1. Introduction

Planet formation in circumstellar disks relies largely on the core accretion to explain most of the planets observed to date. The slow buildup of solids to form planetesimals and planetary cores is sufficient to explain most discovered planets within ~10 au of their host star (Safronov 1969; Lissauer 1993; Mordasini et al. 2012). The formation timescale of core accretion increases with distance from the star, making the formation of gas giant planets at wide radii more difficult. Fragmentation through gravitational instability (GI) is a model that forms stellar companions quickly and typically only works in the cool outer reaches of the circumstellar disk (Kuiper 1951; Boss 1997). When GI fails to produce a companion star, it may result in the formation of bound companions with masses on the order of several Jupiter masses.

Gravitational instability becomes relevant when a circumstellar disk has enough gas mass to form local overdensities that may collapse under their own self-gravity. Collapse only proceeds when the disk cools efficiently enough to suppress thermal pressure support (Gammie 2001; Durisen et al. 2007). These two conditions are measured by the Toomre parameter $Q$ and a simple cooling timescale $t_c$, respectively. The Toomre parameter quantifies the balance of stabilizing effects of rotation and pressure versus the destabilization generated by gas overdensities and is defined in a linear 2D approximation by Toomre (1964) as

$$Q = \frac{c_s \Omega}{\pi G \Sigma},$$  \hspace{1cm} (1)  

where $\Omega = \sqrt{GM/\Sigma^3}$ is the orbital Keplerian frequency of the disk, stabilizing the disk to density perturbations at larger wavelengths, and $c_s$ is the local speed of sound, which stabilizes the disk to density perturbations at smaller wavelengths (Durisen et al. 2007). Theoretical models of protoplanetary disks often assume a thin disk, so quantities are vertically integrated, i.e., surface density $\Sigma = \rho H$, where $H$ is the scale height of the disk. A critical balance between stability and instability is defined for $Q = 1$. Areas with $Q \lesssim 1$ are unstable to axisymmetric density perturbations in the linear regime, which may contract and fragment following the nonlinear growth of the perturbations, while a region with $Q \gtrsim 1$ will remain stable to collapse (Toomre 1964). When marginally stable, a disk tends to stay close to $Q \gtrsim 1$. Falling below $Q = 1$ generates strong shock heating as structures begin gravitationally contracting, increasing temperatures (Balbus & Papaloizou 1999). The balance of shock heating and cooling results in a gravitoturbulent self-regulating disk, subject to instabilities, but unable to collapse into fragments without an additional constraint.

With a disk balancing on the edge of instability, the cooling timescale is second requirement for fragmentation. The cooling timescale as used here defines the rate at which the disk loses all of its internal energy. Thus, if the disk can cool efficiently, comparable to the dynamical timescale $\Omega^{-1}$, thermal pressure support is removed quickly enough to allow for collapse by self-gravity. When cooling is too slow, thermal saturation will provide pressure strong enough to push back against collapse. Therefore, using a simple prescription for the cooling timescale

$$t_c = \beta \Omega^{-1},$$  \hspace{1cm} (2)
where $\beta$ is constant of around unity for efficient cooling and leads to fragmentation with a critical boundary at $\beta_{\text{crit}} = 3$ (Gammie 2001). This critical cooling threshold between the gravitoturbulent state and fragmentation is dependent on adiabatic index $\gamma$ which will push the critical $\beta$ value up slightly for lower values of $\gamma$ (Rice et al. 2005).

Both SPH and grid simulations have failed to converge to a critical cooling parameter (Lodato & Clarke 2011; Meru & Bate 2011; Paardekooper 2012). Potential resolutions to this issue have varied from changes to the implementation of viscosity (Rice et al. 2014; Klee et al. 2017), to varied cooling methods (Rice et al. 2012; Baehr & Klahr 2015), but one likely reason for the problem lies in the smoothing of the self-gravity potential (Young & Clarke 2015). Without smoothing self-gravity to approximately the scale height of the disk (Hure & Pierens 2009; Muller et al. 2012), self-gravity is exaggerated in 2D, the effect being more pronounced with increasing resolution, resulting in easier runaway collapse with each increase in resolution. Thus, 3D models are needed to clarify the critical cooling boundary for disk fragmentation including what value the boundary takes and whether or not it converges in 3D with a simple cooling law including irradiation. It is still prudent, however, to monitor the effect of various dissipation strengths on the behavior of fragmentation.

