RAA 2014 Vol. **14** No. **4**, 433–455 doi: 10.1088/1674–4527/14/4/006
http://www.raa-journal.org *http://www.iop.org/journals/raa*

Research in Astronomy and Astrophysics

Gap formation in a self-gravitating disk and the associated migration of the embedded giant planet *

Hui Zhang[1], Hui-Gen Liu[1], Ji-Lin Zhou[1] and Robert A. Wittenmyer[2]

[1] School of Astronomy and Space Science & Key Laboratory of Modern Astronomy and Astrophysics in Ministry of Education, Nanjing University, Nanjing 210093, China; *huizhang@nju.edu.cn*

[2] Department of Astrophysics, School of Physics, University of New South Wales, UNSW Sydney NSW 2052, Australia

Received 2013 March 10; accepted 2013 November 13

Abstract We present the results of our recent study on the interactions between a giant planet and a self-gravitating gas disk. We investigate how the disk's self-gravity affects the gap formation process and the migration of the giant planet. Two series of 1-D and 2-D hydrodynamic simulations are performed. We select several surface densities and focus on the gravitationally stable region. To obtain more reliable gravity torques exerted on the planet, a refined treatment of the disk's gravity is adopted in the vicinity of the planet. Our results indicate that the net effect of the disk's self-gravity on the gap formation process depends on the surface density of the disk. We notice that there are two critical values, $\Sigma_{\rm I}$ and $\Sigma_{\rm II}$. When the surface density of the disk is lower than the first one, $\Sigma_0 < \Sigma_{\rm I}$, the effect of self-gravity suppresses the formation of a gap. When $\Sigma_0 > \Sigma_{\rm I}$, the self-gravity of the gas tends to benefit the gap formation process and enlarges the width/depth of the gap. According to our 1-D and 2-D simulations, we estimate the first critical surface density to be $\Sigma_{\rm I} \approx 0.8$ MMSN. This effect increases until the surface density reaches the second critical value $\Sigma_{\rm II}$. When $\Sigma_0 > \Sigma_{\rm II}$, the gravitational turbulence in the disk becomes dominant and the gap formation process is suppressed again. Our 2-D simulations show that this critical surface density is around 3.5 MMSN. We also study the associated orbital evolution of a giant planet. Under the effect of the disk's self-gravity, the migration rate of the giant planet increases when the disk is dominated by gravitational turbulence. We show that the migration timescale correlates with the effective viscosity and can be up to 10^4 yr.

Key words: planets and satellites: formation — planetary systems: formation — planetary systems: protoplanetary disks

1 INTRODUCTION

To date, more than 900 exoplanets have been confirmed. The great diversity in the orbital characteristics of exoplanets reveals complicated physical and dynamical processes in their formation and evolution. One of the most important dynamical processes is the interaction between exoplanets and

* Supported by the National Natural Science Foundation of China.

the protostellar disk in which they are embedded. The physical properties of the protostellar disk usually dominate the initial conditions of the subsequent orbital evolution of the exoplanet system. Thanks to the improvement of direct imaging methods, a number of protostellar or debris disks interacting with exoplanets have been resolved, e.g. Fomalhaut (Kalas et al. 2008) and HR 8799 (Marois et al. 2008). Their detailed structures, such as the gaps created by the embedded planets, may be revealed in the near future.

According to the general theory of disk-planet interaction, a planet embedded in a protostellar disk will generate density waves within it. For a planet with a few Earth masses ($M_{\oplus}$), the response of the disk is linear and the structure of the disk is almost unchanged (Goldreich & Tremaine 1979; Ward 1997). On the other hand, for a planet with a mass comparable to that of Jupiter, the response of the disk becomes nonlinear and it usually results in a density gap at the position of the planet's orbit. In this regime, the planet is locked and moves as part of the disk. This is called type II migration (Lin & Papaloizou 1986).

The gap formation process is a key issue in understanding type II migration. In an inviscid disk, the gravitational tidal force exerted on the gas by a giant planet tends to split the disk, while the local fluid pressure resists the creation of any low density region. So, the criterion for gap formation is that the planet's Roche radius exceeds the pressure scale height of the disk. In the case of a viscous disk, the dissipation driven by the viscosity of the gas also tends to replenish the gap. As a result, the gap formation condition usually depends on the planet-star mass ratio $M_{\rm p}/M_*$, the semi-major axis of the planet's orbit $a_{\rm p}$, the scale height of the disk H and the viscosity of the gas α (Lin & Papaloizou 1993)

$$\frac{M_{\rm p}}{M_*} \equiv 40\alpha\Big(\frac{H}{a_{\rm p}}\Big)\,. \tag{1}$$

However, determining the width of a gap is not straightforward. Basically, the width of a gap is determined by the length scale of wave propagation, and should be a decreasing function of the effective viscosity of the gas (Lin & Papaloizou 1993; Takeuchi et al. 1996). One may rigorously define the positions of gap boundaries as the places where the tidal torque of the planet balances the torque raised by the viscous stress, given that all the other effects have already achieved equilibrium, e.g. the gravity of the central star, the thermal pressure and the centrifugal force of the gas. Most of these factors turn out to be strongly coupled with the surface density profile. If the perturbing planet is small, the response of the disk is linearly analyzable. However, a Jupiter sized planet, as we consider here, cannot be treated as a small perturbation. The density waves it excites are shocks and the associated gap formation process is a highly nonlinear process. Thus, numerical simulation is still the most powerful tool to study this process.

Many simulations have been performed to investigate the gap formation process in laminar viscous disks (Takeuchi et al. 1996; Kley 1999; Lubow et al. 1999; D'Angelo et al. 2003) and in MHD turbulent disks (Winters et al. 2003; Papaloizou et al. 2004). There is an important issue that has been poorly investigated so far, which is the self-gravitating effect of the gas (disk). In a non-ionized disk, gravitational turbulence is the most important source of the effective viscosity. In a high density self-gravitating disk, the gravitational turbulence can be very strong and even a Jupiter sized planet may not be able to open a gap (Baruteau et al. 2011). Conversely, in a low density disk, the effect of self-gravity is usually neglected or just treated as an effective viscosity (Gammie 2001). So, it seems that as the surface density of the disk increases, the effect of self-gravity will only result in a higher effective viscosity and monotonically reduce the gap size. When the density is high enough, even a giant planet could not open a gap.

However, the self-gravity potential is in fact coupled with the equilibrium angular velocity of the gas. As the self-gravity potential varies with the surface density profile, the angular velocity required by the equilibrium varies as well. Thus, the gas needs to drift inward or outward to achieve a new equilibrium, especially at the boundaries of the gap where the self-gravity potential changes the most. As a result, the self-gravity of the gas changes the size of the gap and its net effect on

the gap formation process may not be straightforward. Thus, systematic numerical experiments are needed. In this paper, we focus on the gravitationally stable region of the disk's surface density and investigate how the effect of self-gravity really affects the gap formation process. We also focus on the subsequent migration of the embedded giant planet.

We perform both 1-D and 2-D simulations to investigate the disk-planet interactions with effect of the the disk's self-gravity included. Our results show that the self-gravity does not suppress the gap formation process monotonically. Instead, there are two critical surface densities $\Sigma_{\rm I}$ and $\Sigma_{\rm II}$. When $\Sigma_0 < \Sigma_{\rm I}$, where Σ_0 is the initial surface density of disk, the gap formation process is suppressed when the effect of self-gravity is included. When $\Sigma_0 > \Sigma_{\rm I}$, the effect of self-gravity benefits the gap formation process and results in a wider gap. This enlargement enhances until the second critical surface density $\Sigma_{\rm II}$ is reached. When $\Sigma_0 > \Sigma_{\rm II}$, the gravitational turbulence viscosity becomes dominant in the disk and the gap formation process is suppressed again. The exact value of these two critical densities may depend on the many physical settings. In our simulations, we use MMSN (the surface density in the minimum mass solar nebula model from Hayashi 1981) as a unit of disk surface density. Then, the first critical density is around 0.8 and the second one is around 3.5. The associated migration of the giant planet is also studied and we find that the self-gravity of gas accelerates the type II migration when $\Sigma_0 > \Sigma_{\rm II}$. We confirmed that the migration timescale correlates with the effective viscosity in the disk, and can be as short as $\sim 10^4$ yr in a very dense disk with $\Sigma_0 \geq 7$ MMSN.

This paper is arranged as follows: we introduce the models of the 1-D and 2-D simulations in Section 2. The results are described in Section 3. In Section 4, we summarize our conclusions and discussions. The details of a refined treatment of the gravity torques and the calculation of the disk's self-gravity are described in Appendix A and B.

2 NUMERICAL MODEL

2.1 Computational Units

To normalize our calculations, we set the mass of the central star to be the mass unit $M_* = 1$ and the gravitational constant $G = 1$. The length unit is set to be the initial orbital radius of the planet $a_0 = 1$. Thus the orbital frequency of the planet is unity and its initial orbital period is $P_0 = 2\pi$. According to this configuration, our scale is in fact arbitrary. To connect with real physical dimensions, we further set the central star to be one solar mass $M_* = M_\odot$ and the initial orbital radius of the planet is $a_0 = 5.2$ AU. Thus, the time unit becomes $11.2\ {\rm yr}/2\pi$. According to the minimum-mass solar nebula model (MMSN, Hayashi 1981),

$$\Sigma = \Sigma_0 \Big(\frac{a}{1\ {\rm AU}}\Big)^{-3/2}\,, \tag{2}$$

where $\Sigma_0 = 1700\ {\rm g\ cm^{-2}}$. According to our length unit, where $a_0 = 5.2$ AU, the density constant in our model is $\Sigma_0 \approx 140\ {\rm g\ cm^{-2}}$. To be convenient, we set MMSN$= 140\ {\rm g\ cm^{-2}}$ as a surface density unit in this paper. Therefore, 2 MMSN equals $\Sigma_0 \approx 280\ {\rm g\ cm^{-2}}$ at $a_0 = 5.2$ AU.

2.2 1-D Model

In the first series of simulations, we solve the viscous evolution of a 1-D self-gravitating disk which is perturbed by a Jupiter mass planet. Assuming Σ is the surface density of the disk, and Ω and $v_{\rm r}$ are the angular and radial velocity of gas respectively, the 1-D continuity equation is

$$r\frac{\partial \Sigma}{\partial t} + \frac{\partial}{\partial r}(r\Sigma v_{\rm r}) = 0\,, \tag{3}$$

and the equation of angular momentum reads

$$r\frac{\partial(\Sigma r^2\Omega)}{\partial t}+\frac{\partial}{\partial r}(r\Sigma v_{\rm r}\cdot r^2\Omega)=\frac{1}{2\pi}\frac{\partial\tilde{G}}{\partial r}\,, \tag{4}$$

where $\tilde{G}$ is the rate of angular momentum transport. By eliminating $v_{\rm r}$, we obtain the governing equation

$$\frac{\partial\Sigma}{\partial t}=-\frac{1}{r}\frac{\partial}{\partial r}\left[\left(\frac{1}{2\pi}\frac{\partial\tilde{G}}{\partial r}-r^2\Sigma\frac{\partial v_\theta}{\partial t}\right)\Big/\left(\frac{\partial r v_\theta}{\partial r}\right)\right] \tag{5}$$

where $v_\theta=r\Omega$.

