Nonlinear outcome of gravitational instability in an irradiated protoplanetary disc

Shigenobu Hirose,1⋆ and Ji-Ming Shi2

1Department of Mathematical Science and Advanced Technology, JAMSTEC, Yokohama 236-0001, Japan
2Department of Astrophysical Sciences, Princeton University, 4 Ivy Ln, Princeton, NJ 08544

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT
Using local three dimensional radiation hydrodynamics simulations, the nonlinear outcome of gravitational instability in an irradiated protoplanetary disc is investigated in a parameter space of the surface density $\Sigma$ and the radius $r$. Starting from laminar flow, axisymmetric self-gravitating density waves grow first. Their self-gravitating degree becomes larger when $\Sigma$ is larger or the cooling time is shorter at larger radii. The density waves eventually collapse owing to non-axisymmetric instability, which results in either fragmentation or gravito-turbulence after a transient phase. The boundaries between the two are found at $r \sim 75$ AU as well as at the $\Sigma$ that corresponds to the initial Toomre’s parameter of $\sim 0.2$. The former boundary corresponds to the radius where the cooling time becomes short, approximating unity. Even when gravito-turbulence is established around the boundary radius, such a short cooling time inevitably makes the fluctuation of $\Sigma$ large enough to trigger fragmentation. On the other hand, when $\Sigma$ is beyond the latter boundary (i.e. the initial Toomre’s parameter is less than $\sim 0.2$), the initial laminar flow is so unstable against self-gravity that it evolves into fragmentation regardless of the radius or, equivalently, the cooling time. Runaway collapse follows fragmentation when the mass concentration at the centre of a bound object is high enough that the temperature exceeds the $\text{H}_2$ dissociation temperature.

Key words: protoplanetary discs — gravitational — hydrodynamics — radiative transfer — instabilities — turbulence

1 INTRODUCTION
Recent observations, including those by the Atacama Large Millimeter/submillimeter Array (ALMA), have revealed massive protostellar/protoplanetary discs in young stellar class 0 and class I systems (e.g. Andrews et al. 2013; Najita & Kenyon 2014; Pérez et al. 2016). In such massive discs, self-gravity is a very important and relevant aspect of physics. Specifically, when the condition

$$Q \equiv \frac{c_s \kappa}{\pi G \Sigma} < 1$$

is met, the disc is subject to gravitational instability (GI; Toomre 1964). Here, $Q$ is called Toomre’s parameter whilst $c_s$ is the sound speed, $\kappa$ is the epicyclic frequency, and $\Sigma$ is the surface density. In Keplerian discs, $\kappa$ is equal to the orbital frequency $\Omega$.

The nonlinear development of GI generally leads to formation of spiral density waves; especially when they are tightly wound, they may be described as so-called gravito-turbulence in the local approximation (see Kratter & Lodato 2016, for a recent comprehensive review on GI in protoplanetary discs). The shear stress associated with the spiral density waves radially transfers angular momentum, which evolves the radial structure of the disc. Another outcome of GI is fragmentation, or formation of self-gravitationally bound objects, which may eventually become companion stars, brown dwarfs, or gas giant planets. Thus, the nonlinear outcome of GI largely affects the growth and evolution of the disc, but in different forms depending on whether formation of spiral density waves (gravito-turbulence) or fragmentation occur. Therefore, what determines the nonlinear outcome of GI is of great interest, and thus has been widely explored by numerical hydrodynamics simulations.

In the framework of the shearing box, Gammie (2001) first revealed the importance of the cooling time. He showed that fragmentation occurred when cooling was fast enough, where $\beta \equiv t_{\text{cool}} \Omega < 3$, when the cooling time $t_{\text{cool}}$ was assumed constant everywhere for simplicity. Since then, the fragmentation condition in terms of $\beta$ has been the main
focus of interest. It has been extensively studied using various types of numerical methods and cooling prescriptions, both in local and global simulations (e.g. Johnson & Gammie 2003; Stamatello & Whitworth 2009; Cassins et al. 2010; Baehr & Klahr 2015; Riols & Latter 2016), and especially for protoplanetary discs by many authors mostly motivated by the formation of gas giants via GI (Boss 1997, 1998; Durisen et al. 2007; Zhu et al. 2012).

However, the exact value of the critical $\beta$ for fragmentation remains an open question. Non-convergence of the fragmentation criterion may arise from numerical artefacts (Meru & Bate 2010; Lodato & Clarke 2011; Meru & Bate 2012), inherent stochasticity of fragmentation (Paardekooper 2012; Hopkins & Christiansen 2013), the dimension (i.e. 2D vs. 3D) (Young & Clarke 2015) or the fact that there is no physical temperature floor in the $\beta$ cooling prescription (Lin & Kratter 2016). Irradiation can be a main heating source in cool protoplanetary discs subject to GI, and thus may affect the fragmentation criterion (Rice et al. 2011). It has also been suggested that a fragmentation criterion in terms of the $\alpha$ parameter (Shakura & Sunyaev 1973) may be more general than the cooling time $\beta$ (Rice et al. 2005).

On the other hand, some authors have claimed that the cooling time $\beta$ is not necessarily the primary factor for fragmentation. Rogers & Wadsley (2012) proposed that the Hill radius plays an essential role in fragmentation; that is, fragmentation occurs when the width of a spiral density wave is less than the Hill radius, although the width itself may be determined by the balance between cooling and heating. Tsukamoto et al. (2014) also found that fragmentation discs have narrower spiral density waves than non-fragmentation discs, and emphasised that the local minimum of Toomre’s parameter inside the spiral density waves, $Q_{\text{min}}$, determines whether they fragment (for $Q_{\text{min}} \lesssim 0.2$) or not. Takahashi et al. (2016), based on a linear analysis, related the critical Toomre’s parameter below which fragmentation occurs and the width of a density wave, and derived a fragmentation condition as $Q_{\text{min}} \lesssim 0.6$ for typical density waves in their global simulations.

In a series of papers (Hirose & Shi 2017, hereafter Paper I, and this paper), we have examined the fragmentation condition, as well as the gravito-turbulence, in an irradiated protoplanetary disc in the framework of a local shearing box. Given that there are many studies in the literature, our stance is as follows. Because temperature is one of the most important quantities that control GI (see eq. 1), correct thermodynamic analysis is essential to study the nonlinear evolution of GI in realistic protoplanetary discs. Therefore, we perform 3D radiation hydrodynamics simulations with a realistic opacity and a realistic equation of state (EOS), and include the irradiation heating by the central star. This is an extended work from Shi & Chiang (2014), who performed 3D local shearing box simulations using the $\beta$ cooling and simple optically-thin cooling prescriptions. The local shearing box has two physical parameters, the distance from the central star $r$ and the surface density $\Sigma$. In Paper I, we mainly studied the dependence on $\Sigma$ of the nonlinear outcome of GI at a single radius of $r = 50$ AU. It is therefore the goal of this paper to present the nonlinear outcome of GI in a relatively complete $\Sigma$-$r$ parameter space. Especially, we map out the regions in which the disc is laminar, turbulent, or fragmenting in the $\Sigma$-$r$ parameter space, and provide physical interpretations to such a phase diagram. In this sense, this work is also an extension of Johnson & Gammie (2003), who presented similar mapping based on 2D shearing box simulations but lacked a realistic EOS and did not consider irradiation heating.

This paper is organised as follows. After we briefly describe our numerical methods in Section 2, we present the nonlinear outcome of GI and discuss the properties of gravito-turbulence as well as the fragmentation condition in Section 3. In Section 4, we compare our results with previous studies and discuss some implications. Finally, we provide a summary in Section 5.