Global 3D studies of disk fragmentation are common (Mej{\`a} et al. 2005; Boley et al. 2006; Mayer et al. 2007; Michael et al. 2012; Vorobyov 2013; Lichtenberg & Schleicher 2015), but because one simulates the entire disk, it is difficult to resolve the small scales of clump/fragment formation. Using local 3D high-resolution simulations, we can instead focus on a small area of the disk and its interaction with embedded objects. Recently, an increasing number of local 3D simulations have been undertaken, but do not focus on the formation of fragments (Hirose & Shi 2017; Riols et al. 2017). Their focus has been on properties of the gravitoturbulent state, such as the nature of turbulence and non-axisymmetric structure (Mamatsashvili & Rice 2010; Shi & Chiang 2014) and dust concentration (Shi & Chiang 2013; Shi & Chiang 2014) and Hirose & Shi (2017) came the closest to establishing a link to the critical cooling threshold of Gammie (2001) using 3D local simulations, but stopped short of modeling clump collapse or conducting simulations of the fragmentation threshold over multiple resolutions. Studies of similar 3D physics at galactic scales were conducted by Kim et al. (2002), which also found a critical Toomre parameter $Q \approx 0.7$ but did not include thermodynamics or investigate the effect of resolution on the fragmentation behavior.

In this paper, we investigate the fragmentation boundary in 3D local disks over several resolutions to study the convergence behavior of self-gravitating disks. In light of recent unsuccessful (Baehr & Klahr 2015) and successful (Young & Clarke 2015; Deng et al. 2017) attempts at convergence in 2D self-gravitating disks, we seek to determine if 3D modeling of fragmentation with simple cooling including a fixed floor temperature is closer to a realistic treatment the self-gravity of the problem and results in convergence. This paper will proceed with a brief overview of the relevant hydrodynamic equations in Section 2 followed by a short look at the numerical setup in Section 3. Finally, we look at the results of our vertically stratified simulations in Section 4 and discuss the implications and consequences on disk fragmentation and the simulation thereof in Section 5.

2. Theory

We consider a non-magnetized, self-gravitating disk in the shearing box approximation. Accordingly, the disk is modeled locally on a small radial-azimuthal patch of the disk, transformed from global cylindrical coordinates ($r$, $\theta$, $\phi$) to local Cartesian coordinates ($x$, $y$, $z$) co-rotating with the disk (Goldreich & Lynden-Bell 1965). These assumptions allow for the modeling of the local properties of the disk while following the evolution of fragments that form when using periodic boundary conditions. Using this model, the relevant conservation equations are shown in Equations (3)–(5) with terms for the Coriolis effect $\Omega \times \mathbf{u}$ and centrifugal force $q\Omega^{2} \gamma$ as well as heating $H$ and cooling terms $C$ in the energy equation.

$$\frac{\partial \rho}{\partial t} - q\Omega x \frac{\partial \rho}{\partial y} + \nabla \cdot (\rho \mathbf{u}) = f_{\rho}(\rho)$$

$$\frac{\partial \mathbf{u}}{\partial t} - q\Omega x \frac{\partial \mathbf{u}}{\partial y} + \mathbf{u} \cdot \nabla \mathbf{u} = -\frac{\nabla P}{\rho} + q\Omega \gamma y - 2\Omega \times \mathbf{u} - \nabla \Phi + f_{u}(u)$$

$$\frac{\partial s}{\partial t} - q\Omega x \frac{\partial s}{\partial y} + (u \cdot \nabla)s = \frac{1}{\rho T}(2\rho u S - \Lambda + f_{s}(s))$$

Here, $\mathbf{u} = (u_{x}, u_{y}, u_{z})$ is the flow plus shear velocity in the local box, $\rho$ is the gas density, and $s$ is the gas entropy. Viscous heat is generated by $H \approx 2\rho u S$, with rate-of-strain tensor $S$, and radiated away by a simple $\beta$-cooling prescription $\Lambda$, described in Section 2.2.