The rate of angular momentum transport $\tilde{G}$ is mainly determined by two factors: the effective viscosity of the gas $\tilde{G_\nu}$ and the torques exerted by the planet $\tilde{G_{\rm p}}$. For the first factor,

$$\tilde{G_\nu}=2\pi r\cdot\nu\Sigma r\frac{d\Omega}{dr}\cdot r\,, \tag{6}$$

where $\nu=\nu_{\rm art}+\nu_{\rm sg}=(\alpha_{\rm art}+\alpha_{\rm sg})(H/r)^2\Omega$. H/r is the scale height ratio of the disk. To identify the contribution from the disk's self-gravity, the total viscosity is divided into two parts: $\alpha_{\rm sg}$, the effective viscosity caused by the effect of self-gravity and $\alpha_{\rm art}$, an artificial viscosity denoting viscosity which comes from all other effects, e.g. the magnetorotational instability. The typical value of $\alpha_{\rm art}$ ranges from 10^{-3} to 10^{-2}. We are considering a self-gravitating disk, where most of the viscosity is assumed to be caused by the effect of self-gravity, so we adopt a low artificial viscosity: $\alpha_{\rm art}=10^{-3}$.

When a planet is embedded in the disk, its tidal torques cause transport of additional angular momentum (Lin & Papaloizou 1979). The total rate of angular momentum transport becomes

$$\tilde{G}=2\pi r\cdot\nu\Sigma r\frac{d\Omega}{dr}+\tilde{G_{\rm p}}\,, \tag{7}$$

where $\tilde{G_{\rm p}}$ is the torque exerted on the gas by the embedded planet, which contains both linear Lindblad and corotation torques associated with isothermal gas (Paardekooper et al. 2010). Following equations (14) and (15) of Ward (1997), we can obtain the smoothed Lindblad torque density. In the corotational region, the linear corotational torque density is represented in equation (16) of Paardekooper et al. (2010). Hence, we can finally obtain the torques $\Gamma_{\rm r}$ as well as the torque density $\frac{\partial\Gamma_{\rm r}}{\partial r}$ on the gas at radius r due to the planet.

The effects of self-gravity are simulated by two terms: the self-gravitating viscosity $\alpha_{\rm sg}$ and the self-gravity potential on the disk $\Phi_{\rm sg}$. Since the 1-D model could not simulate the gravitational turbulence well, we adopt an analytic description of $\alpha_{\rm sg}$. We will discuss it in detail in Section 3.2.1. Besides the effective viscosity, the self-gravity of the disk $\Phi_{\rm sg}$ can also change the meridional velocity field v_θ on the disk. Considering this to be a quasi-static process, we have

$$\frac{v_\theta^2}{r}=\frac{GM_*}{r^2}+\frac{1}{\Sigma}\frac{dP}{dr}+\frac{d\Phi_{\rm sg}}{dr}\,. \tag{8}$$

In this 1-D model, $\Phi_{\rm sg}$ is calculated by integrating the radial component of the gravity of all the grids on the disk. To avoid a singularity, a softening length $\epsilon=0.1H$ is adopted. We emphasize that the self-gravity potential $\Phi_{\rm sg}$ is not constant. Instead, it changes at every time step as the surface density changes, like the equilibrium angular velocity. This effect may imply there is a change in size of the gap.

The governing equation is diagonalized to a tridiagonal matrix. Methods to solve this kind of linear algebraic equation can be found in Press et al. (1992). The initial surface density Σ_0 is set equal to a series of values from 0.7 MMSN to 2.8 MMSN. The boundary condition is set to be solid where $\Sigma_{\rm bound}=\Sigma_0$.

2.3 2-D Model

In the second series of simulations, we solve the vertical integrated continuity and momentum equations in 2-D cylindrical coordinates by our ANTARES code. The details and convergence tests of our code can be found in Zhang et al. (2008) and Zhang & Zhou (2010a), respectively.

2.3.1 Numerical method

We assume the disk is thin and cold, where $H/r = 0.02$. The vertically averaged equations are solved in 2-D cylindrical coordinates (r, θ), whose origin is located at the central star. To make sure each cell is almost square, we adopt a logarithmic grid along the radial direction with a constant ratio $\beta = \Delta r/(r\Delta\theta) \approx 0.8$.

The major difficulty in numerical experiments that incorporate the effect of self-gravity is poor computational efficiency. It is too time-consuming to solve the Poisson equation describing gravitational potential on a highly perturbed disk, where the density varies quickly in both time and position. Thanks to the application of the Fast Fourier Transform (FFT) method, we could greatly reduce the complexity of this problem from N^2 to $N \ln N$, where N is the total number of grids used to resolve the disk. Despite this improvement, it is still too "expensive" to perform high resolution 2-D or 3-D simulations, when the total $N = N_{\rm r} \times N_\theta > 10^6$. On the other hand, low resolution simulations usually introduce un-physical effects and the results are thus less reliable.

One of the most significant numerical effects on an Eulerian grid is that, since the mass is placed at the center of each cell instead of smoothly spreading over it, the net gravity force exerted on the planet is usually dominated by the mass within only the single cell whose center is immediately adjacent to it. As the planet travels through a series of cells, the net torque exerted on it experiences un-physically large variations. When the resolution of the grid is high enough, this effect could be partly reduced by a well-chosen softening length. However, choosing the value of the softening length is difficult. On one hand, it should be small. It is usually smaller than the scale height of the disk or the Hill radius of the planet. On the other hand, it needs to be large enough that the softening region can be resolved by the grid size. It usually requires a large number of grids to resolve the immediate vicinity of the planet, e.g. the corotation zone of the planet (Masset & Ogilvie 2004). This also necessitates a high resolution for the grid. To balance the computational efficiency and accuracy, we adopt a relatively low mesh resolution $N_{\rm r} \times N_\theta = 256 \times 512$ and a refined treatment of the gravity torque in the vicinity of the planet (see Appendix A).

The velocity is denoted by $v = (v_{\rm r}, v_\theta)$, where $v_{\rm r}$ is the radial velocity and v_θ is the velocity in the azimuthal direction. The vertically averaged continuity equation is given by

$$\frac{\partial \sigma}{\partial t} + \frac{1}{r}\frac{\partial (r\sigma v_{\rm r})}{\partial r} + \frac{1}{r}\frac{\partial (\sigma v_\theta)}{\partial \theta} = 0\,. \tag{9}$$

The momentum equations in the radial and azimuthal directions are

$$\frac{\partial (\sigma v_{\rm r})}{\partial t} + \frac{1}{r}\frac{\partial (r\sigma v_{\rm r}^2)}{\partial r} + \frac{1}{r}\frac{\partial (\sigma v_{\rm r} v_\theta)}{\partial \theta} = \sigma\frac{v_\theta^2}{r} - \sigma\frac{\partial \Phi}{\partial r} - \frac{\partial P}{\partial r}\,, \tag{10}$$

$$\frac{\partial (\sigma v_\theta)}{\partial t} + \frac{1}{r}\frac{\partial (r\sigma v_{\rm r} v_\theta)}{\partial r} + \frac{1}{r}\frac{\partial (\sigma v_\theta^2)}{\partial \theta} = -\sigma\frac{v_{\rm r} v_\theta}{r} - \sigma\frac{\partial \Phi}{\partial \theta} - \frac{1}{r}\frac{\partial P}{\partial \theta}\,. \tag{11}$$

The external potential Φ is

$$\Phi = \Phi_{\rm S} + \Phi_{\rm p} + \Phi_{\rm D} + \Phi_{N,p} + \Phi_{N,D}\,, \tag{12}$$

where $\Phi_{\rm S} = -GM_\odot/|\boldsymbol{r}|$ is the potential of the central star, $\Phi_{\rm p} = -GM_{\rm p}/(|\boldsymbol{r} - \boldsymbol{r}_{\rm p}| + \varepsilon)$ is the potential of the planet and $\Phi_{\rm D}$ is the potential of self-gravity from the gaseous disk. This is determined by the Poisson equation

$$\nabla^2 \Phi_{\rm D} = 4\pi G\Sigma\,. \tag{13}$$

In our simulation, we directly calculate the self-gravity force $F_{\rm sg}$ by the FFT method (see Appendix B). $\Phi_{N,p}$ is the indirect potential caused by the Jupiter-mass planet

$$\Phi_{N,p} = \frac{M_{\rm p}}{M_\odot + M_{\rm p}} {\Omega_{\rm p}}^2 \boldsymbol{r} \cdot \boldsymbol{r}_{\rm p} \approx \frac{GM_{\rm p}}{{r_{\rm p}}^3} \boldsymbol{r} \cdot \boldsymbol{r}_{\rm p} \,. \tag{14}$$

$\Phi_{N,D}$ is the indirect potential due to the gravity of the gas disk

$$\Phi_{N,D} = G \int_{\rm D} \frac{\boldsymbol{r} \cdot \boldsymbol{r}'}{|\boldsymbol{r}'|^3} dm(\boldsymbol{r}') \,. \tag{15}$$

Since we have assumed the disk is very cold and we focus on the gravitationally stable region, we do not adopt the energy equation in the 2-D model. Instead, we adopt a locally isothermal equation of state

$$p = \Sigma c_{\rm s}^2 \,, \tag{16}$$

where $c_{\rm s}$ is the sound speed which is only a function of r; $c_{\rm s} = (H/r) v_{\rm kep}$ and $v_{\rm kep} = \sqrt{GM_*/r}$ is the local Keplerian velocity. We do not employ any artificial viscosity. However, the numerical viscosity due to the coarse grid is $\nu_{\rm num} \sim 10^{-5}$.

To estimate the numerical viscosity we have performed several short-term simulations to test the diffusion time of a density ring in the disk under different resolutions: 256×512, 512×1024, 1024×2048, 1600×3200 and 2048×4096. The self-gravities of the planet and disk are not included. We find that the diffusion time does not change any further when we change the resolution from 1600×3200 to 2048×4096. We believe that the grid effect is negligible when the resolution reaches 2048×4096. Then we add an artificial viscosity in the 2048×4096 case. When this artificial viscosity increases to 10^{-5}, we found the diffusion time is comparable to the value of the 256×512 case. So, we conclude that the viscosity from the coarse grid is about 10^{-5} in a resolution of 256×512.