2 METHODS

In this section, we explain the methods we used only briefly, because they are the same as those used in Paper I, which the reader may refer to for additional details.

2.1 Basic equations and numerical schemes

The basic equations solved in our simulations are hydrodynamics equations with Poisson’s equation for self-gravity and frequency-integrated angular-moment equations of the radiative transfer:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0, \quad (2)$$

$$\frac{\partial (\rho \mathbf{v})}{\partial t} + \nabla \cdot (\rho \mathbf{v} \mathbf{v}) = -\nabla p - \rho \nabla \Phi + \frac{\kappa B \rho}{c} F, \quad (3)$$

$$\frac{\partial e}{\partial t} + \nabla \cdot (e \mathbf{v}) = -(\mathbf{v} \cdot \nabla) p - (4\pi B(T) - c E) \kappa T \rho, \quad (4)$$

$$\frac{\partial (\kappa T \rho)}{\partial t} + \nabla \cdot (e \mathbf{v}) = -\nabla v \cdot P + (4\pi B(T) - c E) \kappa T \rho - \nabla \cdot F, \quad (5)$$

$$\nabla \cdot \Phi = 4\pi G \rho, \quad (6)$$

where $\rho$ is the gas density, $e$ is the gas internal energy, $p$ is the gas pressure, $T$ is the gas temperature (assumed to be the same as the dust temperature), $E$ is the radiation energy density, $P$ is the radiation pressure tensor, $F$ is the radiation energy flux, $\mathbf{v}$ is the velocity field vector, $B(T) = \sigma_B T^4/\pi$ is the Planck function ($\sigma_B$ is the Stefan-Boltzmann constant), and $c$ is the speed of light. The flux limited diffusion approximation was employed to close the angular-moment equations, where the first and second moments, $F$ and $P$, are related to the zeroth moment, $E$ (Turner & Stone 2001).

The EOS, $p = \rho_e(\rho_e)$ and $T = T(\rho_e, \rho)$, is an updated version of that used in Tomida et al. (2013) in their star formation simulations. The Rosseland- and the Planck- mean opacity, $\kappa_T(\rho, T)$ and $\kappa_P(\rho, T)$, are the same as those used in Hirose (2015), where the dust and gas opacity are taken from, respectively, Semenov et al. (2003) and Ferguson et al. (2005).

We used the shearing box approximation to model a local patch of an accretion disc as a co-rotating Cartesian frame ($x, y, z$) with the linearized Keplerian shear flow, $\mathbf{v}_K = -(3/2) \Omega \mathbf{e}_\theta \times \mathbf{r}$. The inertial forces in the co-rotating frame and the vertical component of the external gravity by the central star are added as source terms in the equation of motion (3). Shearing-periodic, periodic, and outflow boundary conditions were imposed.
Nonlinear outcome of gravitational instability

conditions are applied to the boundaries in the $x$, $y$, and $z$ direction, respectively (Hirose et al. 2006).

We employed ZEUS (Stone & Norman 1992) to solve the above equation set. An orbital advection algorithm (Stone & Gardiner 2010) was implemented for accurate calculation in a wide shearing box. Poisson’s equation with the vacuum boundary condition in the $z$ direction was solved by Fast Fourier Transforms (Koyama & Ostriker 2009). The irradiation heating rate, evaluated by solving a time-independent radiative transfer equation (ignoring scattering), was added as a source term in equation (4). The nonlinear radiative transfer terms in the energy equations (4) and (5) were coupled to be solved time-implicitly using the Newton-Raphson method. The kinetic energy dissipating either numerically or physically was captured in the form of gas internal energy, which guaranteed conservation of the sum of the kinetic and internal energies (Hirose et al. 2006).

2.2 Parameters and the initial conditions

A stratified shearing box has two physical parameters. One is the orbital frequency $\Omega = \sqrt{GM_*/r^3}$ [s$^{-1}$], which appears in the inertial force terms and the shearing periodic boundary condition. Here $M_*$ is the mass of the central star, $r$ is the distance from the central star, and $G$ is the gravitational constant. The other is the (horizontally-averaged) surface density $\Sigma$ [g cm$^{-2}$], which represents the amount of gas in the box. In our simulations, the value of $\Sigma$ varied from the initial value $\Sigma_0$ due to the outflow boundary condition as well as the density floor (see Hirose et al. 2006, for details). However, because the relative difference was typically small (a few percent in one hundred orbits at largest), we do not explicitly distinguish $\Sigma$ and $\Sigma_0$ in this paper.

The parameters of irradiation heating are the energy flux $F_{\text{irr}} = (R_*/r)^2 \sigma T_4^4$ [erg cm$^{-2}$ s$^{-1}$] and the grazing angle $\theta$, where $R_*$ and $T_4$ are, respectively, the radius and the effective temperature of the central star. We assumed that $T_4 = 4000 \text{ K}$, $M_*/M_\odot = 1$, and $R_* = 1R_\odot$. Also, we fixed the grazing angle as $\theta = 0.02$ for simplicity because the main effect of the irradiation heating (i.e. setting a physical temperature floor near the midplane) only weakly depends on $\theta$ (see eq. (7) below).

The initial disc was set up to be isothermal and in hydrostatic equilibrium ignoring self-gravity, where a mean molecular weight $\mu = 2.38$ and adiabatic exponent $\gamma = 5/3$ were used. The isothermal temperature was evaluated using the radiative equilibrium disc model (Equation 12a in Chiang & Goldreich 1997) as

$$T_0 = \left(\frac{\theta}{3}\right)^{\frac{1}{2}} \left(\frac{R_*}{r}\right)^{\frac{1}{2}} T_*.$$  \hspace{1cm} (7)

The initial radiation field $E_0$ was assumed to be in thermal equilibrium with the gas, where $E_0 = (4\pi\beta\zeta/c)T_0^4$. The initial velocity field was the linearized Keplerian shear flow, whose $x$ and $z$ components were perturbed randomly up to 0.5% of the local sound speed $c_s = \sqrt{\Gamma (\rho/p)}$, where $\Gamma = d\ln p/d\ln \rho$ is the generalised adiabatic exponent.

In all runs, the box size and the number of cells were set as $(L_x, L_y, L_z) = (24H, 24H, 12H)$ and $(N_x, N_y, N_z) = (128, 128, 64)$, respectively. Here and hereafter, the scale height of the initial isothermal disc $H = \sqrt{2RT_0/(\mu\Omega^2)}$ is used as the unit length, where $R$ is the gas constant.
Table 1. List of the selected five runs compared to see dependence of the nonlinear outcome on $\Sigma$ (runs S0, S1, and S2) and $r$ (runs R0, R1, and R2). Run S0 is identical to Run R1.

| Label |
|-------|
| $\Sigma_{0.82}$ (80 g cm$^{-2}$) |
| $\Sigma_{0.55}$ (120 g cm$^{-2}$) |
| $\Sigma_{0.22}$ (300 g cm$^{-2}$) |
| $\Sigma_{0.24}$ (300 g cm$^{-2}$) |
| $\Sigma_{0.78}$ (30 g cm$^{-2}$) |
| Radius $r$ |
| 50 AU |
| 50 AU |
| 50 AU |
| 25 AU |
| 50 AU |
| 90 AU |
| outcome |
| turbulence |
| turbulence |
| runaway collapse |
| turbulence |
| turbulence |
| fragmentation |
| Fig. # |
| 4 |
| 5 |
| 6 |
| 7 |
| 4 |
| 15 |
| 8 |

Figure 2. Time evolution of the local Toomre’s parameter $Q_*$. The upper panel compares runs S0 (blue), S1 (red) and S2 (black), whilst the lower panel compares runs R0 (orange), R1 (blue), and R2 (green); refer to Table 1 for the labels of the runs. The open circle on each curve denotes the end of the growth of the axisymmetric self-gravitating density waves. The horizontal grey line denotes the value of 0.2.

