling on planet–disc interactions, focusing on `B17AD`. The simulations presented in this section are globally isothermal, with $q = 0$. The absorption of AMF from the background occurs in locally isothermal disc (in Figs 1 and 2), but not in globally isothermal discs here. From left to right, Fig. 9 shows the density perturbations for the $M_p = 0.1\ M_{th}$ isothermal disc, and adiabatic discs with cooling time $T_{cool} = 0.01$, 0.1, 1, 10, and 100. The top panels are results from FARGO-ADSG, whereas the bottom panels are results from `Athena++`. They are almost identical, except for the small difference for the corotation features in the $T_{cool} = 100$ cases. We analyse the simulations at $t = 40\ t_p$ for FARGO and 20 t_p for `Athena++` due to a shorter t_{grow} adopted in `Athena++` simulations. At these times, we find that all the discs are stable and have reached quasi-steady states. Discs with shorter dimensionless cooling times T_{cool} are more similar to isothermal

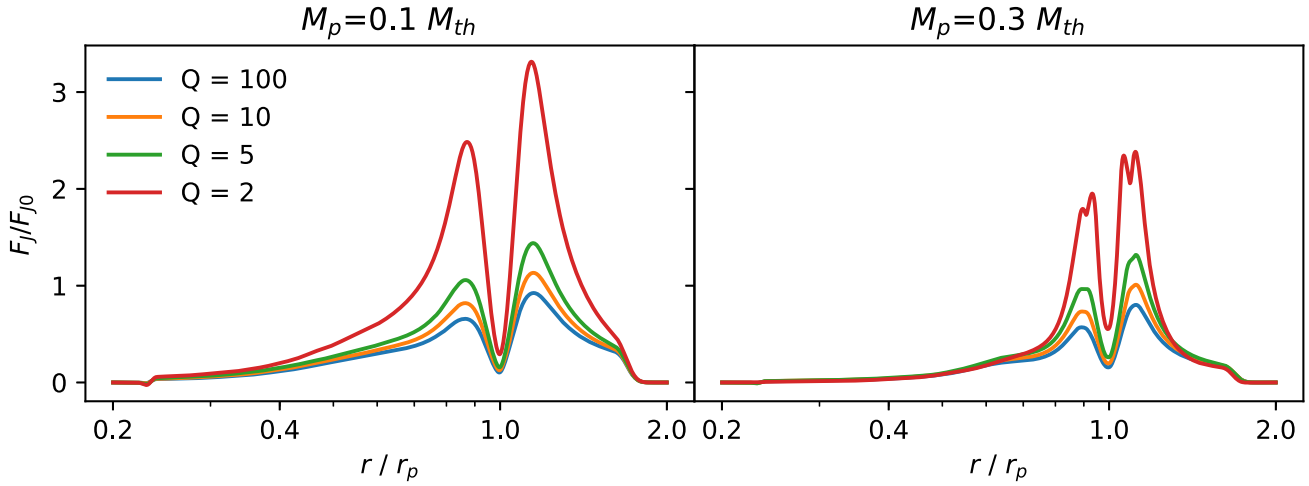

Figure 7. The AMF of the isothermal discs at 20 orbits with $M_p = 0.1$ and 0.3 M_{th}. Given the same disc mass, as the planet mass increases, the normalized AMF becomes lower at any disc with different Q, as shock forms earlier and closer to the planet with the increases of the planet mass.

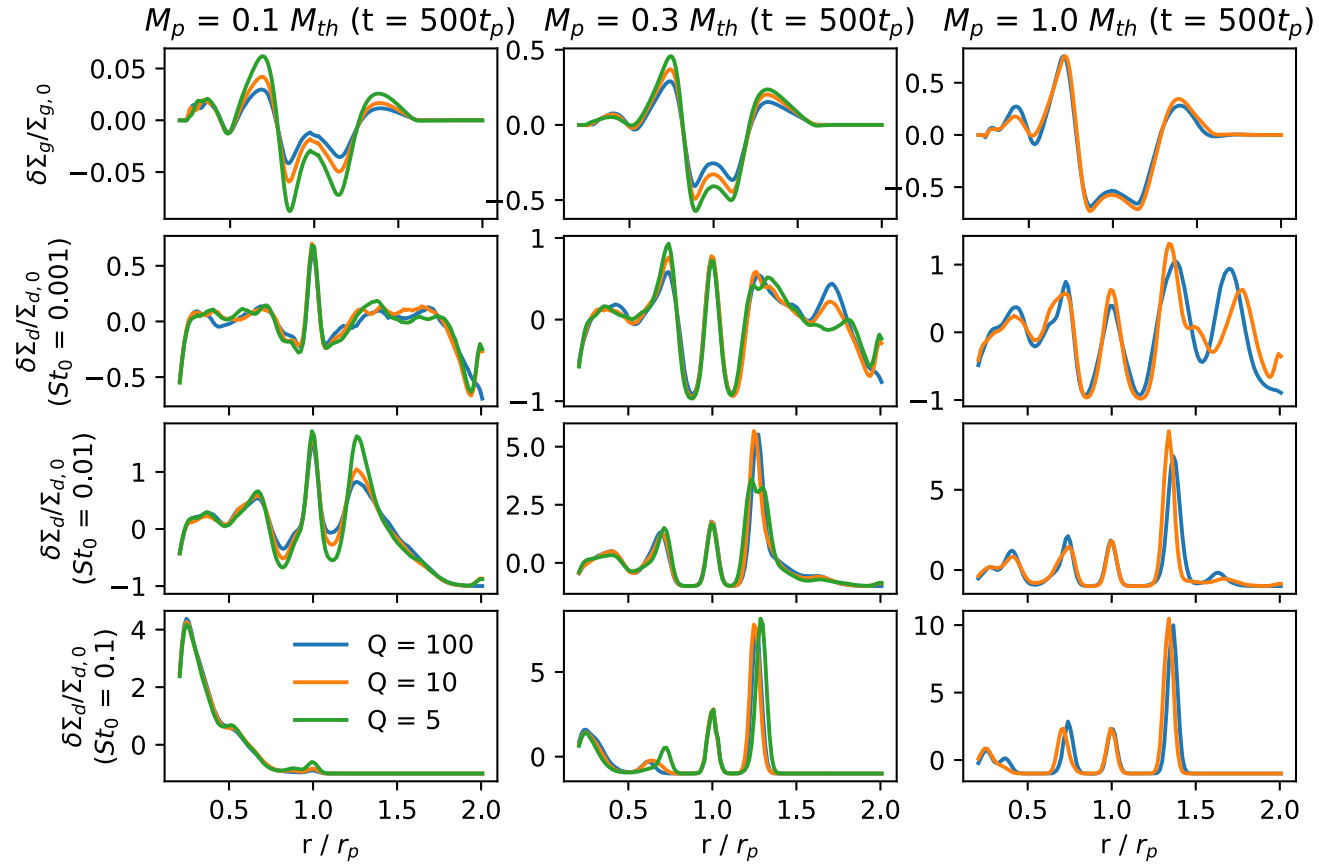

Figure 8. The density perturbations of the gas and dusts with different Stokes number ($St = 0.001$, 0.01, and 0.1) at $t = 500\ t_p$ in globally isothermal disc with $M_p = 0.1$, 0.3, and 1 M_{th} (from left to right). From top to bottom panels, there are density perturbations of the gas, $St = 0.001$, $St = 0.01$ and $St = 0.1$ sized dusts. The blue, orange, and green curves represent $Q = 100$, 10, and 5 discs. Some curves for discs with $Q = 5$ are missing since the disc becomes unstable at this time-steps in the simulations. Overall, the position of the secondary gap does not change with the increase of disc masses. The gap becomes deeper for lower Q values.

discs, whereas discs with larger T_{cool} resemble adiabatic discs. The primary spirals are marked as 'P', the secondary spirals are marked as 'S', and the tertiary spirals are marked as 'T' on the figure.