2.3.2 Initial and boundary conditions

We fix the star at the origin of the frame and let the gas and the planet travel around it. The initial orbit of the planet is circular and its semi-major axis is set to be unity, $a_0 = 1$. To ensure the gas disk starts with an equilibrium state, the initial azimuthal velocity field is set to be $v(r)_{\theta 0} = [1/r + rF_{\rm sg}(r) - c_{\rm s}(r)^2]^{1/2}$, where $F_{\rm sg}(r)$ is the self-gravity of the gaseous disk and $c_{\rm s}(r)$ is the local sound speed. The initial radial velocity of gas v_{r0} is set to be 0.

To reduce the initial impact on the disk, we hold the planet in a circular orbit for 50 orbits and gradually increase its mass from 0.01 to 1 Jupiter mass. Since the initial mass of the planet is very small and the initial velocity of gas has taken the forces of gravity and pressure into account, the disk achieves a steady state well before the planet emerges. Two strong spiral arms emerge after about 30 orbits. When we release the planet, a clear gap has already formed. At the initial state, Toomre Q parameter is greater than 1 over the disk (Fig. 1).

The calculations are actually performed in a wide annulus, with the inner boundary located at $R_{\rm in} = 0.4a_0 = 2.08$ AU and the outer one located at $R_{\rm out} = 2.5a_0 = 13$ AU. We adopt outgoing boundary conditions at both the inner and outer boundaries. It is a wave absorbing boundary condition that the waves are only allowed to propagate out of the computational domain, while the waves traveling inward are set to zero. There are two ghost rings outside the boundaries, whose density and velocity field stay at the initial state. In the self-gravitating model, we include the gravitational potential of these two ghost rings to avoid the un-physical cutting-off of the self-gravity potential at the edges of the disk.

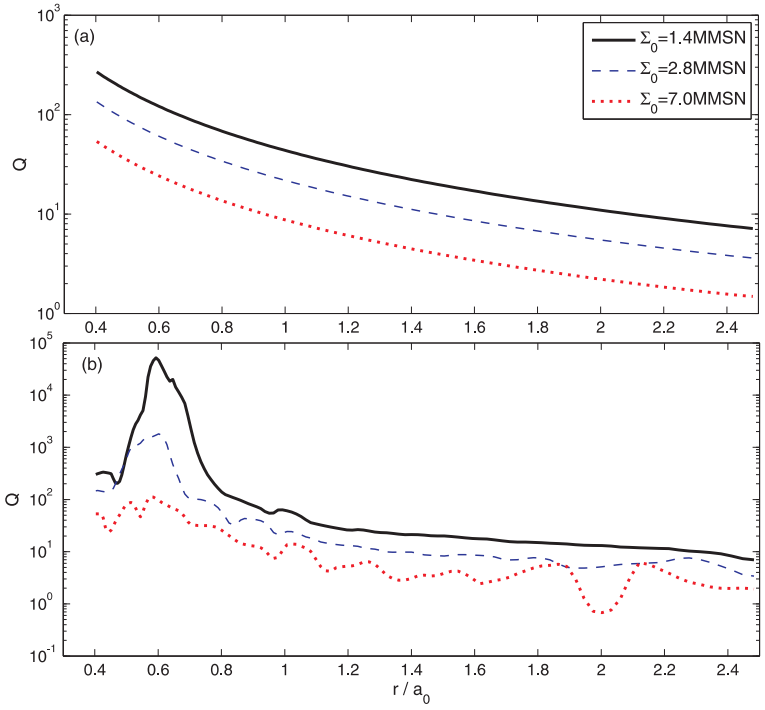

Fig. 1 Q profile on the disk. Panel (a): The initial Q profiles. Panel (b): The final Q profiles on the disk with different surface densities.

2.3.3 Measurement of the gap width

The gap width is a key quantity in this work, although the exact positions of gap boundaries are hard to be determined analytically. Fortunately, we are focusing on the relative changes of gap width in a disk with or without effects of self-gravity. So, we could define the gap width by the disk's surface density profiles. To ensure the comparability, we set the surface density at the initial position of the planet as the reference density. Then, the measurement of the gap width in the 1-D simulation is quite simple. At each side of the planet's orbit, we can find a position where the surface density is equal to the reference density. If we get more than one, the nearest one (to the planet) is chosen. Then we get two positions on both sides of the planet. We define these two radii as the inner and outer boundary of the gap and the width of the gap is the difference in their radial positions. The measurement in the 2-D simulation is similar. The only difference is that we use an azimuthally averaged density profile in the 2-D simulations (panel (b) in Fig. 2).

3 RESULTS

Our numerical simulations consist of two steps. First, we adopt a 1-D model that describes the radial viscous evolution of a self-gravitating disk. The effect of self-gravity in the gas is added both as an additional radial force field and an effective viscosity. Since the 1-D model is not suited to simulate the 2-D gravitational turbulence and the behavior of a gravitationally unstable disk, we concentrate on a range that has low surface density to study variations in the gap in a gravitationally stable disk. Secondly, to reveal variations in the gap within the transition stage (from gravitationally stable to

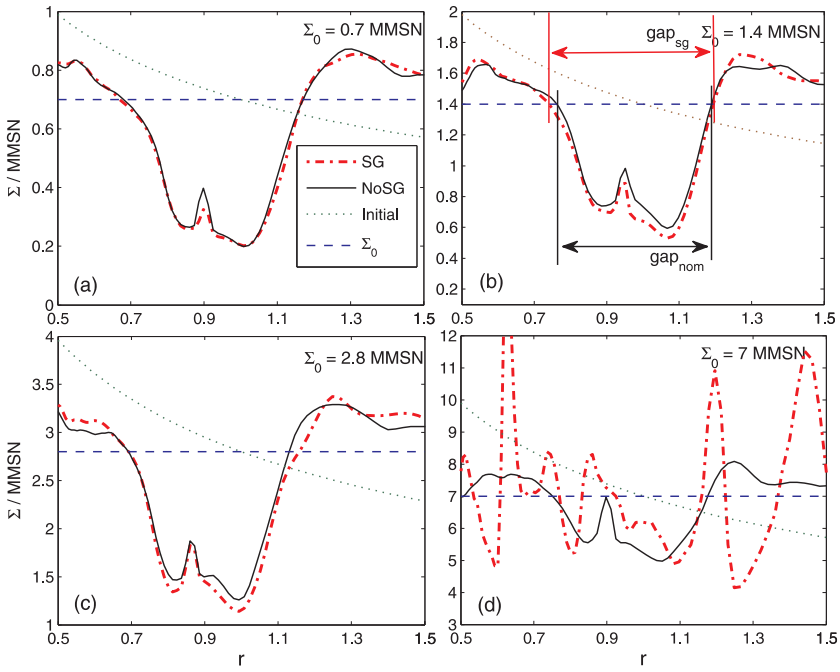

Fig. 2 Cross sections of surface density for the disks with different surface densities. The surface density is averaged over the azimuthal direction. These figures show the gap structures when the disk's self-gravity is included or excluded. When the disk's self-gravity is included, the gap is slightly deeper and wider. Panel (b) shows how we measure the width of the gap. Panel (d): the disk's structure becomes very turbulent in a dense self-gravitating disk. There is no clear gap in that case.

unstable), we further perform a series of fully self-consistent 2-D simulations with the self-gravity of gas included. We then investigate the orbital evolution of the embedded planet associated with the process of gap formation.

3.1 1-D Simulation

Panel (a) of Figure 3 shows the variation in the width of the gap versus evolution time in the self-gravitating and non-self-gravitating models. Our 1-D simulations show that, in a disk without the effect of self-gravity, the width of the gap is almost unchanged when the surface density (or disk mass) increases. This is consistent with the former analysis that when the self-gravity is absent, the width of the gap is determined by the dissipation of the gas and the tidal force of the planet (Goldreich & Tremaine 1980; Lin & Papaloizou 1986). When the self-gravity is included, we find that the width of the gap increases as the surface density increases. When the width of the gap becomes stable, we measure the difference in widths between the two gaps in different models for a series of surface densities (panel (b) of Fig. 3). It clearly shows that there exists a critical surface density around $\Sigma_{\rm I} \simeq 0.85$ MMSN. The self-gravity suppresses the process of gap formation when $\Sigma_0 < \Sigma_{\rm I}$ and enlarges the gap when $\Sigma_0 > \Sigma_{\rm I}$.

During the process of gap formation, the effect of self-gravity plays two opposing roles. On one hand, it drives an effective viscosity (Gammie 2001) which tends to make the disk more dissipative. Therefore, it is more difficult for the gap to be cleared and the process of gap formation is suppressed. On the other hand, the equilibrium at the position of the boundaries of the gap changes as the local

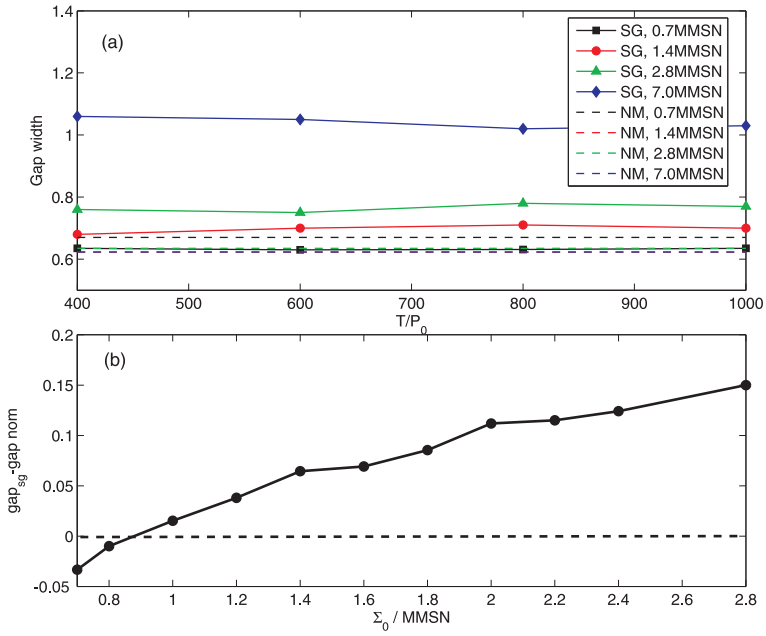

Fig. 3 Widths of the gap in 1-D simulations. Panel (a): Width of gap versus evolution time. From top to bottom the surface density of the disk decreases from 7 MMSN to 0.7 MMSN. The dashed lines denote the non-self-gravitating cases. Since the gap size almost does not change with the disk's density in the non-self-gravitating cases, the dashed lines that denote 1.4 MMSN and 2.8 MMSN are overlapped. Panel (b): The differences in width of the gap between the self-gravitating disk and the non-self-gravitating disk. The surface density increases from 0.7 MMSN to 2.8 MMSN, and the critical surface density is around $\Sigma_{\rm I} \approx 0.85$ MMSN.

self-gravitational potential varies with the surface density there. When the slope of the density becomes sharp at the boundaries of the gap, the local self-gravity potential may change direction and it tends to contract the disk. This effect may lead to enlargement of the gap. The behavior of the width of the gap under these two effects is described below.