Figure 3. Same as Fig. 2, except for the local adiabatic exponent $\Gamma_*$. The horizontal grey line denotes the critical value of $\Gamma = 4/3$.

3 RESULTS

3.1 Diagnostics

For diagnostics, we use simple and density-weighted volume averages for a quantity $f(x, y, z)$, defined as

$$
\langle f \rangle = \frac{\int f(z) \, dz}{\int dz}
$$

simple volume average, \hspace{1cm} (8)

$$
\langle f \rangle_{\text{mid}} = \frac{\int f(z) \rho(z) \, dz}{\int \rho(z) \, dz}
$$

density-weighted volume average, \hspace{1cm} (9)

where $\langle f \rangle \equiv \iint f(x, y, z) \rho(x, y, z) \, dx \, dy \, dz / \iint \rho(x, y, z) \, dx \, dy$ is the horizontal average.

Also, we define locally the Toomre’s parameter and the normalised cooling time in the midplane, respectively, as

$$
Q(x, y) = \frac{c_s(x, y) \kappa}{\pi G \sigma(x, y)},
$$

(10)

$$
\beta(x, y) = \frac{\kappa e(x, y, z = 0) \Omega}{-q'(x, y, z = 0)},
$$

(11)

where $\sigma(x, y) \equiv \int \rho(x, y, z) \, dz$ is the local surface density, $c_s(x, y) \equiv \int c_s(x, y, z) \rho(x, y, z) \, dz / \sigma(x, y)$ is the density-weighted average of the sound speed, and $q' \equiv -\kappa R (4\pi B(T) - \epsilon E)$ is the radiative cooling term in equation.
Nonlinear outcome of gravitational instability

(4). In this paper, when evaluating Toomre’s parameter $Q$, we assume $\kappa = \Omega$ (the Keplerian rotation) except in Section 3.5.

As we are interested in the nonlinear outcome of GI, we often examine quantities evaluated at the cell where the self-gravitational energy, $E_{sg} \equiv \rho \Phi / 2$, takes the minimum value on the midplane. Hereafter, the subscript “∗” denotes a quantity at the cell of minimum $E_{sg}$ on the midplane; for example, $(x_∗, y_∗)$ denotes the horizontal position of that cell.

3.2 Nonlinear development and outcome of GI

We have run 74 simulations in total to explore the parameter ranges of $\Sigma_1 \lesssim \Sigma \lesssim \Sigma_0.2$ and 15 AU $\leq r \leq 90$ AU.\(^1\) Here, $\Sigma_0$ denotes the surface density that corresponds to the initial Toomre’s parameter $Q_0$; that is, $Q_0 = c_0 \Omega / (\pi G \Sigma_0)$, where $c_0$ is the initial sound speed. The nonlinear outcome is summarised as a phase diagram in the $\Sigma-\Gamma$ space in Fig. 1.

In Paper I, we found at $r = 50$ AU that gravito-turbulence is sustained for a certain range of $\Sigma$ whilst GI is not driven below that range and runaway collapse occurs above it. Such dependence on $\Sigma$ can be seen at $r \lesssim 60$ AU. Specifically, GI is driven when $\Sigma$ exceeds $\Sigma_1$ and runaway collapse occurs when $\Sigma$ exceeds $\Sigma_0.2$. On the other hand, at $r = 90$ AU, when GI is driven, the outcome is always fragmentation (or runaway collapse) and no gravito-turbulence is sustained. The outcome at $r = 75$ AU is somewhat intermediate between $r \lesssim 60$ AU and $r = 90$ AU.

Among the total 74 runs, we especially inspect in detail the five runs listed in Table 1 to observe the dependence of the outcome on $\Sigma$ (runs S0, S1 and S2) as well as on $r$ (runs R0, R1 and R2). In Fig. 2, we compare the time evolution of $Q_*$ amongst them, where $Q_*$ is the local Toomre’s parameter evaluated at the cell of minimum $E_{sg}$ as

$$Q_* = \frac{c_2(x_*, y_*)}{\pi G \sigma(x_*, y_*)}. \quad (12)$$

The value of $Q_*$ in the final state is found to provide a good measure for distinguishing the outcome quantitatively as follows:

- gravito-turbulence: $0.1 < Q_*$,
- fragmentation: $0.01 < Q_* < 0.1$,
- runaway collapse (fragmentation): $Q_* < 0.01$.

Here, runaway collapse is a special case of fragmentation in which gas pressure cannot stop the gravitational collapse due to softening of the EOS. Fig. 3 compares the time evolution of $\Gamma_*$ (the adiabatic exponent $\Gamma$ at the cell of minimum $E_{sg}$) amongst the five runs. In run S2, fragmentation was followed by runaway collapse because the core temperature exceeded the hydrogen dissociation temperature and thus $\Gamma$ dipped below the critical value of 4/3. On the other hand, in run R2, although fragmentation did occur, runaway collapse did not occur and a pressure-supported clump survived because $\Gamma$ remained well above the critical value owing to insufficient rise of the core temperature. Once runaway collapse occurred, we stopped the calculation because the following evolution at smaller scales could not be treated in our simulation with a fixed grid.

Regardless of the outcome, the nonlinear evolution of GI followed the same steps:

(i) destabilisation of the initial laminar flow,
(ii) growth of almost axisymmetric density waves,
(iii) non-axisymmetric destabilisation of the density waves,
(iv) collapse of the density waves into a transient phase,
(v) the final outcome.

In Figs. 4, 5, 6, 7, and 8, we show time series snapshots of gas temperature $T(x, y, z = 0)$, density $\rho(x, y, z = 0)$, and the cooling time $\beta(x, y)$ for the selected five runs. In each figure, the top row corresponds to the epoch when the non-axisymmetric deformation of the almost axisymmetric density waves becomes clear by eye-measurement (step (iii)). The middle row shows the following transient phase where complicated interactions between density waves and clumps are apparent (step (iv)), and the bottom row shows the final outcome (step (v)).

Here we note that the almost axisymmetric density waves in step (ii) are not simple nonlinear manifestations of the initial most unstable modes in step (i); rather, they emerged as a result of nonlinear interactions between the initial unstable modes. As the axisymmetric density waves grew, their Toomre’s parameter decreased and they became strongly self-gravitating (Fig. 2). Eventually, without exception, they became unstable against non-axisymmetric perturbations (step (iii)) and collapsed into the transient phase (step (iv)). We will discuss in detail the non-axisymmetric instability of the density waves in Section 3.5.

Hereafter, we refer to the almost axisymmetric density waves in step (ii) simply as axisymmetric density waves, omitting “almost”, for simplicity.

3.3 Gravito-turbulence

Gravito-turbulence is the state where turbulent dissipation and radiative cooling balance, which was established for a finite range of $\Sigma$ at $r \lesssim 75$ AU. In this section, we examine how the properties of the gravito-turbulence depend on $\Sigma$ and $r$. Quantities that will be discussed in this section are the ones time-averaged for a period in which the gravito-turbulence is sustained (see Appendix A for the period). The time interval when recording the numerical data was 0.01 orbits, except for the bottom panel in Fig. 9, where the interval was 1 orbit.