From the isothermal to the adiabatic limit (increasing T_{cool} from the left to the right-hand panels), the spirals become weaker as the cooling time-scale becomes longer until $T_{cool} = 1$ and then become stronger again as T_{cool} continues to increase. The secondary and tertiary spirals also become very weak at $T_{cool} = 1$ and stronger at $T_{cool} = 10$. Nevertheless, the secondary and tertiary spirals are still present at $T_{cool} = 1$ as indicated by the arrows.

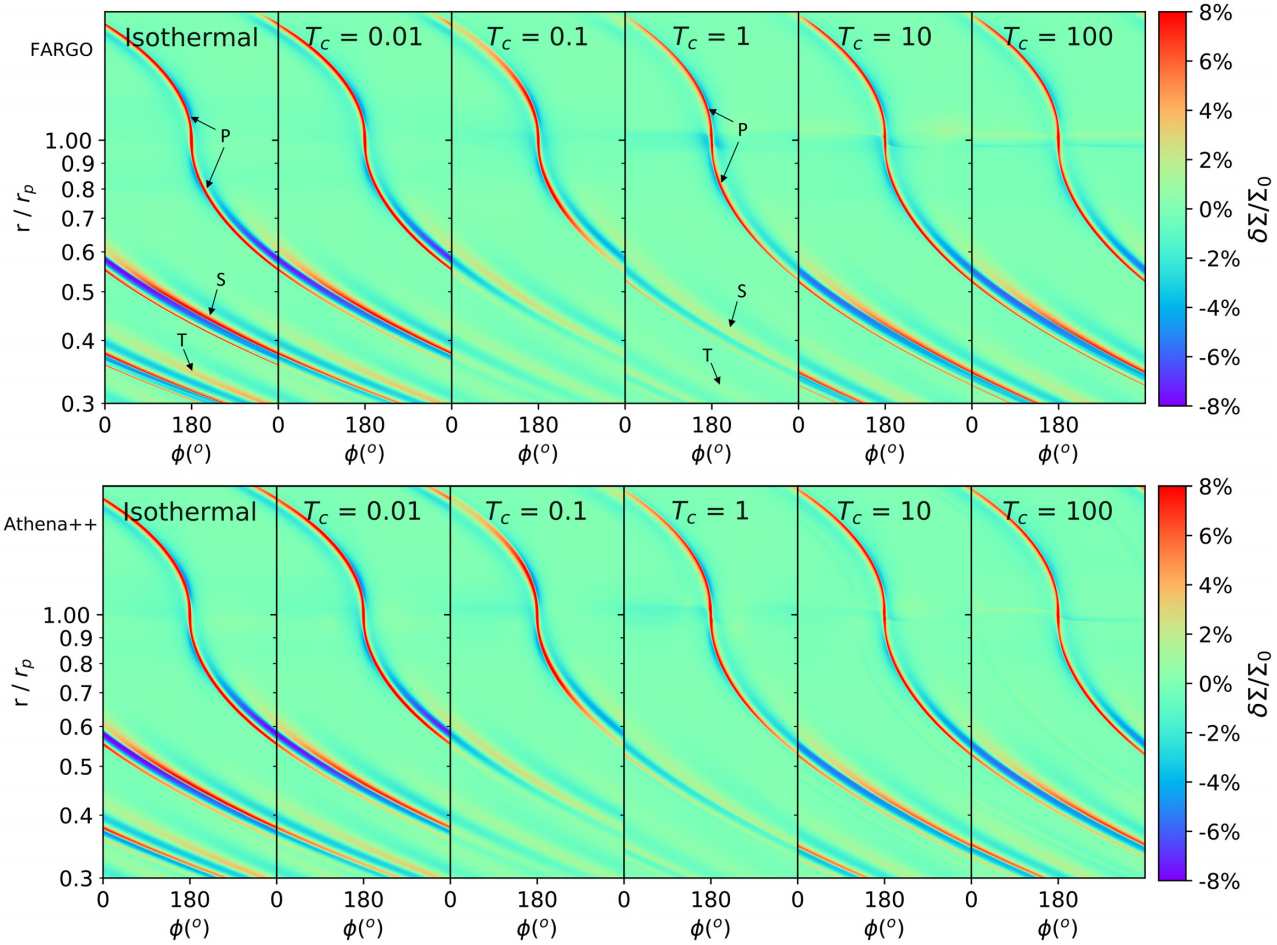

Figure 9. The density perturbations of the adiabatic simulations `B17AD` at 40 orbits (FARGO) and 20 orbits (`Athena++`) with $M_p = 0.1\ M_{\rm th}$. From left to right, the discs represent fast to slow cooling. The isothermal dick is shown in the leftmost panel, then the dimensionless disc cooling parameters $T_{\rm cool}$ are 0.01, 0.1, 1, 10, and 100, respectively. The upper panels show results from FARGO-ADSG and the lower panels show results from `Athena++`. The primary spirals are marked as 'P', the secondary spirals are marked as 'S', and the tertiary spirals are marked as 'T'.

The spirals become more open as the cooling time increases. This is due to the increases of the sound speed from $c_s = (RT/\mu)^{1/2}$ in the isothermal limit to $c_s = (\gamma RT/\mu)^{1/2}$ in the adiabatic limit. In other words, the effective γ in equation (30) increases from 1 to 1.4 as $T_{\rm cool}$ becomes larger.

The change of the amplitude and position of the spiral arms with respect to the cooling time can be seen more clearly in 1D plots. Fig. 10 (left-hand panels) shows the density perturbations versus azimuthal angle (ϕ) cut at $r = 0.45\ r_{\rm p}$ and $r = 0.32\ r_{\rm p}$. The density perturbations due to the primary spiral and the secondary spiral are marked as 'P' and 'S' on the figure. The one due to the tertiary spiral is also marked as 'T' in the $r = 0.32\ r_{\rm p}$ plot. The positions of the primary, secondary, and tertiary spirals in $T_{\rm cool} = 0.01$ discs are similar to those in the isothermal disc. At $r = 0.45\ r_{\rm p}$, the amplitude of the spiral in the $T_{\rm cool} = 0.01$ disc is slightly lower than that in the isothermal disc

There is a significant change of the spiral's amplitude when $T_{\rm cool}$ increases from $\lesssim$0.01 to 0.1. The spiral's amplitude becomes much smaller at $T_{\rm cool} = 0.1$ with a slight shift of the primary spiral position. The tertiary spiral almost disappears at $r = 0.32\ r_p$, but we can still see its presence with a amplitude of $\sim$ 1 per cent. But the biggest change of the spiral's position occurs when $T_{\rm cool}$ increases from 0.1 to 1. At $r = 0.45\ r_{\rm p}$, the primary spiral shifts $\sim 45^\circ$ between $T_{\rm cool} = 0.1$ to $T_{\rm cool} = 1$ cases, although the amplitude does not change much. The tertiary spiral at $r = 0.32\ r_{\rm p}$ is now the hump from 160° to 250° with a amplitude of $\sim$ 0.5 per cent. At $T_{\rm cool} = 10$, the amplitude of the spiral becomes much higher. The positions are very similar to those in $T_{\rm cool} = 1$. The tertiary spiral also becomes stronger. But the spirals are still weaker than those in the isothermal disc. From $T_{\rm cool} = 10$ to $T_{\rm cool} = 100$ discs, the perturbations seem to converge to the limit of an adiabatic disc. Overall, from isothermal discs ($T_{\rm cool} \lesssim 0.01$) to adiabatic discs ($T_{\rm cool} \gtrsim 10$) the azimuthal angles of the spirals shift to smaller values by $\sim 50^\circ$ at $r = 0.45\ r_{\rm p}$. This is consistent with the finding in 2D plots that the spirals become less tightly wound as $T_{\rm cool}$ increases.[2]

Since spirals are directly related to gap opening, we want to study how the properties of the induced gaps are related to radiative cooling. The top panel of Fig. 11 shows the azimuthally averaged density perturbations of discs with different cooling times. $T_{\rm cool} = 0.01$ disc has a similar profile to the isothermal disc, but with

[2]Aside from the interesting behaviour at $T_{\rm cool} \sim 1$, we want to note that at two extremes, the spiral at the adiabatic limit is still weaker than that at the locally isothermal limit as expected. This is related to the increase of the effective sound speed with a longer cooling time. When the sound speed increases, the thermal mass increases and ratio between the planet mass and the thermal mass decreases so that the planet is less capable of exciting spiral waves.