When the disk surface density is low, the dynamics of the gas are mostly determined by the central gravity GM_*/r^2. Although the global self-gravity potential of the disk is weak, the gas exchanges angular momentum more effectively with immediate neighbors by the local mutual gravity. This can be expressed as an effective viscosity which suppresses the process of gap formation. As the density of the disk increases, the global self-gravitational potential begins to make measurable influences on the central gravity. For example, we could just look at the outer boundary of the gap, where $r = r_{\rm ob}$. When the gap is stable, there is an equilibrium

$$\frac{v_\theta^2}{r_{\rm ob}} = \frac{GM_*}{r_{\rm ob}^2} + \frac{1}{\Sigma}\frac{dP}{dr}|_{\rm ob}\,, \tag{17}$$

given that the tidal force of the planet is balanced by the viscosity dissipation. When the self-gravity $\Phi_{\rm sg}(r,t)|_{\rm ob}$ is included, the equilibrium becomes

$$\frac{v_\theta^2}{r_{\rm ob}} = \frac{GM_*}{r_{\rm ob}^2} + \frac{1}{\Sigma}\frac{dP}{dr}|_{\rm ob} + \frac{d\Phi_{\rm sg}}{dr}|_{\rm ob}\,. \tag{18}$$

As the gas is being cleared in the gap, the gradient of self-gravitational potential becomes very sharp at the boundaries. Thus, $F_{\rm sg}(r_{\rm ob},t) = -d\Phi_{\rm sg}/dr|_{\rm ob}$ increases from negative (directed inward)

to positive (directed outward). By assuming that the tidal force of the planet and the viscous dissipation remain balanced, we may find that when $F_{\rm sg}(r_{\rm ob}, t)$ increases, the angular velocity required by the equilibrium decreases. During this transition stage, the angular velocity of the gas at $r_{\rm ob}$ is greater than that required by the equilibrium, so the gas tends to drift outward. Meanwhile, the pressure gradient and viscous dissipation try to push the gas back. However, the disk has not been dominated by the gravitational turbulence yet—the effective viscosity is still too low: $\alpha_{\rm sg} \sim 10^{-3}$. The viscous timescale is as long as 10^6 yr and is much longer than the variation timescale of $F_{\rm sg}(r_{\rm ob}, t)$ which is only dozens of orbits for a Jupiter-mass planet. To retain the equilibrium, the surface density profile needs to become sharper to generate a stronger pressure gradient, $(1/\Sigma)dP/dr|_{\rm ob}$, at the gap boundaries. However, the sharper gradient of the surface density also enhances the gradient of the self-gravity potential at the boundaries. Finally, the outer boundary moves outward until the angular velocity of the gas matches the required value and a new equilibrium is achieved. A similar process occurs at the inner boundary of the gap but results in an inward drift of the gas. This combined effect behaves like a 'self-gravitational contraction' of the two parts of the disk and makes the gap become wider and deeper. Furthermore, since the pressure effect decreases as Σ increases, this effect is more pronounced as the disk becomes denser (Fig. 3).

3.2 2-D Simulation

Our 1-D simulations suggest that when the surface density exceeds $\Sigma_{\rm I}$, the width of the gap increases monotonically (for Σ up to 2.8 MMSN). To ensure that this is a trend in a fully described self-gravitating disk, a series of 2-D hydrodynamic simulations is performed. The orbital evolution of the giant planet embedded in a self-gravitating disk is also studied. Since the 2-D simulation is very time consuming when the disk self-gravity is included, we only choose four typical surface densities: 0.7, 1.4, 2.8 and 7 MMSN.

3.2.1 Gap formation

Panels (a) and (b) in Figure 4 show the evolution of width of the gap versus time in the self-gravitating and non-self-gravitating disks. Different surface densities are denoted by corresponding marks. Note that the decrease in the width of the gap after about $200P_0$ is due to the decrease in the Hill radii when the planet is migrating inward ($a_{\rm p}$ decreases). It is clear that the surface density does not change the width of the gap when the self-gravity is excluded, but the width of the gap in a self-gravitating disk strongly depends on the surface density (see panel (c) of Fig. 4). Panel (a) of Figure 5 shows the evolution of the normalized gap differences: $({\rm gap}_{\rm sg} - {\rm gap}_{\rm nom})/{\rm gap}_{\rm nom}$. All the widths of the gap have been normalized by the corresponding semi-major axis of the planet to eliminate the effect of migration. In a self-gravitating disk, the width of the gap increases as the disk's surface density increases. However, it is not a linear relation. Furthermore, the 2-D simulations show that the enlargement of the gap decreases when the surface density exceeds ~ 2 MMSN and becomes negative when $\Sigma_0 > 3.5$ MMSN (panel (b) of Fig. 5). The width of the gap is recorded every 10 orbits. When the simulation is complete, we sum all the widths together and find the averaged value. Note that widths of the gap from the first 100 orbits are dropped, since the gap is not well formed before that.

Figure 2 shows the gap structures under different situations and how we measure the width of the gap. The size of the gap is almost identical when the surface density is low. When the disk becomes denser, the gap is slightly deeper and wider in the self-gravitating disks. We measure the differences of the averaged width of the gap between the self-gravitating and non-self-gravitating models and interpolate these data (panel (b) of Fig. 5). The results suggest that there is another critical surface density which is around $\Sigma_{\rm II} \simeq 3.5$ MMSN. When $\Sigma_0 \geq \Sigma_{\rm II}$, the self-gravity suppresses the formation of the gap again. Notice that, for a very dense disk $\Sigma_0 \geq 7$ MMSN, the gap is not cleared.

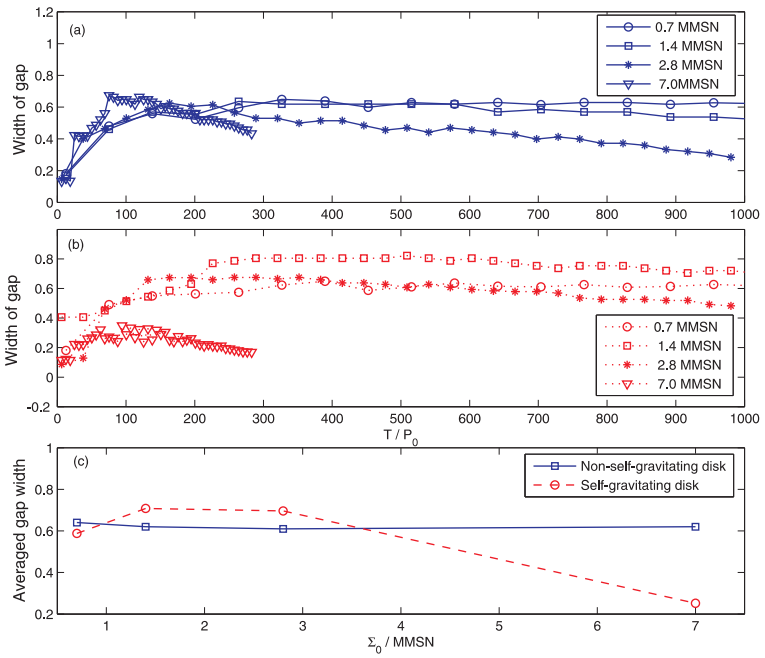

Fig. 4 Evolution of width of the gap in 2-D simulations. Panel (a): width of the gap versus evolution time in non-self-gravitating disks. The gap is cleared around 100 orbits and reaches the maximum value. Then the width of the gap decreases as the planet migrates inward. All the widths are azimuthally averaged. Panel (b): width of the gap versus evolution time in self-gravitating disks. Panel (c): differences in width of the of gap in self-gravitating and non-self-gravitating cases. The width of the gap does not change with the surface density of the non-self-gravitating disks.

So, it is significantly smaller than the others (Panel (d) of Fig. 2 and Fig. 6). The widths shown in Figure 4 are the azimuthally averaged value.

When the surface density exceeds 2.8 MMSN, the gravitational turbulence becomes significant. Figure 6 shows the density contours of the disk under different surface densities. The three figures in the left column are the normal disks. Their disk structures do not change much when their surface density increases from 0.7 MMSN to 7 MMSN. The three figures in the right column are the self-gravitating disks. When the surface density increases to 2.8 MMSN, turbulence emerges at the outer part of the disk, where the Toomre Q parameter is relatively low. As the surface density increases further, the gravitational turbulence becomes stronger. When the disk's surface density exceeds $\Sigma_{\rm II}$, the disk becomes gravitationally unstable (the bottom figure in the left column). At such a high surface density, the effective viscosity caused by the self-gravitational turbulence will overcome the effect of 'self-gravitational contraction' and dominate the gap formation process.

Compared with our 1-D simulations, there are two major differences. One is that our 2-D simulations indicate a smaller value of the first critical surface density $\Sigma_{\rm I} \sim 0.8$. This suggests that the 'self-gravitational contraction' is stronger in a 2-D disk. This is probably because, in the 1-D simulations, we adopt an artificial viscosity $\nu_{\rm art}$, which turns out to be slightly larger than the numerical viscosity $\nu_{\rm num}$ in our 2-D simulations, which makes the total effective viscosity in the 1-D simulation slightly larger than the one in the 2-D simulation. However, the difference is quite small (our 1-D results indicate $\Sigma_{\rm I} \sim 0.85$) and does not change our main results.