Self-gravity is modeled by solving the Poisson equation

$$\nabla^{2} \Phi = 4\pi G \rho, \quad (6)$$

in Fourier space by transforming the surface density to find the potential at the scale of wavenumber $k$ and transforming the solution back into real space. The solution to the Poisson equation in Fourier space is

$$\Phi(k, t) = -\frac{2\pi G \rho(k, t)}{k^{2}}. \quad (7)$$

In 2D simulations, it is useful to smooth the self-gravity potential at small scales. When using a Fourier method, this means limiting the wavenumbers included in the calculation of the potential to exclude those corresponding to gravitational attraction at scales smaller than a pressure scale height. This is used in 2D to ascribe a thickness to the disk; however, as our simulations include a vertically stratified density distribution, there should be no need for smoothing in these simulations.

Finally, we use an ideal equation of state, with internal energy $\epsilon$, and specific heat ratio $\gamma$:

$$P = (\gamma - 1)\epsilon \rho. \quad (8)$$

We use a ratio of $\gamma = 5/3$, which corresponds to a 2D value of $\gamma = 1.8$ (Gammie 2001).

2.1. Viscous Stress

Accretion driven by turbulence during the early self-gravitating phase of a disk is expected to be strong, and according to the $\alpha$-formalism (Shakura & Sunyaev 1973), typically approaches unity $\alpha \approx 0.1−1$, a few orders of magnitude higher than that of magnetorotational instability (Balbus & Hawley 1991). To measure the $\alpha$-stress in our
simulations, we will need to consider the contribution of the gravitational and Reynolds stresses:

\[ \alpha = -\left( \frac{d \ln \Omega}{d \ln r} \right) \left( \frac{\langle G_{xy} \rangle + \langle R_{xy} \rangle}{\rho \nu_c^2} \right), \]  

(9)

where

\[ \langle G_{xy} \rangle = \int_{-\infty}^{\infty} \frac{\delta \rho \delta u_x}{4 \pi G} dz, \]  

(10)

and

\[ \langle R_{xy} \rangle = \langle \rho u_x u_y \rangle. \]  

(11)

An analytic expectation of \( \alpha \) in a gravitoturbulent thin disk is given by Gammie (2001):

\[ \alpha = \frac{4}{9} \frac{1}{\gamma (\gamma - 1) \tau_c \Omega}. \]  

(12)

Shi & Chiang (2014) compare the standard formulation of \( \alpha \) (Equation (9)) with one based on density-weighted values of the pressure and stress, and find that while the density-weighted \( \alpha' \) is a factor of 1.25–2 higher, the usual formulation more closely follows the theoretical expectation (12). From Equations (9) and (12), one can compare theoretical expectations of viscous stresses in the gravitoturbulent state with the stresses generated by our simulations. It is important to note that as a local description of viscous stress, \( \alpha \) does not necessarily model gravitationally unstable disks, as gravity inherently acts on non-local scales (Balbus & Papaloizou 1999).

### 2.2. Thermodynamics

Proper modeling of disk cooling is complex even without using sophisticated ray-tracing and flux-limited diffusion schemes, depending significantly on the local temperature and density of the disk from midplane to atmosphere (Hubeny 1990; Rafikov 2005):

\[ \tau_c = \frac{4}{9} \frac{\Sigma c^2_s}{\sigma (T^4 - T_{int}^4)} f(\tau), \]  

(13)

where \( \sigma \) is the Stefan–Boltzmann constant, \( f(\tau) = \tau + 1/\tau \) for optical depth \( \tau \), \( T_{int} \) is the reference temperature and adiabatic index \( \gamma \).