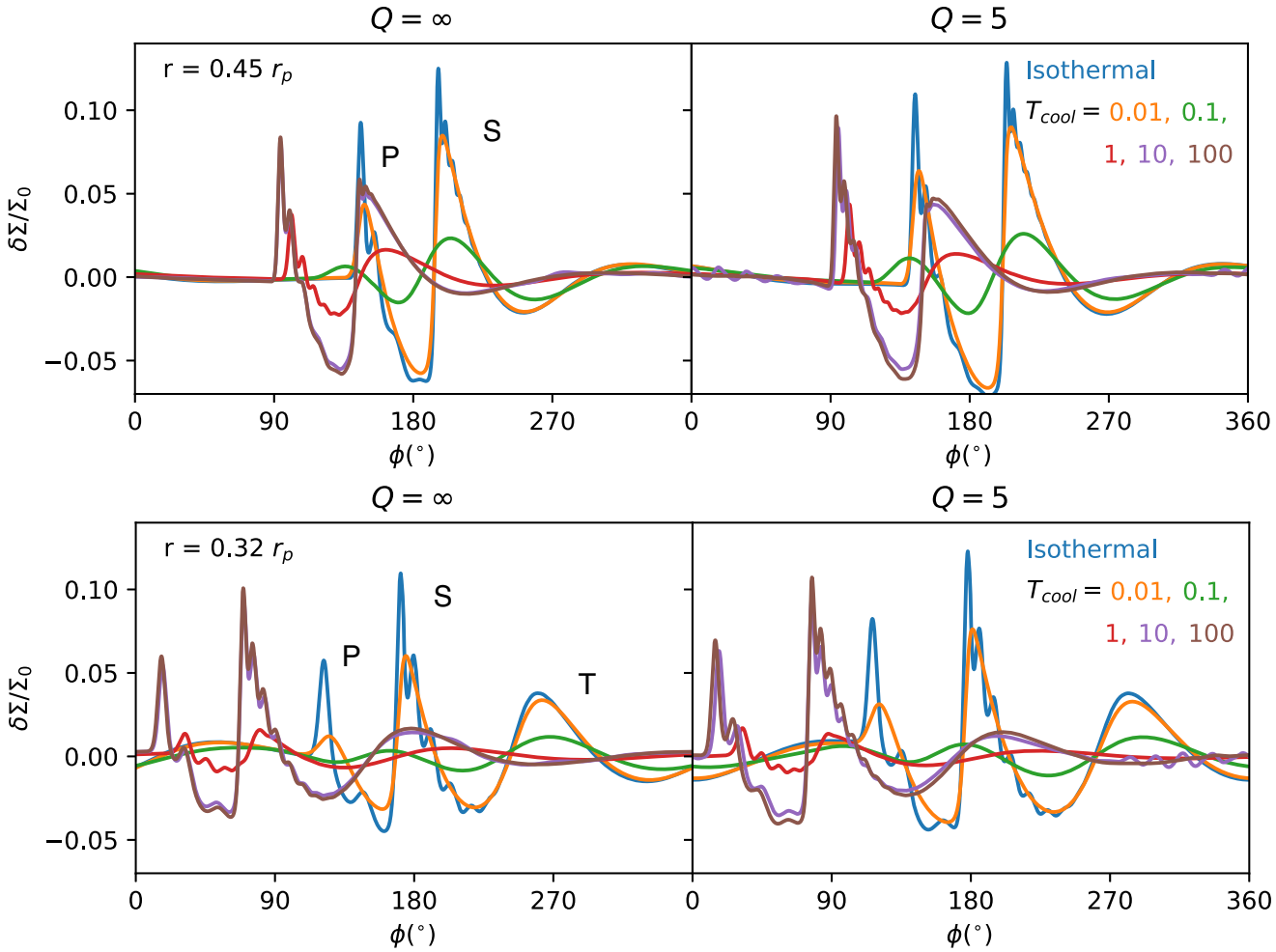

Figure 10. The density perturbations along $r = 0.45\,r_{\rm p}$ (upper panels) and $r = 0.32\,r_{\rm p}$ (lower panels) for `B17AD` (left-hand panels) and `B17ADSG` (right-hand panels) discs at $t = 40\,t_{\rm p}$. The positions of the primary, secondary, and tertiary spirals for small $T_{\rm cool}$ are marked as 'P', 'S', and 'T' on the left-hand panels.

smaller amplitudes. The secondary gap at $T_{\rm cool} = 0.1$ and 1 disappears (the disappearance of the gap in gas does not necessarily lead to that in dust; see Section 5.3). It reappears at $T_{\rm cool} = 10$, but the position changes from 0.5 $r_{\rm p}$ to 0.4 $r_{\rm p}$ and the amplitude becomes smaller. Since spirals carve gaps in discs, the inward shift of the secondary gap position is due to the more open secondary spiral for higher $T_{\rm cool}$. On the other hand, the non-monotonically transition – the trend of high, low, and high amplitudes of the secondary gap depth from the fast cooling to the slow cooling cases – is due to weaker spirals at intermediate $T_{\rm cool}$. Overall, for the intermediate massive planet ($M_{\rm p} \sim 0.1\,M_{\rm th}$), discs with $T_{\rm cool} \lesssim 0.01$ can be treated as isothermal discs and discs with $T_{\rm cool} \gtrsim 10$ can be seen as purely adiabatic discs, whereas the discs undergo a sharp transition between these limits, characterized by low amplitude spirals and shallow secondary gaps. Note that this $T_{\rm cool} \lesssim 0.01$ criterion is more stringent with a low mass planet in the linear regime. As shown in Fig. 2, only when $T_{\rm cool} = 10^{-4}$ the AMF is close to the isothermal case with $M_{\rm p} = 0.01\,M_{\rm th}$. Both the temperature gradient ($\mathrm{T} \propto r^{-0.5}$) and the linear damping contribute to the AMF profile, but the linear damping is the dominant effect here.

Since such conditions with intermediate cooling times are very likely to occur in realistic protoplanetary discs (see Section 2), we are interested in whether these much shallower secondary gaps can still trap enough dust particles to be observable. Examining the previous models in Zhang et al. (2018), we find the current set of simulations has very similar setup to the $h/r = 0.07$, $\alpha = 10^{-4}$, $M_{\rm p}/M_* = 11 M_\oplus/M_\odot$ model in Zhang et al. (2018) (the upper-left plots in 2D figures therein), despite the α is lower here. However, the secondary gap cannot be found in any dust configurations in that model, even though those simulations assume the locally isothermal EoS. Thus, we conclude that the models being analysed in this subsection with such low planet mass is not suitable for testing the observability of the secondary gap. To that end, we run simulations with higher planet mass, 1 $M_{\rm th}$, close to the setup of the AS 209 simulation, α varying model (Zhang et al. 2018). We direct the discussion to Section 5.3.