3.3.1 Cooling time $\beta$ and Toomre’s parameter $Q$

In the top panel of Fig. 9, the space-averaged cooling time defined as

$$\beta_{ave} = \frac{\langle \epsilon \rangle_{mid}}{-\langle q \rangle_{mid}} \quad (13)$$

is shown in terms of $\Sigma$ and $r$. At each radius, $\beta_{ave}$ is almost constant except in some small $\Sigma$ cases, where it is relatively large because extra heating by irradiation raised the midplane thermal energy $\langle \epsilon \rangle_{mid}$. On the other hand, $\beta_{ave}$ strongly depends on $r$, as explicitly shown in the inset. Specifically, $\beta_{ave}$ is as small as $\sim 1.5$ at $r = 75$ AU, whilst

\(^1\) The specific value of $\Sigma$ as well as that of $r$ in each simulation are given in Appendix A.
it is as large as $\sim 400$ at $r = 15$ AU. As shown in the middle panel of Fig. 9, the space-averaged Toomre’s parameter defined as

$$Q_{\text{ave}} \equiv \frac{\langle c_s \rangle_{\text{mid}}}{\pi \Omega \Sigma}$$

(14)

is around unity for all runs, and it depends on $\Sigma$ and $r$ only weakly.

The strong dependence of $\beta_{\text{ave}}$ on $r$ is derived as follows (e.g. Clarke 2009; Paardekooper 2012). Because the disc is optically thick, the cooling time is evaluated as the vertically integrated thermal energy $T_{\text{mid}} \Sigma$ divided by the radiative diffusion cooling rate $(T_{\text{mid}} / \kappa_R \Sigma)$, where $T_{\text{mid}}$ represents the midplane temperature. Then,

$$\beta \propto \frac{(T_{\text{mid}} \Sigma) \Omega}{\Omega_{\text{mid}}^2 \kappa_R \Sigma} \propto T_{\text{mid}}^{-1} \Sigma^2 \Omega \propto \left(Q^{-2} \Omega^{-2} \Sigma^{-2}\right) \Sigma^2 \Omega \propto Q^{-2} \Omega^2 \Sigma^0.$$  

(15)

Here we used the dependence of opacity on temperature $\kappa_R \propto T^2$ as shown in the bottom panel of Fig. 9 as well as the definition of Toomre’s parameter $Q \propto T_{\text{mid}}^{1/2} \Omega \Sigma^{-1}$, to eliminate $T_{\text{mid}}$. As we stated above, $Q_{\text{ave}}$ only weakly depends on $\Sigma$ and $r$. Therefore, if we ignore $Q$ in equation (15), it becomes $\beta \propto \Omega^2 \Sigma^0$, which is roughly consistent with the dependence shown in the inset of the top panel.
3.3.2 Shear stress and $\alpha$

The upper panel of Fig. 10 shows the dependence on $\Sigma$ and $r$ of the $\alpha$ parameter defined as

$$\alpha = \frac{\int \langle W_{xy} \rangle \, dz}{\int \langle P_{\text{thermal}} \rangle \, dz},$$

(16)

where the thermal pressure $P_{\text{thermal}}$ is the sum of the gas and radiation pressures, and the shear stress $W_{xy}$ is the sum of the gravitational and Reynolds shear stresses. The value of $\alpha$ widely ranges from $\sim 4 \times 10^{-3}$ at $r = 15$ AU to $0.7$ at $r = 75$ AU, and is roughly proportional to $\Omega^{-3}$, as shown in the inset. Comparing the upper panel of Fig. 10 with the top panel of Fig. 9, the dependence of $\alpha$ on $\Sigma$ and $r$ is almost opposite to that of $\beta_{\text{ave}}$. This is expected from the thermal balance condition, in that the vertically-integrated cooling rate $\int \langle (\epsilon) + (E) \rangle d\Omega/\beta_{\text{ave}}$ equals the vertically-integrated stress work $\frac{3}{2} \Omega \int \langle W_{xy} \rangle \, dz$, which requires that $\alpha \propto \beta_{\text{ave}}^{-1}$ (Gammie 2001).

Because the stress work $\frac{3}{2} \Omega \int \langle W_{xy} \rangle \, dz$ is equated to the release rate of gravitational energy $\frac{3}{16} \dot{M} \Omega^2$, the mass accretion rate $\dot{M}$ can be evaluated as

$$\dot{M} = \frac{\int \langle W_{xy} \rangle \, dz}{\Omega/2\pi},$$

(17)

which strongly depends on both $\Sigma$ and $r$, as shown in the lower panel of Fig. 10. The dependence of $\dot{M}$ on $\Sigma$ and $r$ can be derived as

$$\dot{M} \propto \alpha(T_{\text{mid}}/\Sigma) \Omega^{-1} \propto \Omega^{-3} \left( (Q^2 \Omega^2 \Sigma^2) \Omega^{-1} \propto Q^2 \Omega^{-6} \Sigma^3 \right),$$

(18)

where $\alpha \propto \Omega^{-3}$ is substituted. Again, if we ignore the very weak dependence of Toomre’s parameter on $\Sigma$ and $\Omega$, Eq. (18) becomes $\dot{M} \propto \Sigma^3 \Omega^{-6}$, which is roughly confirmed in the inset of the lower panel. Eq. (18) also states that the gravito-turbulence can sustain accretion flows of larger $\dot{M}$ at larger radii. Specifically, we note that the maximum accretion rate of $\dot{M} \sim 10^{-4} M_\odot \text{yr}^{-1}$ is realised at $r = 50 \sim 60$ AU, whilst the minimum accretion rate of $\dot{M} \sim 10^{-7} M_\odot \text{yr}^{-1}$ is realised at $r = 15 \sim 20$ AU.

Figure 5. Same as Fig. 4, except for run S1.

\[\text{Figure S5. Same as Figure S4, except for run S1.}\]
Also, the fraction of the gravitational stress to the total stress is generally around $\sim 0.5$ with a slight decrease with $\Sigma$ at each radius, as shown in the lower panel of Fig. 10.

### 3.3.3 Time variations

In this section, we examine the temporal behaviour in the gravito-turbulence at different radii. The top panel of Fig. 11 shows time variations of volume-averaged internal energy $\langle e \rangle$ of a representative run at different radii. At larger radii ($r \gtrsim 50$ AU), the time variation appears stochastic, but, as the radius decreases, it becomes more quasi-periodic, with longer periodicity.

To quantify the typical time scale of variation of $\langle e \rangle$, we computed the one-sided power spectral density of $\langle e \rangle$,

$$P(f) = \left| \int_0^\infty \langle e \rangle e^{-2\pi ift} dt \right|^2 + \left| \int_0^\infty \langle e \rangle e^{2\pi i(-ft)} dt \right|^2,$$

where $f$ is the frequency, and then take its cumulative fraction

$$C(f) = \frac{\int_0^f P(f')df'}{\int_0^\infty P(f')df'}$$

(20)

where $P(f = 0)$ is excluded in the total power (the denominator). The result is shown in the middle panel of Fig. 11, and we observe a clear trend where, as the radius decreases, the power at longer periods provides a major contribution to the total power.

To quantify the trend above, we measure a frequency $f_0$ below which the cumulative fraction exceeds a critical value of 0.3 (that is, $C(f) > 0.3$ for $f < f_0$), supposing that the inverse of the frequency $f_0$ represents the typical time scale of $\langle e \rangle$. The choice of the critical value is arbitrary, but here it is chosen so that the trend can be best observed. The bottom panel of Fig. 11 shows that the typical time scale $1/f_0$ has a fairly strong dependence on $r$. Also, the figure shows that it has good correlation with the space-averaged...
cooling time $\beta_{\text{ave}}$, indicating that the time variation of $\langle e \rangle$ is mainly determined by the space-averaged cooling time.