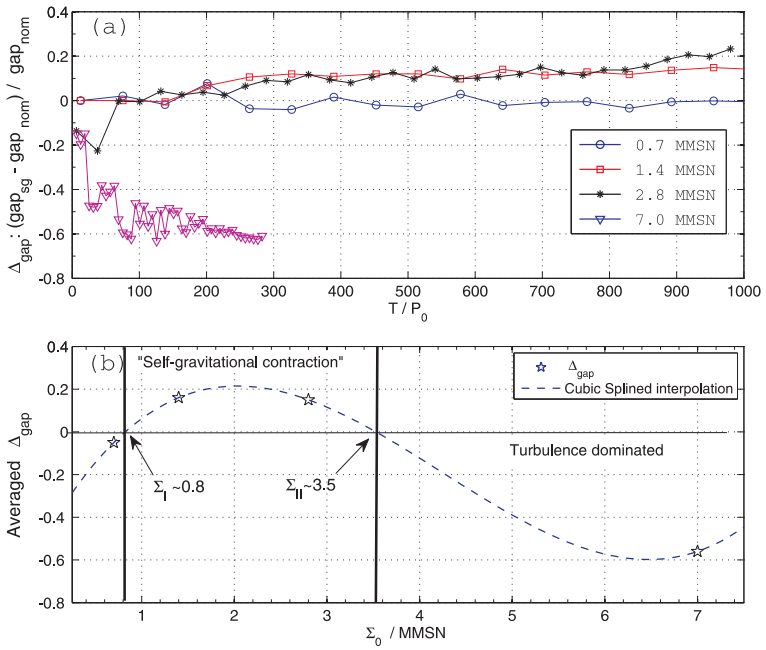

Fig. 5 The relative differences in width of the gap in 2-D simulations. Panel (a): The relative differences in width between self-gravitating and non-self-gravitating disks. Width of the gap has first been normalized by the corresponding semi-major axis of the planet to eliminate the effect of migration. Panel (b): The time averaged relative differences versus the surface density of the disk. The dashed lines show the cubic spline interpolations and we found the second critical surface density $\Sigma_{\rm II} \approx 3.5$ MMSN. The first one $\Sigma_{\rm I} \approx 0.8$ MMSN is in agreement with our 1-D results. When $\Sigma_{\rm II} > \Sigma_0 > \Sigma_{\rm I}$ the "self-gravitating contraction" dominates the process of gap formation. When $\Sigma_0 > \Sigma_{\rm II}$, the viscosity from gravitational turbulence becomes dominant. In a self-gravitating disk, width of the gap reaches a maximum when $\Sigma_0 \approx 2$ MMSN. All the widths we adopt are the azimuthally averaged value. Notice that, for a very dense disk $\Sigma_0 \geq 7$ MMSN, the gap is not cleared, so it is relatively small.

The other difference is that our 1-D results suggest that the gap size increases monotonically as the surface density increases from 0.7 MMSN to 2.8 MMSN. However our 2-D results show that the increasing trend decreases around 2 MMSN. For the 1-D simulations, the angular momentum exchange caused by the effect of self-gravity was only described by an effective viscosity $\nu_{\rm sg}$. In this description, $\nu_{\rm sg} = \alpha_{\rm sg} c_{\rm s}^2/\Omega$ (Gammie 2001), where

$$\alpha_{\rm sg} = \frac{4}{9\gamma(\gamma-1)t_{\rm cool}\Omega}\,. \tag{19}$$

The cooling timescale is determined by the internal energy per unit area U and the cooling function Λ,

$$t_{\rm cool} = \frac{U}{\Lambda} = \frac{c_{\rm s}^2\Sigma}{\gamma(\gamma-1)\Lambda}\,, \tag{20}$$

and (Hubeny 1990)

$$\Lambda = \frac{16\sigma}{3}(T_c^4 - T_0^4)\frac{\tau}{1+\tau^2}\,. \tag{21}$$

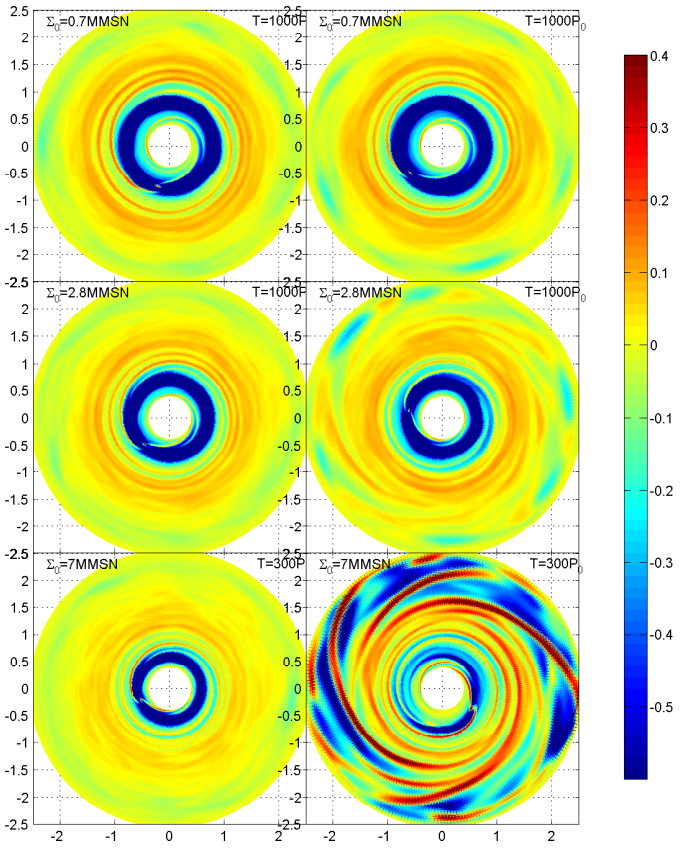

Fig. 6 The evolution of surface density. The left panels show the non-self-gravitating disk and the right panels show the self-gravitating disk. The surface density of the disk is $\Sigma_0 = 1.4$, 2.8 and 7 MMSN from top to bottom, respectively. The gravitational turbulence clearly emerges when $\Sigma_0 > 2.8$ MMSN.

$T_c = 280\,\mathrm{K}(a/1\mathrm{AU})^{-1/2}$ is the mid-plane temperature of the disk and $T_0 = 10\,\mathrm{K}$ is a minimum temperature of background sources (Stamatellos et al. 2007). Using the analytic approximation of the Rosseland mean opacity for molecules (Bell & Lin 1994),

$$\kappa = \kappa_0 \Big(\frac{\Sigma}{2H}\Big)^{2/3} T_c^3 \,, \tag{22}$$

we can get the optical depth τ (Rice et al. 2010),

$$\tau \approx H\kappa\Big(\frac{\Sigma}{2H}, T\Big)\frac{\Sigma}{2H} = H\Big(\frac{\Sigma}{2H}\Big)^{5/3} T_c^3 \,. \tag{23}$$

Then we get

$$\Lambda = \frac{16\sigma}{3}(T_c^4 - T_0^4)\frac{H\Big(\frac{\Sigma}{2H}\Big)^{5/3} T_c^3}{1 + \Big[H\big(\frac{\Sigma}{2H}\big)^{5/3} T_c^3\Big]^2} \,. \tag{24}$$

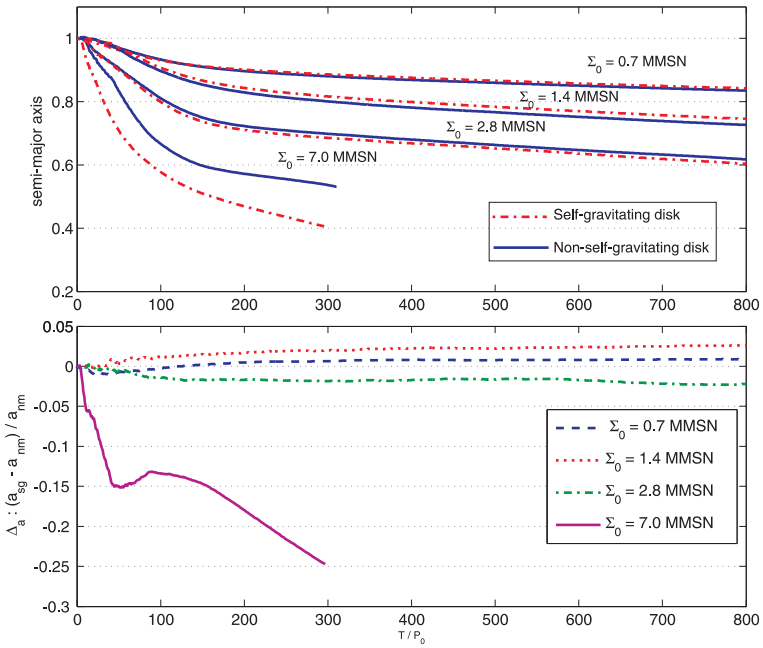

Fig. 7 *Upper panel*: orbital migrations of the planet with or without the disk's self-gravity. Blue solid lines show the migration in non-self-gravitating disks, while the red solid lines show the migration in self-gravitating disks. *Lower panel*: the differences of migrations between the self-gravitating cases and non-self-gravitating cases (normal cases). The value has been normalized by the value from corresponding normal cases. The difference in the migration is significant only when the disk is very dense.

At the location of the giant planet, where $a = 5.2$ AU, $H/r = 0.02$ and $\Sigma = 1$ MMSN, we found that

$$\left[H\left(\frac{\Sigma}{2H}\right)^{5/3} T_c^3\right]^2 \approx 10^3 \gg 1\,. \tag{25}$$

Thus, we have $\Lambda \propto \Sigma^{-5/3}$. This gives us $t_{\rm cool} \propto \Sigma^{8/3}$ and $\nu_{\rm sg} \propto \Sigma^{-8/3}$. So, as the surface density Σ increases, the dissipation in the disk becomes weaker and the gap forms more effectively. This result could be valid when Q is much larger than unity (Rice et al. 2010 estimated that $Q \geq 2$). In some high-density simulations, however, Q is close to unity after several hundred orbits (Fig. 1), so we believe that the real effect of self-gravity in a dense disk should be consistently calculated by the realtime density distribution, and the 2-D simulations should be more self-consistent.

3.2.2 Migration of the giant planet

Besides formation of the gap, the orbital migration of the planet is another important outcome of the disk-planet interactions. The upper panel of Figure 7 shows the migration of the planet embedded in a series of disks. The dashed lines are the results with the self-gravity of the gas included, while the solid lines are those results without the self-gravity of gas. From top to bottom, the surface density of the disk increases from 0.7 MMSN to 7 MMSN. One may find that all the migrations experience two stages. At the first stage, the giant planet is still surrounded by the gas and undergoes the type I (or type-I-like) migration whose timescale should be inversely proportional to the disk's surface

density (Tanaka et al. 2002),

$$\tau = (2.7 + 1.1\gamma)^{-1} \frac{M_*}{M_\mathrm{p}} \frac{M_*}{\Sigma_0 a_\mathrm{p}^2} \Big(\frac{c_\mathrm{s}}{a_\mathrm{p}\Omega_\mathrm{p}} \Big)^2 \Omega_\mathrm{p}^{-1} \propto \Sigma_\mathrm{p}^{-1} \,. \tag{26}$$

Our results show that at this stage, the migration rates of the planet are greater as the disk becomes denser (with differing slopes of the migration curve in the upper panel of Fig. 7 and panel (a) in Fig. 8). It is qualitatively consistent with the analytic predictions we mentioned above and this could demonstrate the consistency of our simulations. The lower panel of Figure 7 shows the relative differences in migration (semi-major axis vs. time) between the self-gravitating cases and the normal (non-self-gravitating) cases. The differences are normalized by the values from the corresponding normal cases.