We also plot the AMF (in solid curves) and the integrated torque (in dashed curves) of these discs in the lower panel of Fig. 11. The torque is integrated from the planet's position towards either side of the disc. The AMF decreases monotonously close to the planet ($|r - r_{\rm p}| \lesssim 0.1\,r_{\rm p}$) due to the torque cutoff. The AMF far away from the planet (e.g. at $r = 0.3\,r_{\rm p}$ and $1.5\,r_{\rm p}$) decreases to the lowest values at $T_{\rm cool} = 0.1$ and 1 and increases back to the isothermal values at $T_{\rm cool} = 10$ and 100. On the other hand, from the fast cooling (and isothermal) to the slow cooling cases, the torque first increases and reaches the maximum at $T_{\rm cool} = 1$ and then decreases

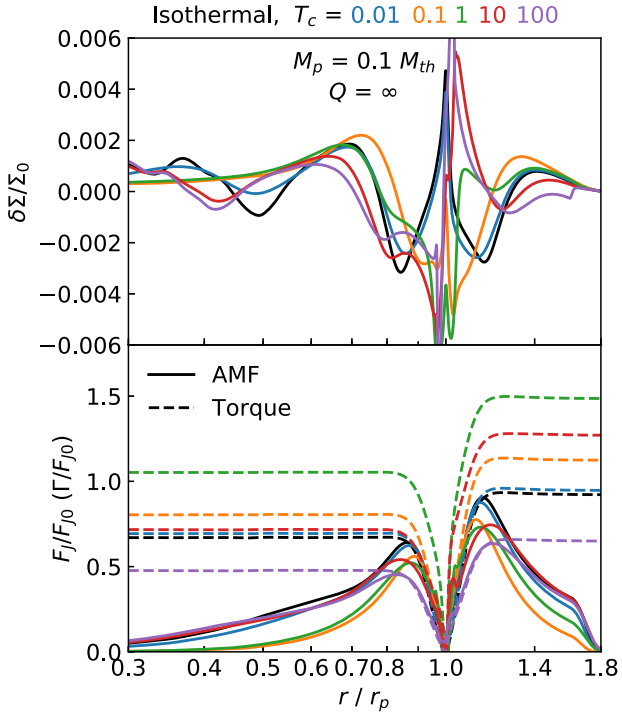

Figure 11. Top panel: The azimuthally averaged density perturbations in Fig. 9, upper panels. Bottom panel: The normalized AMF (solid curves) and torque (dashed curves) for different cooling times, $T_{\rm cool}$. The black curves represent isothermal disc, whereas blue, orange, green, red, and purple curves represent adiabatic discs with $T_{\rm cool}$ = 0.01, 0.1, 1, 10, and 100, respectively. The changes of the AMF distribution should be responsible for the change of the density perturbations.

and reaches the minimum at $T_{\rm cool}$ = 100. The AMF and torque at several h away from the planet (e.g. at r = 0.85 $r_{\rm p}$ or 1.2 $r_{\rm p}$) have similar values in both the very small (e.g. $T_{\rm cool}$ = 0.01) and very large (e.g. $T_{\rm cool}$ = 100) $T_{\rm cool}$ cases. However, the torque there is much higher than the AMF at $T_{\rm cool}$ = 0.1, 1, and 10. This indicates that, for intermediate cooling cases, the waves fail to launch or are damped right after they are launched, indicating radiative cooling is important for these cases. Note that the normalization (characteristic value, F_{J0}) should have varied with different cooling times, since $F_{J0} \propto h_p^{-3}$. In adiabatic EoS, the sound speed, $c_s = \left(\gamma \frac{dp}{d\rho}\right)^{1/2}$ and the scale height $h \propto c_s \propto \gamma^{1/2}$. Thus, $F_J \propto \gamma^{-3/2}$. The effective normalization is 0.6 times the F_{J0} in equation (6) at the slow cooling limit when γ = 1.4. This value should be in between 1 and 0.6 in transition from small to large $T_{\rm cool}$. Nevertheless, we just use F_{J0} as equation (6) and compare the absolute values of the AMF. We will have more detailed discussions in Section 5.1.

4.3 Disks with both self-gravity and radiative cooling

We also explore a situation that both self-gravity and radiative cooling are included, even though we do not seek to carry out a parameter space study. The parameters we choose are Q = 5, `B17ADSG` discs with various $T_{\rm cool}$. The density perturbations at two radii are shown in Fig. 10's right-hand panels. The profiles of the spirals are very similar to the non-self-gravitating ones on the left-hand panels. All the trends found in Section 4.2 still apply. At the same cooling time, the perturbation amplitudes become stronger and the spirals become tighter if Q becomes smaller, which are consistent with the results in Section 4.1. This means that when both of these two physical processes (disk self-gravity and radiative cooling) are operating, they play similar roles as those that they play individually, at least for low mass planets. From the magnitudes of the changes due to these two processes, we can conclude that when the disc is not too massive ($Q \gtrsim 5$), the effects of the self-gravity on the discs are less important than those of the cooling.

5 DISCUSSION

5.1 AMF versus Q and $T_{\rm cool}$

To quantify the relationships between AMF versus Q and AMF versus $T_{\rm cool}$, we measure the AMF at given positions in Figs 7 and 11 for `B17SG` and `B17AD` discs with $M_{\rm p}$ = 0.1 $M_{\rm th}$. We plot the AMF as a function of Q (the left-hand panel) and $T_{\rm cool}$ (the right-hand panel) in Fig. 12. The AMF is measured at 0.87 $r_{\rm p}$ for `B17SG` simulations, where it reaches the maximum at the inner disc. The AMF increases as the disc becomes more massive (with a smaller Q). This increase becomes quite fast when Q is approaching 2. The advective AMF in the Q = 2 simulation is almost five times the AMF in the Q = 100 simulation, which is roughly consistent with the analytical expectation (equation 26).

As for the radiative cooling, we measure the AMF at 0.9 $r_{\rm p}$ and 0.6 $r_{\rm p}$. At the position very close to the planet (e.g. 0.9 $r_{\rm p}$), the AMF decreases as the cooling time increases. As mentioned in Section 4.2, this is most likely due to the transition of the effective γ (that enters the sound speed) from 1 to 1.4. The ratio between the AMF at $T_{\rm cool}$ = 100 and that at $T_{\rm cool}$ = 0.01 is ∼0.6, which is very close to the expected change due to our definition of F_{J0}. However, if the AMF is measured farther away from the planet (e.g. 0.6 $r_{\rm p}$), it first decreases and then increases as $T_{\rm cool}$ increases, reaching the minima at 0.1 and 1 $T_{\rm cool}$. The decrease of the AMF with the distance away from the planet is due to the wave damping as the wave propagates. When the wave propagates to 0.6 $r_{\rm p}$, the damping is the strongest at $T_{\rm cool} \sim 0.1 - 1$, implying that radiative cooling plays an important role on the wave damping/dissipation when $T_{\rm cool} \lesssim 1$. We are expecting a stronger wave damping with a smaller $T_{\rm cool}$. However, when $T_{\rm cool}$ becomes very small and approaches the isothermal limit, the temperature of the wave becomes the background disc temperature, and shock damping in this case behaves similar to the shock damping in the adiabatic limit (Goodman & Rafikov 2001; Dong, Rafikov & Stone 2011a). At the position that the secondary spiral should form, the wave becomes already too weak to perturb the gas for the $T_{\rm cool}$ = 0.1 − 1 cases. Thus, the secondary gap can even disappear at these intermediate cooling times.

5.2 Implications for observations

To quantify the properties of the primary gap, we measure the gap width and depth in the gas density perturbation maps for `B17SG` and `B17AD`, $M_{\rm p}$ = 0.1 $M_{\rm th}$, using the definitions in Zhang et al. (2018). Different from Zhang et al. (2018), we only measure the gaps at ∼50 orbits and from the gas density perturbation profiles, because we do not run the simulations for very long time and the planet mass is still too small to open significant gaps in gas and dust. However, these measurements suffice to demonstrate the trend.