### 3.4 Boundaries between gravito-turbulence and fragmentation

As seen in Fig. 1, there are apparently two types of fragmentation boundaries. One is at $r \sim 75$ AU and the other is at $\Sigma \sim \Sigma_{0.2}$.

Starting from Gammie (2001), a consensus has been established that fragmentation occurs when the cooling time $\beta$ is less than a critical value of order unity. This is because a clump contracts to a bound object if the stochastic shock heating fails to catch up to the imposed cooling, which is especially expected when the cooling time $\beta$ is short (e.g. Paardekooper 2012). As we discussed in Section 3.3.1, the space-averaged cooling time scales as $\beta_{\text{ave}} \propto \Omega^2 \Sigma^0$ and is as short as $\beta_{\text{ave}} \sim 1$ at $r = 75$ AU. Therefore, the fragmentation boundary apparent at $r \sim 75$ AU in our simulations corresponds to the minimum cooling time ($\beta_{\text{ave}} \sim 1$) that can sustain the gravito-turbulence without fragmentation. Alternatively, because $\alpha \propto \beta_{\text{ave}}^{-1}$ (as we discussed in Section 3.3.2), the minimum cooling time can be redefined by the maximum $\alpha$ ($\sim 1$) that arises in the gravito-turbulence.

A short cooling time in the gravito-turbulence has an important consequence. That is, the density fluctuation is expected to be anti-correlated to the averaged cooling time as

$$\frac{\delta \Sigma}{\Sigma} \propto \frac{1}{\sqrt{\beta_{\text{ave}}}}$$  \hspace{1cm} (21)

which is derived from equating the cooling rate with the shock heating rate based on wave mechanics (Cossins et al. 2009; Rice et al. 2011). We see that this anti-correlation roughly holds in our results by comparing the top panel in Fig. 9 and the lower panel in Fig. 12, or as is explicitly shown in the inset of the lower panel of Fig. 12, where the density...
fluctuation is computed as

\[
\frac{\delta \Sigma}{\Sigma} = \frac{\int \sqrt{\left(\rho z\right)^2 - \left(\rho \Sigma\right)^2} \, dz}{\int \rho \, dz}.
\]  

(22)

As a consequence of the large density fluctuation, transient clumps in the gravito-turbulence have small \( Q_\star \) (see eq. 12). In our simulations, as \( r \) increases, \( \delta \Sigma/\Sigma \) increases and reaches values as large as \( \sim 2 \) at \( r = 75 \) AU (lower panel in Fig. 12) whilst \( Q_\star \) decreases and reaches values as small as \( \sim 0.2 \) (upper panel in Fig. 12). With such small \( Q_\star \), a transient clump can be easily driven to a bound object by acquiring a small amount of mass via collision with other clumps or by accretion of ambient gas. Such fragmentation process is actually seen in the gravito-turbulence at \( r = 75 \) AU as shown in Fig. 13. Note that the temperature of the clump does not decrease in the fragmentation process, which indicates that the fragmentation (i.e. reduction of \( Q_\star \)) is caused by an increase in the local surface density, rather than by a decrease in the temperature due to cooling.

Because the cooling time \( \beta_{\text{cool}} \) (or \( \alpha \)) does not depend on \( \Sigma \), as shown in Figs. 9 and 10, the cooling time (or \( \alpha \)) cannot play a role in determining another fragmentation boundary at \( \Sigma \sim \Sigma_{0.2} \) (\( \Sigma \) corresponding to the initial Toomre’s parameter \( Q_0 = 0.2 \)). A simple explanation for this boundary is that the initial disc is too unstable against GI (i.e. \( Q_0 \) is too small). This is because the initial temperature is set as the radiative equilibrium temperature (eq. 7), which does not depend on \( \Sigma \). Therefore, the larger \( \Sigma \) is, the smaller \( Q_0 \). However, the fragmentation boundary is not quite specific to that initial temperature. We have also evaluated cases in which the initial temperature is set so that \( Q_0 \) is kept at unity. The results did not change significantly because, as the simulation begins, the temperature quickly decreases to the radiative equilibrium temperature anyway owing to radiative cooling before GI develops. Therefore, when realistic radiative cooling is applied, a laminar disc of \( \Sigma \geq \Sigma_{0.2} \) would not evolve into the gravito-turbulence.

Figure 8. Same as Fig. 4, except for run R2.
Figure 9. Space-averaged cooling time \( \beta_{\text{ave}} \) (top) and Toomre’s parameter \( Q_{\text{ave}} \) (middle) versus \( \Sigma \), and Rosseland-mean opacity \( \langle x \rangle_{\text{mid}} \) versus gas temperature \( \langle T \rangle_{\text{mid}} \) (bottom). The inset in the top panel shows \( \beta_{\text{ave}} \) versus \( \Omega \) (the dotted line denotes \( \propto \Omega^3 \)). In the bottom panel, the dotted line denotes \( \propto \langle T \rangle_{\text{mid}} \), whilst the grey curve denotes the opacity model used in Clarke (2009). The different colours correspond to the different radii. The transition runs from gravito-turbulence to fragmentation are shown as open circles.

Figure 10. The \( \alpha \) parameter versus \( \Sigma \) (upper) and the accretion rate \( \dot{M} \) versus \( \Sigma \) (lower). In the lower panel, the small number at the right of a symbol denotes the fraction of the gravitational stress to the total stress. The inset in the upper panel shows \( \alpha \) versus \( \Omega^3 \) (the dotted line denotes \( \propto \Omega^3 \)), whilst that in the lower panel shows \( \dot{M} \) versus \( \Sigma \Omega^6 \) (the dotted line denotes \( \propto \Sigma \Omega^6 \)). The usage of colours and symbols is the same as in Fig. 9. In the bottom panel, the filled circles with a larger open circle are the solutions used in constructing a steady accretion disc model at \( \dot{M} = 7 \times 10^{-6} M_\odot \text{yr}^{-1} \) (denoted by the grey horizontal line) in Section 4.1.3.

3.5 Initial self-gravitating density waves

3.5.1 Stability of self-gravitating density waves

Here we examine again Fig. 2, where we compare the time evolution of \( Q_\star \) amongst the runs listed in Table 1. In every run, after the initial plateau, there is a period of monotonic decrease in \( Q_\star \), which corresponds to nonlinear growth of the axisymmetric density waves. As \( Q_\star \) decreases in time, the density waves become more strongly self-gravitating, and eventually become unstable when \( Q_\star \) decreases to a critical value \( Q_{\text{crit}} \). Note that the value of \( Q_{\text{crit}} \) depends on \( \Sigma \) and \( r \); the larger \( \Sigma \) or the larger \( r \) is, the smaller \( Q_{\text{crit}} \) becomes, which is also seen in the top rows of Figs. 14 and 15).

To understand the above stability of the initial self-gravitating density waves, we consult the linear analysis given in Takahashi et al. (2016). According to their analysis, a density wave (precisely, a two-dimensional density
Figure 11. Time variations of $\langle e \rangle$ of selected runs that sustain gravito-turbulence at different radii (top), cumulative fraction of the power spectral density (PSD) of $\langle e \rangle$ versus orbit for all runs that sustain gravito-turbulence (middle), and the orbit below which the cumulative fraction of PSD shown in the middle panel exceeds a critical value of 0.3 versus radius (bottom). The dotted line in the middle panel denotes the critical value. In the bottom panel, the small dots at each radii show $\beta_{\text{ave}}$. The usage of colours and symbols is the same as in Fig. 9.