As the gas located in the gap region is cleared, the migration of the giant planet steps into the second stage when the migration rate of the planet is significantly reduced. This is usually called type II migration. According to linear analysis, the timescale of type II migration is supposed to be inversely proportional to the effective viscosity in the disk. From Figure 7 we can find that the type II migrations in different surface densities have almost the same slope when the self-gravity of the disk is excluded. This is reasonable since the effective viscosity should not depend on the surface density. However, we find that the migration rate in the denser disk is indeed larger than the rate in the thinner disk (also in panel (b) of Fig. 8). The reason is that there is an inner boundary in our disk model. When the planet is getting close to the inner boundary, most of the inner disk has flow outside our inner boundary. As a result, the torque from the inner disk (positive torque) is weakened and the net negative torque is greater. That means the planet will drop to the central star faster when it gets closer to the inner boundary in our simulations. Meanwhile, a planet migrates faster in a denser disk than in a thinner disk before the gap is cleared. So, when the migration steps into the type-II regime, a planet embedded in a denser disk will be closer to the inner boundary of the disk and has a larger inward migration rate. However, in the self-gravitating disk, the type II migration rate changes compared to the disk's surface density, because the effective viscosity is now related to the disk's surface density. When the surface density is low, the difference in the semi-major evolution is very small: $< 2\%$ (lower panel of Fig. 7). When the surface density is higher ($\Sigma_0 \geq 7.0$ MMSN), the difference becomes very significant.

In this paper, we concentrate on the variations of the migration rate under the effect of a disk's self-gravity which is the source of the turbulent viscosity. Panel (a) of Figure 8 shows the evolution of the migration rate ($\dot{a_\mathrm{p}}$) with the effect of self-gravity included. After about 300 orbits, the migration rate reaches different stable values according to the surface density of the disk. We measure this stable migration rate in each run and the results are shown in panel (b) of Figure 8. The circles are the results with the self-gravity of gas included and the squares are the results without the self-gravity. The migration rate weakly increases with the surface density in a non-self-gravitating disk. This indicates that the effective viscosity barely changes with the surface density when the self-gravitating effect is excluded. However, in a self-gravitating disk, the migration rate increases quickly as the surface density increases. Our results suggest that, in a self-gravitating disk, the migration of a giant planet is slightly slowed (almost identical with the non-self-gravitating case) when the surface density is moderate. However, the migration of the giant planet becomes faster than that in the non-self-gravitating disk when the surface density exceeds 2.8 MMSN. In a very dense disk with $\Sigma_0 = 7$ MMSN, the migration of the giant planet could be very fast and the timescale could be as short as $\sim 10^4$ yr (Fig. 8).

The quick increase of the migration rate indicates that the effective viscosity is mostly determined by the gravitational turbulent viscosity and increases with the surface density of a self-gravitating disk. We sum the angular momentum of the whole disk and measure its rate of variation. Since the size of the disk does not change with time, variation in its angular momentum is only determined by variations in the radial mass flow and angular velocity, which are both the results

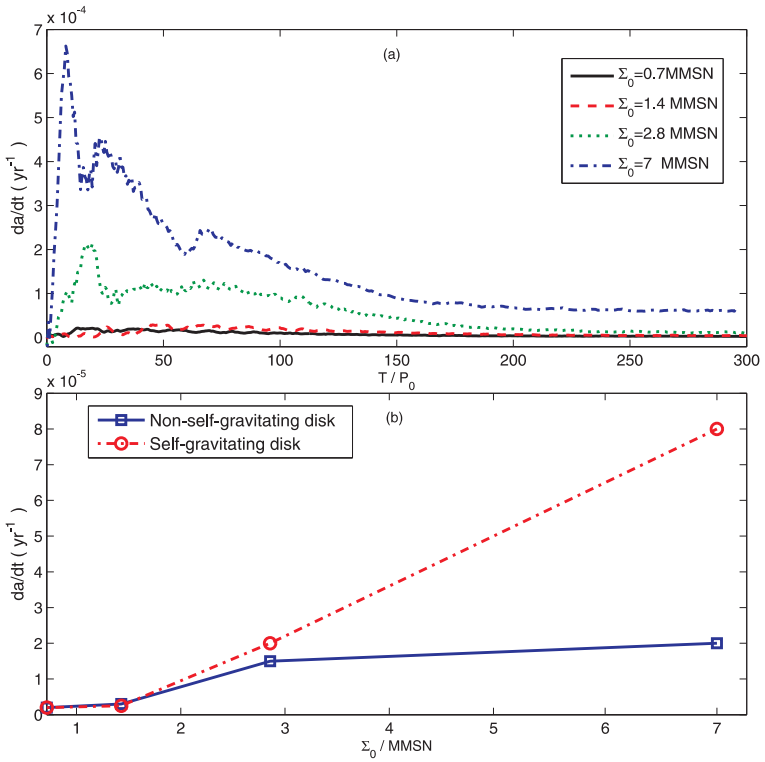

Fig. 8 The migration rate $\dot{a}$ versus surface density of a disk. Panel (a): The migration rate versus evolution time in the self-gravitating disk with various surface densities. $\dot{a}$ reaches a steady value after the gap is cleared ($t = 200$ orbits). Panel (b): The migration rate versus surface density of the disk. In the self-gravitating disk, $\dot{a}$ is proportional to Σ_0 (*circles*). However, in the non-self-gravitating disk, this relation is very weak (*squares*).

of the viscous dissipation when the gap is stable. Therefore, this variation in the rate of angular momentum will roughly indicate the effective viscosity $\nu_{\rm eff}$ in the disk. The associated results are shown in Figure 9. Panel (a) shows variations in the angular momentum versus time in a disk where $\Sigma_0 = 2.8$ MMSN. The result with the disk's self-gravity included is denoted by the dashed line, and the one with self-gravity excluded is denoted by the solid line. The large rate of variation before $t = 300\,P_0$ is the result of the process of gap formation, where gas is driven away by the tidal torque of the planet and results in a sharp decrease in the total mass of the disk. When the planet migrates significantly ($t > 500\,P_0$), the gap moves close to the inner boundary of the disk. The total angular momentum of the disk increases as the gap moves out of the disk's inner boundary (total mass increases). We only estimate the averaged rate of dissipation for the steady state of each run ($300\,P_0 < t < 500\,P_0$) and the results are shown in panel (b) of Figure 9. Since we do not adopt any artificial viscosity, for a non-self-gravitating disk $\nu_{\rm eff} = \nu_{\rm num}$, and for a self-gravitating disk $\nu_{\rm eff} = \nu_{\rm num} + \nu_{\rm sg}$. Our results show that the effective viscosity $\nu_{\rm eff}$ increases with Σ_0 in the self-gravitating disk. For the non-self-gravitating disk, the $\nu_{\rm eff}$ only slightly increases with Σ_0. Then we find that $\nu_{\rm sg}$ is roughly proportional to Σ_0 (stars in panel (b) of Fig. 9).

These results are in very good quantitative agreement with the migration rates we obtained above, except for the very high surface density $\Sigma_0 = 7$ MMSN, where the migration timescale ($\sim 1.2 \times 10^4$ yr) is much shorter than the viscous timescale ($\sim 2.1 \times 10^4$ yr). In fact, in such a dense

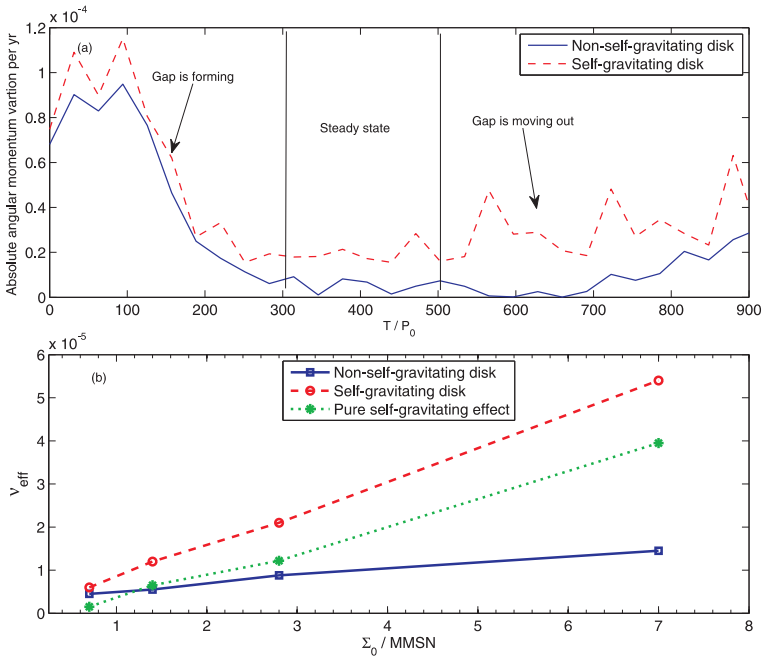

Fig. 9 Panel (a): The variations in absolute angular momentum versus evolution time. In a self-gravitating disk, the variation in angular momentum (*dashed line*) is always larger than that in a non-self-gravitating disk (*solid line*), where $\Sigma_0 = 2.8$ MMSN. Panel (b): The effective viscosity versus surface density of the disk. In a non-self-gravitating disk, the effective viscosity is mainly due to the numerical dissipation $\nu_{\rm eff} = \nu_{\rm num}$ (*squares*). However, in a self-gravitating disk, the gravitational turbulence is the main source of dissipation $\nu_{\rm eff} = \nu_{\rm num} + \nu_{\rm sg}$ (*circles*). The net gravitational viscosity effect $\nu_{\rm sg}$ is then shown by stars, which is proportional to Σ_0.