In the upper panels of Fig. 13, the gap width–Q relation is shown on the left-hand panel, whereas the depth–Q relation is shown on the right-hand panel. Both the x and y axes are in logarithmic scales. Since the exponents in the gap depth relations are ∼5 times larger than that of gap width relations (see Table 1 and 2 in Zhang

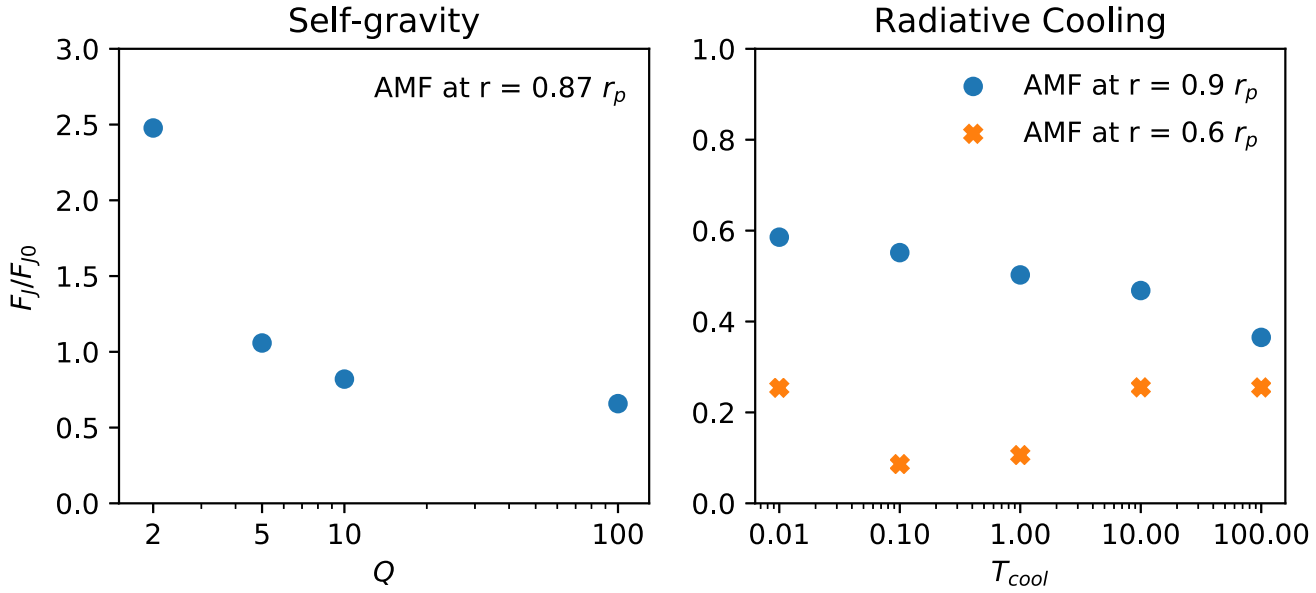

Figure 12. AMF versus Q (left-hand panel) and AMF versus $T_{\rm cool}$ (right-hand panel). The AMF is measured at $r = 0.87\,r_{\rm p}$ for the SG runs and at $r = 0.6\,r_{\rm p}$ (marked as 'X's) and $0.9\,r_{\rm p}$ (marked as dots) for the AD runs.

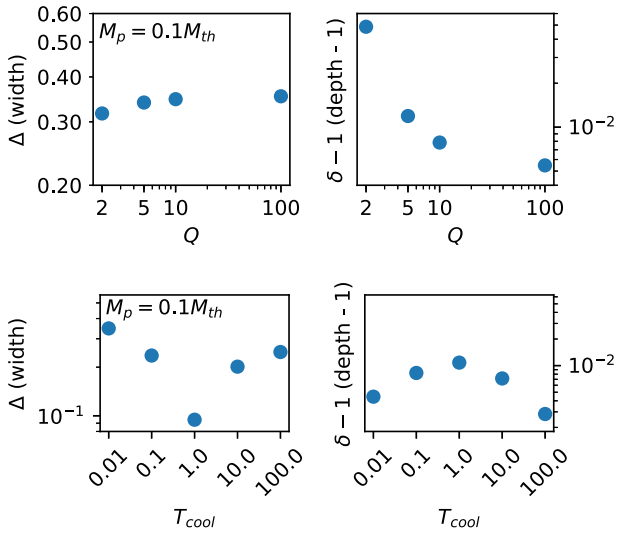

Figure 13. Gap width/depth versus Q (top) and Gap width/depth versus $T_{\rm cool}$ (bottom). Δ and δ - 1 are proportional to $M_{\rm p}$. The range for the gap (depth - 1) is five times than that in the width. Under this factor, the y-axes on the left-hand and right-hand panels have approximately the same scaling to $M_{\rm p}$. Thus, the y-axes of the left-hand and right-hand panels can be directly taken as the planet mass in an arbitrary unit.

et al. 2018 for details), the plotted y-axis range for the depth ($\delta - 1$) is 5 times of the range for the width (Δ). In this way the relative magnitude read on both y-axes is roughly proportional to the planet mass ratio. As Q changes to lower values, the gap depth has more significant changes than the width. The planet mass inferred from the gap depth without considering disc self-gravity will be overestimated. Adopting common exponents 0.25 and 1.25 for the width–$M_{\rm p}$ and depth–$M_{\rm p}$ relations (more precisely, width–K' and depth–K in Zhang et al. 2018), the ratio of the inferred masses between $Q = 2$ and $Q = 100$ discs is 1.5 using the width–$M_{\rm p}$ relation, and 6 using the depth–$M_{\rm p}$ relation. Thus, if one wants to infer the planet mass, the gap width is a better observable than the gap depth since it is less sensitive to the disc mass.

The lower panels of Fig. 13 show the width and depth as functions of the $T_{\rm cool}$. The gap width and depth are not very well defined due to the irregular gap shape. We try to neglect the horseshoe region around $r_{\rm p}$ when we measure the gap width. It is better to use the gap depth as a proxy for the planet mass. If one uses the gap width to measure the planet mass in locally isothermal simulations, the planet mass will be underestimated if $T_{\rm cool} \gtrsim 0.01$.

Note that these trends found in the low planet mass regime do not necessarily apply to the high planet mass regime, as we will discuss in the next subsection, where we find much less change in the gap shape for various cooling times.

5.3 AS 209: cooling in the high planet mass regime

We have two goals in this subsection. One is to explore the effects of radiative cooling on the gap properties when the planet mass is relatively large reaching the thermal mass. Meanwhile, we decide to choose a model that resembles a realistic disc so that we are able to constrain its surface density via estimating its $T_{\rm cool}$. Of all the DSHARP discs, AS 209 features many intricate gaps and rings (Guzmán et al. 2018). Using simple (without radiative cooling and self-gravity) hydrodynamical simulations, Zhang et al. (2018) find an excellent match to the AS 209 disc with a model whose $(h/r)_{\rm p} = 0.05$, $\alpha = 3 \times 10^{-4}(r/r_{\rm p})^2$, and $M_{\rm p}/M_* = 10^{-4}$, or 1 $M_{\rm th}$.[3] The dust size distribution is a power law $n(a) \propto a^{-3.5}$, $a_{\rm max} = 0.68$ mm, and the gas surface density $\Sigma_{g,0} = 6.4\,{\rm g\,cm^{-1}}$. This model can explain up to five gaps in the disc, not only their locations, but also the intensity. Since we have found that self-gravity and radiative cooling have the potential to change the gap shape and the secondary gap position in the low planet mass regime, we would like to explore

[3]We also run simulations with constant α as the model (b) in Zhang et al. (2018). Same as previous models, they reproduce AS 209's 1D profile, but are non-axisymmetric in 2D. We do not see more differences due to the viscosity profile.