Figure 12. Local Toomre’s parameter $Q_0$ (upper) and the fractional density fluctuation $\delta \Sigma / \Sigma$ (lower). The usage of colours and symbols is the same as in Fig. 9.

ring) of line mass $M_L$ and finite width $2W$ is unstable to non-axisymmetric perturbations when the Toomre’s parameter of the density wave satisfies the following condition:

$$\tilde{Q} = \frac{c_s(2\Omega)}{\pi G M_L / (2W)} < Q_{\text{crit}}(l),$$

where

$$l = \frac{(2W)}{c_s/(2\Omega)}$$

is the width of the density wave normalised by $c_s/(2\Omega)$. Here, the tilde symbol of $\tilde{Q}$ denotes that it is evaluated with $\kappa = 2\Omega$ assuming that the density wave is rigidly rotating. The instability condition (23) states that the critical value $Q_{\text{crit}}$ depends on its normalised width $l$. Specifically, as shown in Fig. 16, as $l$ decreases, $Q_{\text{crit}}$ decreases. This is because the long-range effect of self-gravity to drive the non-axisymmetric instability is reduced in narrower waves and thus more surface density is required for instability. The decrease of $Q_{\text{crit}}$ is notable when $l \lesssim 1$, that is, when the width ($2W$) becomes comparable to or less than the scale height ($\sim c_s/(2\Omega)$). On

The numerical data of $Q_{\text{crit}}(l)$ was kindly provided by Sane-michi Takahashi.
the other hand, the critical value \( Q_{\text{crit}}(l) \) becomes unity in the limit of infinite width \( (l \to \infty) \), which corresponds to the usual Toomre's condition, \( Q < 1 \).

To compare our simulation results with the linear stability condition, we evaluate the quantities \( \tilde{Q} \) (eq. 23) and \( l \) (eq. 24) for the density wave that contains the cell of the minimum self-gravitational energy, \((x_*, y_*, z_*) = 0\). Specifically, we assume that the density wave is parallel to the y-axis (see the top rows of Figs. 14 and 15) and approximate the local surface density distribution across the density wave as Gaussian, \( \sigma(x, y_*) \approx \sigma_{\text{max}} e^{-(x-x_*)/\Delta x^2/2} \). The width of a density wave of such Gaussian profile is somewhat arbitrary, but here we define \( W \equiv 1.5\Delta x \) (following Takahashi et al. 2016), where the line mass is computed as \( M_L \equiv \int_{y_*}^{y_*+W} \sigma(x, y_*) dx \sim 1.44\sigma_{\text{max}} W \).

### 3.5.2 Comparison with linear analysis

In Fig. 16, the time evolution of the evaluated quantities \((l, \tilde{Q})\) is represented as a trajectory in the \( l-\tilde{Q} \) plane for the five cases listed in Table 1. The density wave evolves from the upper right to the lower left by increasing its surface density \( (\propto 1/\tilde{Q}) \) as well as by decreasing its width \( (\propto l) \). Note that the left edge of each trajectory, which corresponds to the epoch when the density wave becomes unstable, is always found near the linear stability boundary, \( \tilde{Q} = Q_{\text{crit}}(l) \). This indicates that the behaviour of the initial density waves in our simulations is approximately explained by the linear theory. (Perfect agreement is not expected as, for example, a single isolated density wave is assumed in the linear theory whilst this is not the case in our simulations.)

Fig. 16 also shows that, as \( \Sigma \) increases (upper panel) or \( r \) increases (lower panel), the trajectory shifts leftward in the \( l-\tilde{Q} \) plane. Increasing \( r \) here also represents decreasing the cooling time, as shown in the bottom-right panels in Figs. 7.
4, and 8. (Specifically, $\beta_\star$ is $\sim 500$ at $r = 25$ AU, $\sim 30$ at $r = 50$ AU, and $\sim 6$ at $r = 90$ before the collapse.) Therefore, when $\Sigma$ is larger or the cooling time is shorter at larger radii, the density waves grow narrower and thus the critical Toomre’s parameter $Q_{\text{crit}}$ decreases. It is notable that regardless of changing $\Sigma$ or $r$, when the density waves become as narrow as $l \sim 1$, their collapse results in fragmentation. Therefore, the two fragmentation boundaries ($\Sigma \sim \Sigma_{0.2}$ and $r \sim 75$ AU) may be translated to the single condition, $l \sim 1$, in terms of the width of the initial density waves (c.f. Tsukamoto et al. 2014).

The dependence of the normalised width $l$ on $\Sigma$ or $r$ can also be observed in Fig. 17, where snapshots of density in the midplane $\rho(x, y, z = 0)$ as well as in the $x$-$z$ plane $\rho(x, y = y_\star, z)$, at the epoch when the density waves become unstable, are compared amongst the five runs. The normalised width $l = 2W/(c_\star/2\Omega)$ is regarded as the ratio of the width to the thickness of the density wave (except for a numerical factor of order unity). Then, we see clearly that as $\Sigma$ increases or $r$ increases, the density waves become narrower in terms of their thickness.

3.6 Transition from fragmentation to runaway collapse

In some cases at $r = 75$ and 90 AU, a transition from fragmentation to runaway collapse was observed. In Fig. 18, we compare the time evolution of $\Gamma_\star$ in three cases involving different outcomes at $r = 75$ AU, which are a fragmentation case ($\Sigma = 55$ g cm$^{-2}$), a transition case ($\Sigma = 70$ g cm$^{-2}$), and a runaway collapse case ($\Sigma = 80$ g cm$^{-2}$).

In the smallest $\Sigma$ case, a pressure-supported clump was formed and survived a few hundreds of orbits, where the core temperature $T_\star$ was maintained at the value that corresponds to $\Gamma_\star \sim 1.4$. In the intermediate $\Sigma$ case, a pressure-supported clump was also formed, but $T_\star$ was larger (and
thus $\Gamma_\ast$, was smaller) owing to stronger self-gravity. Furthermore, because the clump was not completely isolated, it accreted mass from the ambient medium and $T_\ast$ rose accordingly. Eventually, at $t \sim 73$ orbits, $T_\ast$ exceeded the $H_2$ dissociation temperature to cause $\Gamma_\ast < 4/3$, leading to runaway collapse. In the largest $\Sigma$ case, owing to the strongest self-gravity, $T_\ast$ quickly exceeded the $H_2$ dissociation temperature and thus $\Gamma_\ast$ became smaller than $4/3$ just after fragmentation occurred.

In summary, because a pressure-supported clump formed in fragmentation is not equilibrated with the ambient medium, it inevitably evolves by accretion or radiative cooling. Especially, when the core temperature of a formed clump is close to the $H_2$ dissociation temperature, runaway collapse may eventually occur as a result of such evolution.

4 DISCUSSION

4.1 Comparison with previous studies

4.1.1 Fragmentation boundaries in phase diagrams

In this section, we compare our phase diagram of the nonlinear outcome of GI (Fig. 1) with two previous studies, Johnson & Gammie (2003) and Clarke (2009).

More direct comparison is possible with Johnson & Gammie (2003), who performed 2D local shearing box simulations with a cooling function based on a one-zone model of optically thick disks. Unlike our simulations, a simple EOS was used and irradiation was not taken into account in their simulations. In Fig. 19, we plot our results on the $\Sigma-\Omega$ plane for direct comparison with their Fig. 7. The two solid curves drawn are taken from their Fig. 7, where the lower curve connects runs that show no signs of fragmentation whilst the upper curve connects those showing definite fragmentation. When compared over a common range of $\Omega$, ...
4.1.2 Stability of density waves and fragmentation

As we discussed in Section 3.5, the correlation between the width of the initial density waves and the critical Toomre’s parameter $Q_{\text{crit}}$ found in our simulations is mostly consistent with the linear stability analysis presented by Takahashi et al. (2016) (see Fig. 16). On the other hand, there is a notable difference between our fragmentation condition and theirs, which apparently comes from the difference between their global and our local simulations. They claimed that fragmentation occurs if and only if spiral density waves become unstable against the non-axisymmetric instability. Namely, the fragmentation condition is identical to the instability condition of the density waves (c.f. their Fig. 1). In contrast, in our simulations, the initial axisymmetric density waves always became unstable and collapsed. However, it is only when the density waves grow as narrow as $Q_{\text{crit}}(l = 1) \sim 0.4$ that the collapse results in fragmentation, which takes place either when $\Sigma > \Sigma_{0.2}$ or when $r > 75$ AU.