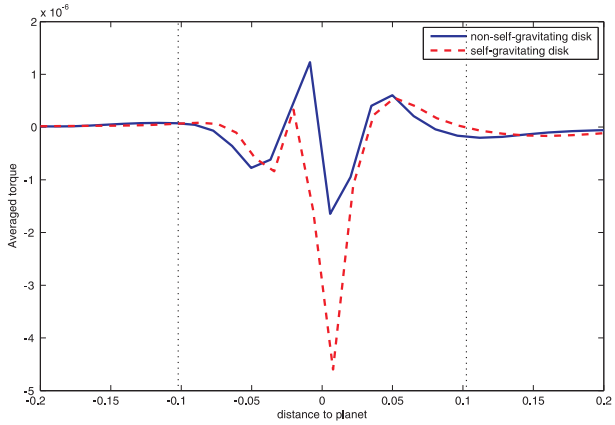

Fig. 10 Azimuthally averaged torque in the vicinity of the giant planet embedded in a very dense disk. The surface density of the disk is 7 MMSN. The torque is almost symmetric with the planet's position in the non-self-gravitating disk (*solid line*). However, in a self-gravitating disk, the planet suffers a net negative corotation torque (*dashed line*).

disk, the planet cannot clear a gap before it reaches the inner boundary (Fig. 6). As the Toomre Q parameter decreases along the radius of the disk, the gravitational turbulence becomes stronger as the radius increases. This generates a specific vorticity gradient across the corotation region of the giant planet and exerts a large negative corotation torque on the planet (Masset & Papaloizou 2003). We further calculate the torques exerted on the planet.

Figure 10 shows the azimuthally averaged torque as a function of the distance to the planet. It clearly shows that, in a non-self-gravitating disk, the torque density is almost symmetric with the position of the planet. There is a great negative torque within the corotation region of the planet, which drags the planet inward even faster. This result is consistent with that obtained by Baruteau et al. (2011).

4 CONCLUSIONS AND DISCUSSION

In this paper, we concentrate on the gap formation process under the effect of a disk's self-gravity. We first perform a series of 1-D simulations, where the disk's self-gravity is modeled by a gravitational effective viscosity $\nu_{\rm sg}$ and a time dependent azimuthally averaged self-gravity potential. We find that when the surface density of the disk is low, the self-gravity potential is too weak to affect the process of gap formation and the gravitational effective viscosity suppresses the growth of the gap. As we increase the surface density of the disk, the self-gravitational potential becomes stronger. It leads to a 'self-gravitational contraction' effect at each boundary of the gap and tends to enlarge the size of the gap. When the surface density exceeds a critical value, $\Sigma_0 > \Sigma_{\rm I}$, the net effect of self-gravity begins to benefit the gap formation process and the width of the gap increases with the surface density of the disk. We estimate that this critical surface density is around $\Sigma_{\rm I} \approx 0.8$ MMSN (Sect. 3.1). Since we recognize that the gravitational turbulence viscosity could not be described consistently in a 1-D simulation, we further perform a series of 2-D simulations where the disk's self-gravity is fully calculated by the real-time density distribution on the disk. We find that the width of the gap does not monotonically increase with the surface density in a self-gravitating disk. The gravitational turbulence becomes stronger as the disk's surface density increases and the associated effective viscosity overwhelms the effect of 'self-gravitational contraction' when the surface density of the disk exceeds another critical value $\Sigma_{\rm II}$. We estimate $\Sigma_{\rm II} \approx 3.5$ MMSN (Sect. 3.2.1). The values of $\Sigma_{\rm I}$ and $\Sigma_{\rm II}$ depend on the disk settings. Here we only give typical cases. In particular for $\Sigma_{\rm II}$, to find its exact value more surface densities beyond 2.8 MMSN need to be tested.

The associated rate of migration of the giant planet is also studied in this paper. Our 2-D simulations show that the rate of migration of the giant planet is slightly reduced in a self-gravitating disk with a moderate surface density ($\Sigma_0 < 2$ MMSN, see Fig. 7). However, it increases with the surface density of the disk where the gravitational turbulence becomes dominant. When the planet is still able to open a clear gap on the disk, its rate of migration is just proportional to the effective viscosity due to the gravitational turbulence. Furthermore, in a very dense disk with $\Sigma_0 > 7$ MMSN, the strong effective viscosity prevents the gap from forming, even for a Jupiter mass planet. The migration timescale then becomes much shorter than the viscous timescale of $\sim 10^4$ yr. This is caused by a large negative corotation drag which is the result of the specific vorticity gradient around the planet (Sect. 3.2.2).

According to our results we find that: (1) the effect of self-gravity may not be treated as simply an effective viscosity, especially for a moderate surface density. Our simulations reveal that the self-gravity plays two opposite roles in the process of gap formation at the same time and the net effect depends on the surface density of the disk. (2) The gravitational viscosity and the associated rate of migration for the giant planet increase with the surface density in a dense self-gravitating disk ($\Sigma_0 > 2.8$ MMSN). For a very dense disk with $\Sigma_0 \geq 7$ MMSN, where giant planets usually form, the gravitational effective viscosity is too strong to allow a clear gap to form and the migration timescale of a giant planet could be much shorter than the type II migration.

So, a giant planet is unlikely to stay at a large distance from the central star if the disk is still dense after the planet has formed. This is not a problem for the core accretion model. A planetary core usually needs $10^6 - 10^7$ yr to reach $10\,M_\oplus$ (Mizuno 1980), but the gas disk would be dispersed within 10^6 yr (Wolk & Walter 1996). If the giant planet could successfully form, its migration would be very slow or would even be stopped since the disk is already too thin to generate a large gravitational viscosity and could not effectively deliver enough angular momentum. The problem is, because of the long timescale required by the growth stage of the core, a giant planet is unlikely to form in a wide orbit by the core accretion model (Dodson-Robinson et al. 2009). For a multiple-planet system, if the outer planet is smaller than the inner one, the two inward migrating planets may become trapped into mutual mean motion resonance and migrate outward together (Zhang & Zhou 2010a; Zhang & Zhou 2010b). This could be an effective way to form giant planets at a large distance from their host star. For a single giant planet, however, it is still a problem. Some studies have shown that the radiative effect may affect the direction of the migration and could result in outward migration (Kalas et al. 2008; Bitsch & Kley 2010).

If a giant planet forms through the gravitational fragmentation of a very dense disk, it would probably migrate inward quickly. However, we emphasize that we do not adopt any cooling process in our 2-D simulations. This is because we do not want to introduce any poorly understood factors into our simulations, which would add too many uncertainties to the results. In our 2-D simulations, we assume a very cold disk with H/r fixed at 0.02 and adopt a locally isothermal equation of state. The cooling in our model is, therefore, perfect. Hence, the effective viscosity due to the self-gravitational turbulence increases with the surface density of the disk and results in fast inward migration in a dense disk. If a proper cooling process were included, the gravitational viscosity would become less effective, slowing the migration rate of the giant planet. This should be fully considered in future work.

We also notice that the existence of the giant planet may trigger the onset of gravitational instability in the disk. Strong spiral structures caused by the giant planet may generate a local minimum in Q and cause global instability when the averaged Q is still far above unity (Fig. 1 and Fig. 6). This effect depends on the mass of the giant planet and the disk where it is embedded. The details are also the subject of future work under preparation.

Acknowledgements This work is supported by the National Natural Science Foundation of China (Nos. 10833001, 10925313, 11003010 and 11078001) and the Research Fund for the Doctoral Program of Higher Education of China (Nos. 20090091110002 and 20090091120025). We are grateful to the High Performance Computing Center (HPCC) of Nanjing University for doing the numerical calculations in this paper on its IBM Blade cluster system. RW is supported by a UNSW Vice-Chancellor's Fellowship.

Appendix A: REFINED TREATMENT OF GRAVITY IN THE VICINITY OF A PLANET

When we calculate the torque exerted on the planet by a single cell of gas, the mass of this cell is usually treated as a point mass located at its center. When the planet travels very close to the center of the cell, we get a gravitational singularity and the planet would suffer extremely large gravitational force. However, since the density is uniform within a cell, the net force exerted on the planet should vanish because of the symmetry of the cell. A softening length is always needed to avoid the singularity, $\Phi_{\rm p} = -GM_{\rm p}/(|\boldsymbol{r} - \boldsymbol{r}_{\rm p}| + \varepsilon)$.

The softening length ε could only reduce the amplitude of the gravitational impulses. However, it could not result in the real gravity exerted on the planet. The choice of softening length is very tricky: a small ε could not effectively reduce the singularity, but a large one would eliminate too

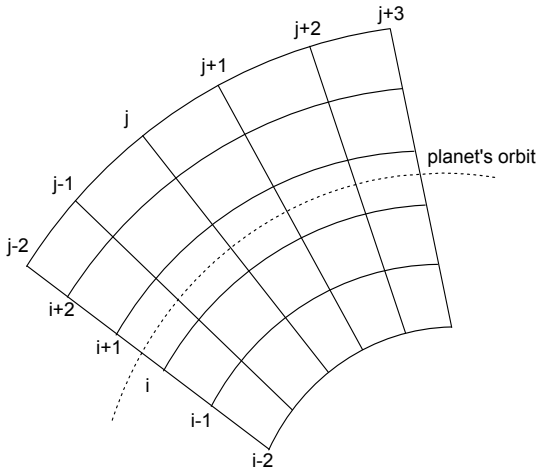

Fig. A.1 A 5×5-grid in the vicinity of the planet. When we calculate the torque exerted on the planet by this area, each cell is treated as a uniform area instead of a mass point. The dotted line shows the track of the planet in the comparison tests.

many physical effects. It is usually set to be a large fraction (e.g. $0.6 - 0.8$) of the scale height of the disk or the Hill radius of the planet. However, in a low resolution grid, the Hill radius only covers a few cells. Many local physical interactions between the planet and disk would be concealed if we chose ε to be comparable to the Hill radius. To more reliably model the gravity felt by the planet, we treat a single cell as a uniform area and the gravity exerted on the planet is an integration over this area, e.g. the force of gravity in the θ direction reads

$$F_{\theta,i,j} = GM_{\rm p}\sigma \int_{r_{i-\frac{1}{2}}}^{r_{i+\frac{1}{2}}} \int_{\theta_{j-\frac{1}{2}}}^{\theta_{j+\frac{1}{2}}} \frac{r^2 \sin(\theta - \theta_{\rm p})}{[r_{\rm p}^2 + r^2 - 2rr_{\rm p}\cos(\theta - \theta_{\rm p}) + \varepsilon]^{3/2}} dr d\theta \,. \tag{A.1}$$

ε is now a softening parameter for the integration which is very small. In our simulations, we set $\varepsilon = 10^{-4}$ in dimensionless units (the radius of the Roche lobe is now ~ 0.07 and the grid size is ~ 0.01). This treatment is performed in 5×5 cells around the cell where the planet is located. The cells outside this 5×5 region are treated as point masses as usual.