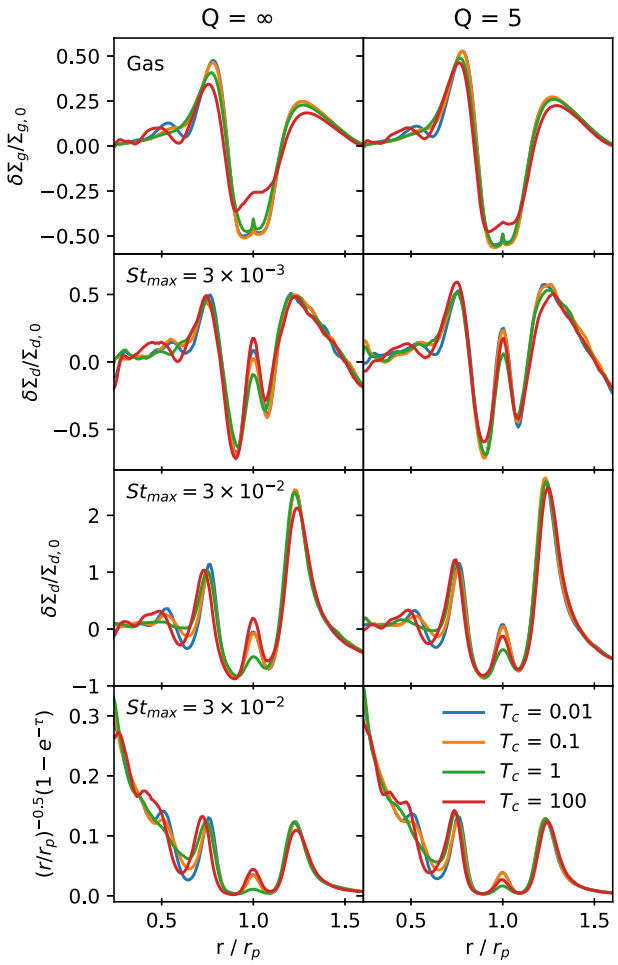

Figure 14. Azimuthally averaged density perturbations and brightness temperature of AS209AD and AS209ADSG at 400 t_p. From top to bottom, they are the density perturbations of the gas, dusts with $St_{max} = 3 \times 10^{-3}$ and $St_{max} = 3 \times 10^{-2}$. The bottom panels show the brightness temperature of the $St_{max} = 3 \times 10^{-2}$ ($a_{max} = 0.5$ mm) dust models. The left-hand panels show the cases without self-gravity, whereas the right panels show the cases with $Q = 5$. The blue, orange, green, and red curves represent $T_{cool} = 0.01$, 0.1, 1, and 100, respectively. The $T_{cool} = 0.01$ curves on the left-hand panels are almost identical to the locally isothermal cases in Zhang et al. (2018) (the slight difference is mainly due to the different setups in resolution, temperature profile, indirect term, and smoothing length).

how the planet mass for AS 209 will change considering these two physical processes.

Fig. 14 shows the azimuthally averaged density perturbation and brightness temperature of the AS209 models with different cooling times ($T_{cool} = 0.01$, 0.1, 1, and 100) and different disc masses ($Q = \infty$ and 5) at 400 t_p. We choose not to show $T_{cool} = 10$ case to avoid crowdedness of the figure, but its feature is similar to the $T_{cool} = 100$ case with shallower gaps in dust. The left-hand panels show the $Q = \infty$ cases, whereas the right-hand panels show the $Q = 5$ cases. The blue, orange, green, and red curves represent $T_{cool} = 0.01$, 1, 10, and 100 cases, respectively. From the top to bottom panels, it shows the density perturbations of the gas, the dust with $St_{max} = 3 \times 10^{-3}$, $St_{max} = 3 \times 10^{-2}$, and the brightness temperature of the $St_{max} = 3 \times 10^{-2}$ models. We assume that the temperature in the disc is proportional to $r^{-0.5}$ in the last row and the procedure to obtain the optical depth is detailed in Zhang et al. (2018). As the T_{cool} becomes larger, the perturbation at the major gap (where the planet locates) becomes shallower in gas. However, this is not the case in dust. The gap width and depth are almost the same across different cooling times. This result is different from that in the low planet mass regime, where the gap shapes are quite different across different cooling times. For dust profiles, the horseshoe region has the smallest density for $T_{cool} = 1$. The secondary gap disappears for $T_{cool} = 1$ in gas, but it is still noticeable in dust for $St_{max} = 3 \times 10^{-2}$, with a lower amplitude than $T_{cool} = 0.01$, 0.1, and 100 cases. This indicates that, due to the gaseous bump at the inner edge of the primary gap, dust particles can still be trapped there, forming the secondary gap in dust even if the secondary gap disappears in the $T_{cool} = 1$ case. The location of the secondary gap shifts from $\sim 0.65\, r_p$ to $\sim 0.6\, r_p$ when T_{cool} increases. The tertiary gap is also present in $T_{cool} = 0.01$ and 100 cases. It moves inwards from $\sim 0.43\, r_p$ to $\sim 0.37\, r_p$. Although the tertiary gap is not present in the $T_{cool} = 1$ case, a change of slope can been seen at $\sim 0.4\, r_p$ (this is more obvious in $Q = 5$ case). The $Q = 5$ discs are similar to the $Q = \infty$ discs but with slightly stronger perturbations. The differences between different T_{cool} curves are slightly smaller in $Q = 5$ discs than those in $Q = \infty$ discs.

Note that aside from the EoS, the simulation setups in this subsection are different from those in Zhang et al. (2018) in several aspects (e.g. the resolution, domain size, temperature profile, indirect term, smoothing length, and planet growing time). Thus, the gap shape and position are expected to be slightly different from the simulations in Zhang et al. (2018). Also, as time evolves, the position of the secondary and tertiary gaps will move slightly inwards. Thus, we do not try to compare these profiles with AS 209 observations. Instead, we use this example to demonstrate the observable differences due to different cooling times.

We also want to emphasize that without the self-gravity, the location of the planet and the gaseous surface density from the simulation can be scaled to any value. However, when the disc self-gravity is included, the gas density and the location of the planet are related. The relation is $\Sigma_{g,0} = (M_*/r_p^2)(h/r)_p/(\pi Q_p)$. Taking $M_* = 0.83\, M_\odot$, $r_p = 99$ au, $Q_p = 5$ and $(h/r)_p = 0.05$, we get $\Sigma_{g,0} = 2.4\, \mathrm{g\,cm^{-2}}$. In Zhang et al. (2018), the best-fitting model has $Q \approx 2$ at the planet's position in the initial condition.

Assuming the gaps at 24, 35, and 61 au in AS 209 are related to the planet at 99 au, we can use these gaps to constrain the AS 209 disc mass. Using equation (39), and adopting the stellar luminosity as $1.4\, L_\odot$ and the mass as $0.87\, M_\odot$ from Andrews et al. (2018), $\phi = 0.02$, $T_f = 20$ K, and $\kappa_R = 0.21\, \mathrm{cm^2 g^{-1}}$, we have

$$T_{cool} = 0.02\left[1 + 0.01\left(\frac{\Sigma}{1\ \mathrm{g\,cm^{-2}}}\right)^2\right]\begin{cases} 2.5 & r \le 46\ \mathrm{au} \\ \left(\frac{r}{99\ \mathrm{au}}\right)^{-3/2} & r > 46\ \mathrm{au} \end{cases}. \quad (43)$$

To produce an observable tertiary gap, we need low or high cooling times ($T_{cool} \lesssim 0.1$ or $T_{cool} \gtrsim 10$). However, $T_{cool} \gtrsim 10$ requires $\Sigma \gtrsim 200\, \mathrm{g\,cm^{-2}}$, which is impossible since Q would be $\ll 1$. If $T_{cool} \lesssim 0.1$, we can constrain $\Sigma \lesssim 20\, \mathrm{g\,cm^{-2}}$. While it is an independent constraint, this does not give tighter constraint on the density for this disc, since from gravitational instability criterion (equation 8), the disc will become unstable if $\Sigma_{g,0} > 12\, \mathrm{g\,cm^{-2}}$ at 99 au. Nevertheless, this constraint from the disc cooling is consistent with that from the disc instability.