Our conclusion that fragmentation occurred when the density waves became narrow enough appears to be similar to the fragmentation condition determined by the Hill radius proposed by Rogers & Wadsley (2012). Their fragmentation condition can be rewritten as (c.f. Takahashi et al. 2016),

$$Q < Q_{\text{crit}}(l) \equiv \frac{32}{3\pi} \frac{l}{r}. \quad (25)$$

which is also plotted in Fig. 16. As shown, the initial axisymmetric self-gravitating density waves in our simulations kept growing without fragmentation even after they entered the above unstable region (eq. 25). Therefore, our simulation results may not be explained in terms of the Hill radius as proposed by Rogers & Wadsley (2012).

4.1.3 Steady accretion driven by gravito-turbulence

Next we examine the gravito-turbulence in thermal equilibrium. Fig. 10 shows that steady accretion driven solely by gravito-turbulence is possible for a range of radii, which depends on the mass accretion rate (c.f. Rice & Armitage 2009). For example, mass accretion of $M = 7 \times 10^{-3} M_{\odot} \text{yr}^{-1}$ would be possible for $20 \lesssim r \lesssim 75$ AU. To construct a steady accretion disc for that accretion rate, we pick up a solution of $M \sim 7 \times 10^{-3} M_{\odot} \text{yr}^{-1}$ at each radius between 20 and 75 AU and combine them. The radial profiles of $\Sigma$, $T_{\text{mid}}$, and $\dot{M}$ are shown in Fig. 10.
Figure 17. Snapshots of density at the epoch when the axisymmetric density waves become unstable for the runs shown in Fig. 14 (upstairs), and for runs shown in Fig. 15 (downstairs). In each pair of panels, the upper one shows $\rho(x, y, z = 0)/\rho_0$ whilst the lower one shows $\rho(x, y, z)/\rho_0$. The cell of the minimum $E_{\text{eg}}$ is shown as a white cross in the snapshots.

Figure 18. Same as Fig. 3, except it compares cases for three different values of $\Sigma$ at $r = 75$ AU: $\Sigma = 55$ (grey), 70 (dark grey), and 80 g cm$^{-2}$ (black).

Figure 19. Comparison with Johnson & Gammie (2003). The nonlinear outcomes of GI shown in Fig. 1 are replotted in the same frame as their Fig. 7 (the vertical and horizontal axes are exchanged). The two black curves are their “critical curves”.

Johnson & Gammie (2003)
\[ \alpha \] for such a steady accretion disc based on our simulations are plotted in Fig. 21.

These profiles are directly compared with those of the analytical model by Clarke (2009) because they assume a central star of 1\( M_\odot \) as we did. In Fig. 21, their profiles at \( M = 7 \times 10^{-6} \, M_\odot \, \text{yr}^{-1} \) are also plotted. Among the three quantities, it is notable that the midplane temperatures are consistently higher in Clarke (2009), indicating that radiative cooling is less effective in their case. This may be because they adopted the midplane opacity for cooler upper layers, which would overestimate the optical thickness of the disk because \( \kappa \propto T^2 \) (below the ice melt temperature). If this is the case, they would also overestimate the cooling time \( \beta \), which leads to underestimating the \( \alpha \) value (\( \propto 1/\beta \) in thermal equilibrium) and then to overestimating \( \Sigma \). These naturally explain the discrepancies between Clarke (2009)'s model and ours seen in Fig. 21, although both models use a similar opacity (see the bottom panel in Fig. 9). Therefore, resolving the vertical structure with appropriate radiative transport is essential in determining the radial profile of the disc.

### 4.2 Formation of bound clumps via gravitational instability

Using 3D global disc simulations adopting the \( \beta \) cooling, Boss (2017) showed that low \( Q_0 \) discs can fragment for high \( \beta \) whilst high \( Q_0 \) discs can be stable for small \( \beta \), which indicates the equal importance of the initial Toomre’s parameter to the cooling time for fragmentation. Our simulations qualitatively agree with their results regarding the importance of the initial Toomre’s parameter. Namely, Fig. 1 shows that fragmentation is possible at any radius, or at any cooling time, provided that the surface density is as large as \( \Sigma_{\text{crit}} \sim \Sigma_{0.2} \), or the initial Toomre’s parameter is as small as 0.2. On the other hand, beyond \( r \sim 90 \, \text{AU} \), the critical surface density is relaxed to \( \Sigma_{\text{crit}} \sim \Sigma_1 \). That is, when the cooling time is as short as \( \beta_{\text{ave}} \lesssim 1 \), fragmentation always occurs at any value of \( Q_0 \) less than unity.

Using the value of \( \Sigma_{\text{crit}} \), we make a crude estimate of the mass of a bound clump formed in the collapse of the
initial density waves as per (c.f. Rafikov 2005)

\[ M_{\text{clump}} \sim (fH)^2 \Sigma_{\text{crit}} \sim \left( \frac{f_{20}^2}{\Omega} \right) \left( \frac{\epsilon_{\text{damp}} Q_{0}}{\pi G Q_{0}} \right) \]

\[ \frac{M_{\text{clump}}}{M_3} \sim \left\{ \begin{array}{ll}
\frac{5}{3} & \left( \frac{Q_{0}}{2} \right)^{-1} \left( \frac{r}{25 \text{ AU}} \right)^{\frac{3}{2}} \quad (r \leq 75 \text{ AU}), \\
\frac{3}{5} & \left( \frac{Q_{0}}{1} \right)^{-1} \left( \frac{r}{90 \text{ AU}} \right)^{\frac{3}{2}} \quad (r \geq 90 \text{ AU}).
\end{array} \right. \]

which can be written in terms of the Jupiter mass \( M_3 = 2 \times 10^{30} \) g as

The numerical factor \( f \) here stands for the size of such a clump in terms of the scale height \( H \), for which we employ a value of \( \sim 6 \) based on our simulations. The above equation indicates that the minimum mass of a bound clump formed in the non-axisymmetric instability is several to ten times \( M_3 \) for the radii we have explored.

So far as we have investigated, a pressure-supported clump once formed was never dissolved by the velocity shear, either surviving or being followed by runaway collapse. This indicates that the realistic cooling is so efficient that a formed clump remains compact enough to resist velocity shear.

### 4.3 Dependence on the box size and the spacial resolution for fragmentation cases

In our simulations shown above, the box size and the spacial resolution were fixed as, respectively, \((L_x, L_y, L_z) = (24H, 24H, 12H)\) and \((N_x, N_y, N_z) = (128, 128, 64)\). They are the same as those used in the fiducial run in Paper I, where we showed that the results do not strongly depend on them when gravito-turbulence is established. Here we discuss how the results could depend on the box size or the spacial resolution when fragmentation occurs (i.e. \( Q < 1 \) in the final state).

Firstly we examine the box size dependence. Fig. 22 compares the time evolution of \( Q_x \) in the case of \( \Sigma = 300 \) g cm\(^{-2} \) at \( r = 50 \) AU (run S2) as well as that in the same case but with a halved box size, i.e. \((L_x, L_y, L_z) = (12H, 12H, 6H)\).