A comparison between different treatments of gravity is performed. We set a region with a 5×5-grid whose surface density is uniform and $\Sigma_0 = 1$ MMSN. Outside this region the surface density is set to be 0 (Fig. A.1). As the planet travels through this region, the gravity exerted on it should change smoothly and symmetrically from the positive to the negative extrema, and vanishes at the center of this area. The results are shown in Figure A.2. It is clear that the gravity is over-smoothed by the large gridsize of $\varepsilon \sim 1$ while the smaller one of $\varepsilon = 0.1 - 0.2 R_{\rm Roche}$ introduces nonphysical gravity impulses (panel (b) of Fig. A.2). Only the integration results with small $\varepsilon \sim 10^{-4 \sim -5}$ could void the nonphysical gravity impulses (panel (a) of Fig. A.2).

We also test the net torque of the whole disk under different treatments. The result is shown in Figure A.3. When we treat a cell in the disk as a point mass, the mutual gravity between the planet and the cell is very sensitive to the distance between them. When the planet travels through a high density cell and is very close to the center of the cell, its net torque will be dominated by this single cell. As the planet keeps passing by these point masses, the net torque exerted on it oscillates violently (the blue line in Fig. A.3). By contrast, when we treat a cell as a continuous uniform area, the net gravity from the cell vanishes when the planet is located at the center. As a result, the net torque becomes more smooth and reliable (the red line in Fig. A.3).

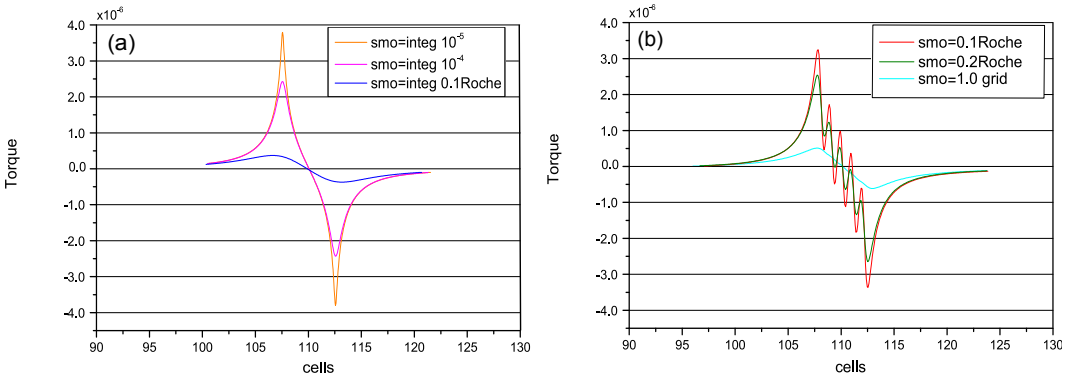

Fig. A.2 A test of the gravity torques exerted on the planet under different treatments. The planet travels in a circular orbit and the disk is divided into 256×512 cells. The center of the test area — a 5×5-grid (A.1) — is located at $x = 110$. The x-axis denotes the cell numbers (grid). Panel (a): cells around the planet are treated as uniform areas. We perform integration over each of these cells to find the torques they exert on the planet. The smoothing length used in the integration is shown in the legend. $\mathrm{smo} = \mathrm{integ}\ 10^{-5}$ means the softening length used in the integration is 10^{-5} in our unit. Panel (b): cells are treated as point masses. We assume the mass of a cell is concentrated at its center. The gravity between the planet and a cell center is calculated with a softening length to avoid a singularity. $\mathrm{smo} = 0.1\ \mathrm{Roche}$ means the softening length is one-tenth of the initial Roche radius of the planet which is ~ 0.069 in our units. $\mathrm{smo} = 1\ \mathrm{gridsize}$ means the softening length is equal to the grid size. Treating cells as point masses usually results in large gravity impulses or over-smoothed gravity, while treating cells as areas gives more smooth results and avoids any un-physical oscillations.

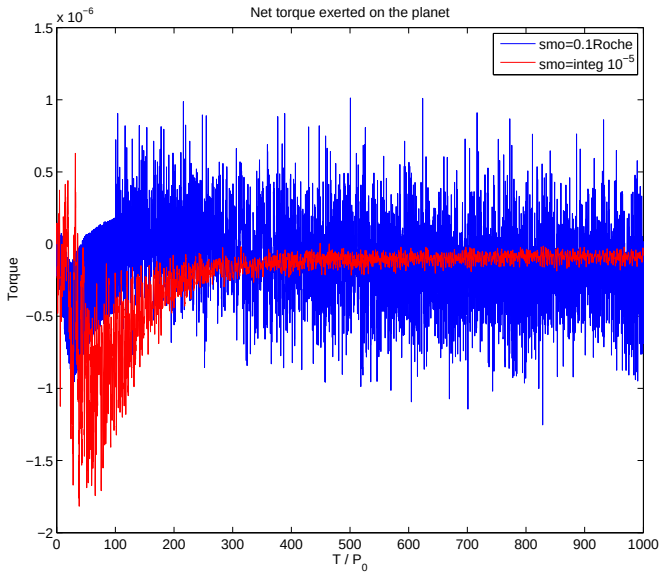

Fig. A.3 Another test of the gravity torques exerted on the planet under different treatments. We test the net torques exerted on the planet by the whole disk. The blue line shows the torque when we treat the cells as point masses, while the red line shows the torque when the cells around the planet are treated as uniform areas. The large oscillations in the former torque show the results when the planet travels through some dense cells (close to a large point mass). It is clear that the latter treatment gives more reliable results.

Appendix B: SELF-GRAVITY FORCE OF THE DISK

The effect of self-gravity from gas is included in the evolution of the disk. As the density distribution is changing with time, the gravitational potential of the disk evolves and needs to be determined by solving the Poisson equation at each time step: $\nabla^2\Phi_{\rm D} = 4\pi G\Sigma$. Integrating it over the disk in polar coordinates we get

$$\Phi_{\rm D}(r,\theta) = 2G\int\int \frac{\Sigma(r',\theta')}{[r^2+r'^2-2rr'\cos(\theta'-\theta)]^{1/2}} r'dr'd\theta' \,. \tag{B.1}$$

However, solving this equation directly is very "expensive" even in coarse resolution and the FFT method is one of the best choices.

The self-gravity force exerted on each cell in the radial direction reads

$$S_{\rm r}(r,\theta) = -2G\int\int \frac{\Sigma(r',\theta')r'[r-r'\cos(\theta'-\theta)]}{[r^2+r'^2-2rr'\cos(\theta'-\theta)]^{3/2}} dr'd\theta' \,. \tag{B.2}$$

Note that the right hand of the above equation is the convolution of $\Sigma(\boldsymbol{r}')$ and $K(\boldsymbol{r}-\boldsymbol{r}')$, where

$$K \equiv -2G\frac{[r-r'\cos(\theta'-\theta)]}{[r^2+r'^2-2rr'\cos(\theta'-\theta)]^{3/2}} \,. \tag{B.3}$$

According to the 'convolution theorem' we can get $S_{\rm r}$ by two Fourier transforms (F) and one reversed Fourier transform (F^{-1}) (Press et al. 1992)

$$S_{\rm r} = F^{-1}[F(\Sigma)F(K)] \,. \tag{B.4}$$

The kernel K in fact does not change with time and only needs to be calculated once at the beginning of the simulation. The self-gravity force in the azimuthal direction can be obtained similarly. Detailed introductions about this method can be found in many handbooks on computational methods, e.g. "Numerical recipes" (Press et al. 1992).

To avoid the self-gravity potential being abruptly cut off at each boundary of the disk, we add two buffer rings immediately outside the boundaries. The width of each buffer ring is 0.3 in our units and their surface densities do not evolve with time. We integrate the radial gravities of these two buffer rings and add them to the total gravity of the disk.

References

Baruteau, C., Meru, F., & Paardekooper, S.-J. 2011, MNRAS, 416, 1971
Bell, K. R., & Lin, D. N. C. 1994, ApJ, 427, 987
Bitsch, B., & Kley, W. 2010, A&A, 523, A30
D'Angelo, G., Kley, W., & Henning, T. 2003, ApJ, 586, 540
Dodson-Robinson, S. E., Veras, D., Ford, E. B., & Beichman, C. A. 2009, ApJ, 707, 79
Gammie, C. F. 2001, ApJ, 553, 174
Goldreich, P., & Tremaine, S. 1979, ApJ, 233, 857
Goldreich, P., & Tremaine, S. 1980, ApJ, 241, 425
Hayashi, C. 1981, Progress of Theoretical Physics Supplement, 70, 35
Hubeny, I. 1990, ApJ, 351, 632
Kalas, P., Graham, J. R., Chiang, E., et al. 2008, Science, 322, 1345
Kley, W. 1999, MNRAS, 303, 696
Lin, D. N. C., & Papaloizou, J. 1979, MNRAS, 188, 191
Lin, D. N. C., & Papaloizou, J. 1986, ApJ, 309, 846

Lin, D. N. C., & Papaloizou, J. C. B. 1993, in Protostars and Planets III, eds. E. H. Levy, & J. I. Lunine (Tucson: Univ. Arizona Press), 749
Lubow, S. H., Seibert, M., & Artymowicz, P. 1999, ApJ, 526, 1001
Marois, C., Macintosh, B., Barman, T., et al. 2008, Science, 322, 1348
Masset, F. S., & Ogilvie, G. I. 2004, ApJ, 615, 1000
Masset, F. S., & Papaloizou, J. C. B. 2003, ApJ, 588, 494
Mizuno, H. 1980, Progress of Theoretical Physics, 64, 544
Paardekooper, S.-J., Baruteau, C., Crida, A., & Kley, W. 2010, MNRAS, 401, 1950
Papaloizou, J. C. B., Nelson, R. P., & Snellgrove, M. D. 2004, MNRAS, 350, 829
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical Recipes in C: The Art of Scientific Computing (Cambridge: Cambridge Univ. Press)
Rice, W. K. M., Mayo, J. H., & Armitage, P. J. 2010, MNRAS, 402, 1740
Stamatellos, D., Whitworth, A. P., Bisbas, T., & Goodwin, S. 2007, A&A, 475, 37
Takeuchi, T., Miyama, S. M., & Lin, D. N. C. 1996, ApJ, 460, 832
Tanaka, H., Takeuchi, T., & Ward, W. R. 2002, ApJ, 565, 1257
Ward, W. R. 1997, Icarus, 126, 261
Winters, W. F., Balbus, S. A., & Hawley, J. F. 2003, ApJ, 589, 543
Wolk, S. J., & Walter, F. M. 1996, AJ, 111, 2066
Zhang, H., Yuan, C., Lin, D. N. C., & Yen, D. C. C. 2008, ApJ, 676, 639
Zhang, H., & Zhou, J.-L. 2010a, ApJ, 714, 532
Zhang, H., & Zhou, J.-L. 2010b, ApJ, 719, 671