The $Q = 5$ models with low T_{cool} might be plausible models for the AS 209 disc, since their density at 99 au is allowed given the constraints derived above. However, with the density $\Sigma_{g,0} = 2.4\, \mathrm{g\,cm^{-2}}$ from the $Q = 5$ disc, we cannot restore the intensity measured from the DSHARP observation. It is much lower than that from the observation. Only $\Sigma_{g,0} \gtrsim 6\, \mathrm{g\,cm^{-2}}$ can explain the

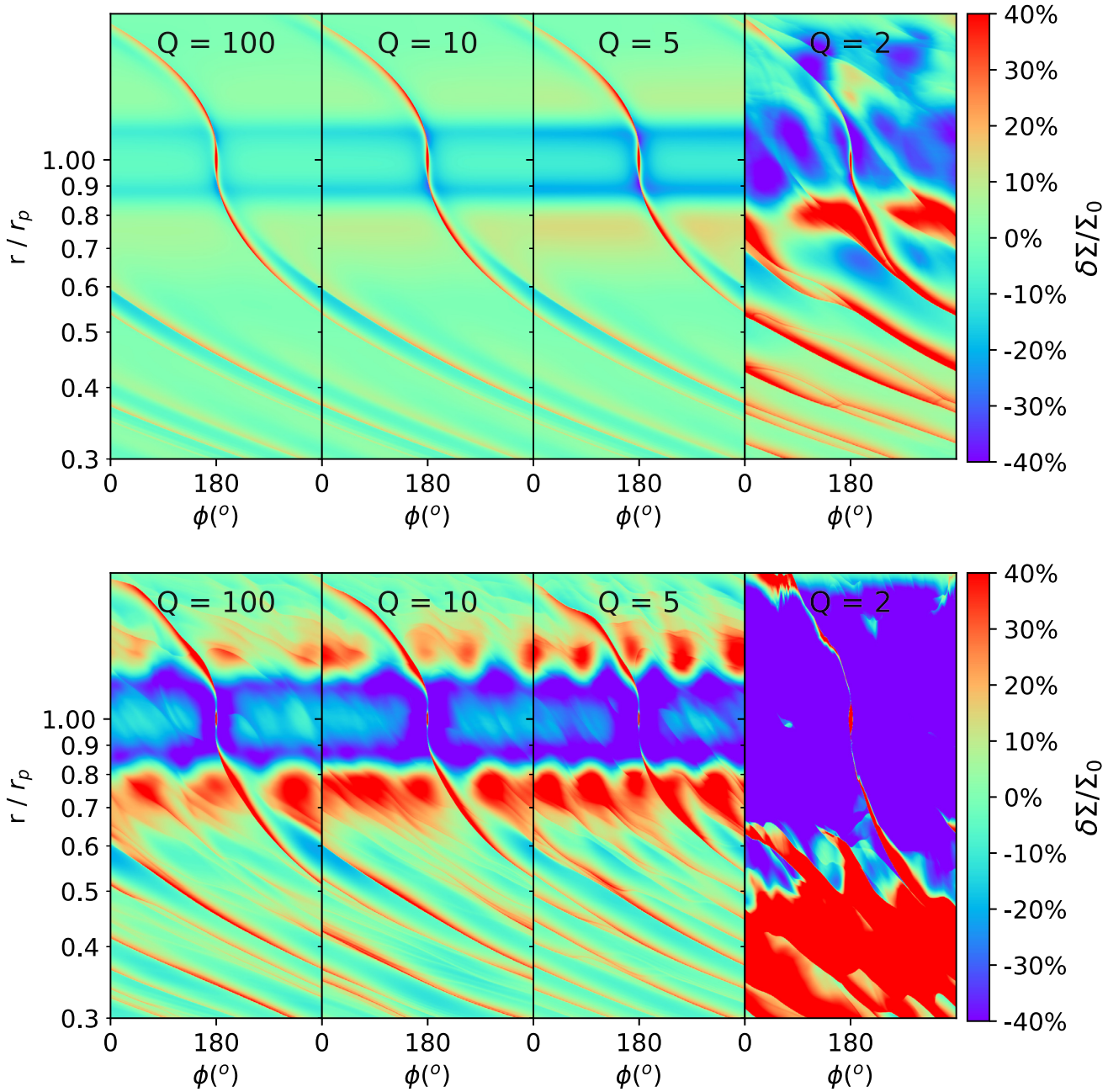

Figure 15. The density perturbations of `B17SG` discs with $M_p = 0.3$ and 1.0 M_{th} (top and bottom) at $t = 100\,t_p$. The disc becomes unstable at $Q = 2$. The lower panel shows that self-gravity can suppress small m vortices, so that the number of the vortices increases as Q becomes smaller.

intensity. This means that either the actual opacity is higher than what we adopt (e.g. the dust-to-gas mass ratio is larger) or the disc has even lower Q. If the latter is true, the T_{cool} can be larger (but still <1), since the low Q that is close to unity can produce much stronger perturbations, compensating for the weaker perturbations due to the larger value of the T_{cool}. Interestingly, Powell et al. (2019) estimate the gas surface density $\Sigma \approx 10\,g\,cm^{-2}$ at 100 au, using the 'dust line' method. This high surface density is actually compatible and consistent with the constraints from disc self-gravity and radiative cooling for the AS 209 disc. If this is the case, it indicates that the outer disc of the AS 209 disc is marginally gravitational stable.

5.4 Long time evolution for self-gravitating discs with planets

On the other hand, in our long time-scale simulations, we find that the self-gravitating discs with planets tend to become unstable at later times. This happens earlier if the Q is small or if the planet mass is high. Fig. 15 shows the density perturbations of `B17SG` discs when $M_p = 0.3$ and 1.0 M_{th} (top and bottom) at $t = 100\,t_p$. $Q = 2$ discs have become unstable at this point. This is why low Q and high M_p curves are missing in Fig. 8. The disc would become more stable if the radiative cooling is included. For instance, the `AS209ADSG` simulations discussed in Section 5.3 can evolve for very long time without any sign of instability even though their $Q = 5$ and $M_p = 1\,M_{th}$. Another interesting behaviour in the lower panels

is that as the disc becomes more massive, the number of vortices at the gap edge becomes larger. This is consistent with the simulation results in Lin (2012), Zhu & Baruteau (2016), as self-gravity can suppress large-scale (small m) vortices produced by Rossby wave instability (Lovelace et al. 1999).

6 CONCLUSIONS

We run 2D hydrodynamical simulations to explore the effects of self-gravity and radiative cooling on the gap and spiral formation in protoplanetary discs. We explore these effects with different Toomre Q parameters and the dimensionless cooling times $T_{\rm cool}$ for low mass planets. As the disc becomes more massive (smaller Q), we find:

(i) The spirals become slightly more tightly wound, especially when $Q \sim 2$, which is consistent with the linear theory.

(ii) The spirals become stronger and the primary and secondary gaps become deeper. This is due to the higher AMF in more massive discs, which is consistent with the linear theory. The advective AMF in the $Q = 2$ disc is almost five times the AMF in the $Q = 100$ disc.

(iii) If $Q \gtrsim 2$, the secondary gap's position does not change, despite gaps being deeper.

In discs with radiative cooling, we find:

(i) Even with a relatively short $T_{\rm cool} = 0.01$, the spiral waves cannot pick up AMF from the background disc when they are propagating to hotter regions, different from locally isothermal discs.

(ii) The spirals become less tightly wound (more open) as the cooling time increases, due to the increase of the disc's effective sound speed. The spiral's openness changes dramatically from discs with $T_{\rm cool} = 0.1$ to $T_{\rm cool} = 1$.

(iii) The spirals become weaker as the cooling time increases from a small value to $1/\Omega$, and then start to become stronger as the cooling time increases. There is a significant change of the spiral's amplitude when $T_{\rm cool}$ increases from $\lesssim 0.01$ to 0.1.

(iv) The AMF of the wave dissipates the fastest during its propagation when $T_{\rm cool} \sim 0.1$ to 1, implying that radiative damping may be important.

(v) With weaker spirals, the secondary gaps due to low mass planets disappear when $T_{\rm cool} = 0.1$ to 1.

(vi) If the secondary gap is present, its position moves from $\sim 0.5\ r_{\rm p}$ to $\sim 0.4\ r_{\rm p}$ in a disc with $(h/r)_{\rm p} = 0.07$ in transition from the isothermal limit to the adiabatic limit.

(vii) For $Q \gtrsim 5$ discs, the effect of radiative cooling is more critical than self-gravity.