In the case of the standard box size, \( Q_x \) decreased below the critical value of \( \sim 0.2 \) and fragmentation occurred, which was followed by runaway collapse at \( r \sim 2.3 \) orbits. On the other hand, in the case of the halved box, although fragmentation occurred similarly, it was not followed by runaway collapse, and a pressure-supported clump survived instead. This means that mass concentration by self-gravity in the halved box, which contained one fourth the amount of mass contained in the standard box, was not enough to raise the core temperature of the clump above the \( H_2 \) dissociation temperature. Therefore, although we may always expect fragmentation beyond some critical \( \Sigma \) at a given radius, whether runaway collapse follows the fragmentation depends on how much mass concentrates by self-gravity, which then may depend on the box size.

Next we examine the spacial resolution. In Fig. 23, we plot the time evolution of the adiabatic exponent \( \Gamma_3 \) for two cases; one is the case of \( \Sigma = 30 \) g cm\(^{-2} \) at \( r = 90 \) AU (run R2; green) and the other is \( \Sigma = 60 \) g cm\(^{-2} \) at \( r = 75 \) AU (purple).

In the former case, a pressure-supported clump was formed and survived for many orbits, where \( \Gamma_3 \sim 1.42 \) (thick green curve). In the latter case, a pressure-supported clump was formed similarly, but \( \Gamma_3 \) was closer to the critical value of \( 4/3 \).

To observe the dependence of the spacial resolution for these two cases, we doubled the number of cells, i.e. \((N_x, N_y, N_z) = (256, 256, 128)\), and restarted the calculation from a snapshot of the standard resolution run. The restarting time was set after a pressure-supported clump was formed. In the former case, the result did not change significantly although \( \Gamma_3 \) in the high-resolution run (thin green curve) was slightly lower, probably because mass concentration was resolved better. On the other hand, in the latter case, the result was changed drastically by doubling the spacial resolution. As shown by the thin purple curve, \( \Gamma_3 \) quickly decreased below the critical value of \( 4/3 \) and runaway collapse occurred. This is because mass concentration enhanced by the doubled resolution was large enough to raise the core temperature above the \( H_2 \) dissociation temperature.

In summary, it is difficult to determine a precise condition for runaway collapse using local shearing box simulations with a fixed spacial resolution because whether runaway collapse occurs does depend on the box size and the spacial resolution. Global disk simulations are needed to determine the amount of mass involved in fragmentation, and a sort of mesh refinement is required to follow mass concentration by self-gravity at smaller scales. On the other hand, the fragmentation condition itself should be obtained by local shearing box simulations if the critical wavelength of GI is contained in the box and is resolved appropriately.

### 5 SUMMARY

Using local three dimensional radiation hydrodynamics simulations, the nonlinear outcome of gravitational instability in an irradiated protoplanetary disc is investigated in a parameter space of the surface density \( \Sigma \) and the radius \( r \).

Starting from laminar flow, axisymmetric self-gravitating density waves grow first. Their degree of self-gravitating becomes larger when \( \Sigma \) is larger or the cooling time is shorter at larger radii. The density waves eventually
We have incorporated into our 3D simulations a realistic EOS, realistic radiative transfer (in the framework of FLD), and consider irradiation heating. These are relevant physics aspects for correct thermodynamic analysis related to protoplanetary discs. Actually, in Section 4.1.3, we showed that resolving the vertical structure with appropriate radiative transport is essential in determining the radial structure of the disc. However, there remain some limitations in our methods, and we add caveats here. Firstly, since we are using the local shearing box approximation, so our results should be valid in the case where the global transport of energy is not important, as discussed in detail in Paper I. Also, as we discussed in Section 4.3, the problem of whether runaway collapse occurs after fragmentation remains subtle, as the mass concentration at the centre of a formed clump is not properly solved in our simulation box with fixed size and resolution. Finally, we note that our study in this paper is dedicated to a particular protoplanetary disc system. Therefore, the fragmentation boundaries presented here may change if, for example, the central star’s irradiation or the dust opacity is changed.

ACKNOWLEDGEMENTS

We thank Sanemichi Takahashi, Shu-ichiro Inutsuka and Yusuke Tsukamoto for valuable discussions, and Kengo Toida for providing us with his EOS tables for our simulations. We also thank the anonymous referee for his/her valuable comments for improving the manuscript. Numerical calculations were carried out partly on Cray XC30 at CICA, National Astronomical Observatory of Japan, and partly on Cray XC40 at YITP in Kyoto University. SH was supported by Japan JSPS KAKENH 15K05040/18K03716 and the joint research project of ILE, Osaka University. JS was supported in part by the National Science Foundation under grant PHY-1144374, "A Max-Planck/Princeton Research Center for Plasma Physics" and grant PHY-0821899, "Center for Magnetic Self-Organization".

REFERENCES

Andrews S. M., Rosenfeld K. A., Kraus A. L., Wilner D. J., 2013, The Astrophysical Journal, 771, 129
Baehrr H., Klahr H., 2015, The Astrophysical Journal, 814, 155
Boss A. P., 1997, Science, 276, 1836
Boss A. P., 1998, The Astrophysical Journal, 503, 923
Boss A. P., 2017, The Astrophysical Journal, 836, 53
Chiang E. I., Goldreich P., 1997, The Astrophysical Journal, 490, 368
Clarke C. J., 2009, Monthly Notices of the Royal Astronomical Society, 396, 1066
Cossins P., Lodato G., Clarke C. J., 2009, Monthly Notices of the Royal Astronomical Society, 393, 1157
Cossins P., Lodato G., Testi L., 2010, Monthly Notices of the Royal Astronomical Society, 407, 181
Durisen R. H., Boss A. P., Mayer L., Nelson A. F., Quinn T., Rice W. K. M., 2007, Protostars and Planets V, pp 607–622
Ferguson J. W., Alexander D. R., Allard F., Barman T., Bodnarik J. G., Hauschildt P. H., Heffner-Wong A., Tamanai A., 2005, "Center for Magnetic Self-Organization".

Figure 23. Same as Fig. 3, except for run R2 (green) and $\Sigma_0 = 60$ g cm$^{-2}$ at $r = 75$ AU (purple). The dotted curves correspond to double-resolution versions of the two cases, restarted from $t = 9$ (purple) and $t = 20$ (green) orbits, respectively.
Nonlinear outcome of gravitational instability

APPENDIX A: LIST OF ALL RUNS

Fig. A1 shows the running period of all 74 runs with the values of the two parameters, \( r \) and \( \Sigma \). The runs with a non-standard box size or resolution discussed in Section 4.3 are excluded here. The running period is colour-coded with green, red, grey, and black for no GI, gravito-turbulence, fragmentation, and runaway collapse, respectively. As stated in the Section 3.2, gravito-turbulence, fragmentation, and runaway collapse are distinguished by the value of \( Q^* \). “No GI” is identified as a state where the fractional density fluctuation \( \delta \Sigma/\Sigma \) defined in equation (22) is less than \( 10^{-2} \); such condition is adopted because the fluctuation \( \delta \Sigma/\Sigma \) is not exactly zero even when GI does not occur, because the initial flow is not in an exact hydrostatic and thermal equilibrium.

This paper has been typeset from a TeX/LaTeX file prepared by the author.
Figure A1. Running period for all runs. The two parameters, $r$ and $\Sigma$, of each run are shown on the left. (Job IDs are shown in the left-most column.) For runs where runaway collapse occurs, a cross is shown on the right of the period.