Our simulations also have implications for the observations. We have run planet–disk simulations with massive planets ($\sim$ thermal mass):

(i) One might overestimate the planet mass using the simulation that neglects self-gravity. The gap width is less sensitive to the disc self-gravity than the gap depth. Thus, it is better to constrain the planet mass using the gap width when disc self-gravity is important.

(ii) One might underestimate the planet mass using the simulation that neglects the radiative cooling. This is especially the case when $T_{\rm cool} \sim 1$.

(iii) For deep primary gaps in the high planet mass regime, their shapes are less affected by the cooling time, but the secondary gap's position and depth for the gas depend on the cooling time. On the other hand, the secondary gap in dust or brightness temperature is less sensitive to the cooling time since dust drift fast to the primary gap edge forming secondary gaps. The tertiary gap becomes unnoticeable in both gas and dust when $T_{\rm cool} = 1$.

(iv) The dependence of the gap properties (e.g. gap depth, width, and secondary/tertiary gap's position and depth) on the cooling time-scale provides a new way to constrain the gaseous disc surface density.

(v) Assuming the gaps at 24, 35, and 61 au in AS 209 are all associated with the planet at 99 au, the gas surface density at $\sim$100 au in AS 209 should be $\lesssim 20\ {\rm g\,cm^{-2}}$ using the $T_{\rm cool} \lesssim 0.1$ cooling constraints, which is consistent with the surface density constraint from $Q \gtrsim 1$.

ACKNOWLEDGEMENTS

SZ and ZZ thank Chao-Chin Yang, Steve Lubow, and Jeffery Fung for helpful discussions. ZZ acknowledges support from the National Aeronautics and Space Administration through the Astrophysics Theory Program with Grant No. NNX17AK40G, the National Science Foundation under CAREER grant No. AST-1753168, and Sloan Research Fellowship. Simulations are carried out with the support from the Texas Advanced Computing Center (TACC) at The University of Texas at Austin through XSEDE grant TG-AST130002, and the NASA High-End Computing Program through the NASA Advanced Supercomputing Division at Ames Research Center.

The following softwares were used in this study: `Dusty FARGO-ADSG` (Baruteau & Masset 2008a,b; Baruteau & Zhu 2016), `Athena++` (Stone et al., in preparation), `Matplotlib` (Hunter 2007), `Numpy` (Van Der Walt, Colbert & Varoquaux 2011), `Scipy` (Jones et al. 01).

REFERENCES

Andrews S. M. et al., 2018, ApJ, 869, L41
Artymowicz P., 1993a, ApJ, 419, 155
Artymowicz P., 1993b, ApJ, 419, 166
Bae J., Zhu Z., 2018a, ApJ, 859, 118
Bae J., Zhu Z., 2018b, ApJ, 859, 119
Bae J., Zhu Z., Hartmann L., 2017, ApJ, 850, 201
Baruteau C., Masset F., 2008a, ApJ, 672, 1054
Baruteau C., Masset F., 2008b, ApJ, 678, 483
Baruteau C., Zhu Z., 2016, MNRAS, 458, 3927
Béthune W., Rafikov R. R., 2019, MNRAS, 487, 2319
Binney J., Tremaine S., 2008, Galactic Dynamics, 2nd edn. Princeton Univ. Press, Princeton
Birnstiel T. et al., 2018, ApJ, 869, L45
Booth A. S., Walsh C., Ilee J. D., Notsu S., Qi C., Nomura H., Akiyama E., 2019, ApJ, 882, L31
de Juan Ovelar M., Min M., Dominik C., Thalmann C., Pinilla P., Benisty M., Birnstiel T., 2013, A&A, 560, A111
Dong R., Rafikov R. R., Stone J. M., 2011a, ApJ, 741, 57
Dong R., Rafikov R. R., Stone J. M., Petrovich C., 2011b, ApJ, 741, 56
Dong R., Zhu Z., Whitney B., 2015, ApJ, 809, 93
Dong R., Li S., Chiang E., Li H., 2018, ApJ, 866, 110
Dullemond C. P. et al., 2018, ApJ, 869, L46
Flock M., Ruge J. P., Dzyurkevich N., Henning T., Klahr H., Wolf S., 2015, A&A, 574, A68
Gammie C. F., 2001, ApJ, 553, 174
Goldreich P., Tremaine S., 1978, ApJ, 222, 850
Goldreich P., Tremaine S., 1979, ApJ, 233, 857
Goldreich P., Tremaine S., 1980, ApJ, 241, 425
Gonzalez J.-F., Laibe G., Maddison S. T., 2017, MNRAS, 467, 1984
Goodman J., Rafikov R. R., 2001, ApJ, 552, 793
Guzmán V. V. et al., 2018, ApJ, 869, L48

Gyeol Yun H., Kim W.-T., Bae J., Han C., 2019, ApJ, 884, 142
Huang J. et al., 2018, ApJ, 869, L42
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jones E., et al., 2001, SciPy: Open source scientific tools for Python. Available at: http://www.scipy.org/
Juhász A., Rosotti G. P., 2018, MNRAS, 474, L32
Kanagawa K. D., Muto T., Tanaka H., Tanigawa T., Takeuchi T., Tsukagoshi T., Momose M., 2015, ApJ, 806, L15
Kanagawa K. D., Muto T., Tanaka H., Tanigawa T., Takeuchi T., Tsukagoshi T., Momose M., 2016, PASJ, 68, 43
Kley W., Nelson R. P., 2012, ARA&A, 50, 211
Li S., Li H., 2016, J. Phys., 719, 012007
Lin M.-K., 2012, MNRAS, 426, 3211
Lin M.-K., 2015, MNRAS, 448, 3806
Lin M.-K., Papaloizou J. C. B., 2011, MNRAS, 415, 1445
Long F. et al., 2018, ApJ, 869, 17
Lovelace R. V. E., Jore K. P., Haynes M. P., 1997, ApJ, 475, 83
Lovelace R. V. E., Li H., Colgate S. A., Nelson A. F., 1999, ApJ, 513, 805
Lynden-Bell D., Kalnajs A. J., 1972, MNRAS, 157, 1
Miranda R., Rafikov R. R., 2019a, ApJ, 875, 37
Miranda R., Rafikov R. R., 2019b, ApJ, 878, L9
Ogilvie G. I., Lubow S. H., 2002, MNRAS, 330, 950
Okuzumi S., Momose M., Sirono S.-i., Kobayashi H., Tanaka H., 2016, ApJ, 821, 82
Pérez L. M. et al., 2016, Science, 353, 1519
Pierens A., Huré J.-M., 2005, A&A, 433, L37
Pinilla P., Pohl A., Stammler S. M., Birnstiel T., 2017, ApJ, 845, 68
Pohl A., Pinilla P., Benisty M., Ataiee S., Juhász A., Dullemond C. P., Van Boekel R., Henning T., 2015, MNRAS, 453, 1768
Powell D., Murray-Clay R., Pérez L. M., Schlichting H. E., Rosenthal M., 2019, ApJ, 878, 116
Rafikov R. R., 2002, ApJ, 572, 566
Rafikov R. R., Petrovich C., 2012, ApJ, 747, 24
Rosotti G. P. et al., 2019, MNRAS, 2689
Takahashi S. Z., Inutsuka S.-i., 2014, ApJ, 794, 55
Teague R., Bae J., Bergin E. A., Birnstiel T., Foreman-Mackey D., 2018, ApJ, 860, L12
van der Marel N., Dong R., di Francesco J., Williams J. P., Tobin J., 2019, ApJ, 872, 112
van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
Ward W. R., 1997, Icarus, 126, 261
Yu S.-Y., Ho L. C., Zhu Z., 2019, ApJ, 877, 100
Zhang K., Blake G. A., Bergin E. A., 2015, ApJ, 806, L7
Zhang S. et al., 2018, ApJ, 869, L47
Zhu Z., Baruteau C., 2016, MNRAS, 458, 3918
Zhu Z., Dong R., Stone J. M., Rafikov R. R., 2015, ApJ, 813, 88
Zhu Z. et al., 2019, ApJ, 877, L18

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