Contraction to a fragment requires that thermal pressure support does not counter the self-gravity of overdense perturbations. Heat generated through viscous heating and shock dissipation should be released on a cooling timescale short enough to keep pressure support from building up, or shorter than a few dynamical timescales (see Equation (2)). Cooling with a \( \beta \)-prescription is implemented in Pencil using a Newtonian cooling scheme,

\[ \Lambda = \frac{\rho (\rho |c_s^2 - c_{int}^2|)}{(\gamma - 1) \tau_c}, \]  

(14)

where \( c_{int}^2 \) is proportional to a non-zero background temperature, corresponding to some constant irradiation, which as demonstrated by Lin & Kratter (2016) adds additional stability against small scale density perturbations, and \( \tau_c \) is a \( \beta \)-cooling timescale (2). The sound speed relates to entropy as

\[ c_s^2 = \frac{P}{\rho} = c_{s,0}^2 \exp(\gamma s/c_p + (\gamma - 1) \ln (\rho/\rho_0)), \]  

(15)

where \( c_{s,0} \) is the initial sound speed \( c_{s,0} = \pi, \) density \( \rho_0 = 0.293 \), and temperatures are derived from the sound speed using \( T = c_s^2/c_p (\gamma - 1) \). This is a simple way to model disk cooling which is not as computationally expensive as more realistic methods, particularly at the resolutions used here.

The cooling described in Equation (14) is balanced by two sources of heat generation: viscous heating and shock dissipation. Viscous heating is calculated as

\[ \mathcal{H} = 2\rho \nu S^2, \]  

(16)

with kinematic viscosity \( \nu = 2.5 H^2 \Omega \) for all simulations, as seen in Equation (5), but ultimately the heating of the simulation will be dominated by the dissipation of shocks produced by the collapse of self-gravitating filaments and clumps. The heat generated by the dissipation of shocks is determined by the strength of the constant of hyperdissipation, so in addition to our standard collection of simulations spanning resolution and cooling timescale, we include a small sample of simulations with various dissipation values to note the effect on fragmentation.

In practice, this may be overestimating the heating resulting from strong shocks generated by self-gravity. It is well understood that in outer disk regions, disk heating is dominated by stellar irradiation instead of viscous accretion (D’Alessio et al. 1999). Additionally, recent work by Rafikov (2016) suggests that shocks are not a dominant source of heat in passive disks such as those considered here.

### 2.3. Vertical Stratification

A proper implementation of a 3D shearing box simulation requires vertical stratification of the gas density. If this is not correctly initialized, the simulation will adjust on its own, causing undesired rapid heating and cooling as the disk swells up or contracts. Thus, we assume the disk begins in hydrostatic equilibrium with self-gravity

\[ \frac{1}{\rho} \frac{d P}{dz} = -\frac{\Omega^2}{4\pi G} \int_0^z \rho(z')dz'. \]  

(17)

The initial vertical density distribution is determined by rewriting the above as a dimensionless, differential equation, following Shi & Chiang (2014):

\[ \frac{d^2 \tilde{\rho}}{d \tilde{z}^2} + \frac{\gamma - 2}{\tilde{\rho}} \left( \frac{d \tilde{\rho}}{d \tilde{z}} \right)^2 + Q_0 \tilde{\rho}^{2-\gamma} + 2 \tilde{h}^{3-\gamma} = 0. \]  

(18)

We use an isothermal approximation to the polytropic equation of state for which Equation (18) was derived (\( \gamma = 1, K = c_s^2 \)), and assume \( \Sigma_0 = 2H \rho \). Solving for vertical hydrostatic equilibrium including self-gravity we find

\[ \frac{d^2 \tilde{\rho}}{d \tilde{z}^2} - \frac{1}{\tilde{h}} \left( \frac{d \tilde{\rho}}{d \tilde{z}} \right)^2 + Q_0 \tilde{\rho} + \frac{2}{h} \tilde{h} \rho^2 = 0, \]  

(19)

where \( \tilde{\rho} = \rho/\rho_0, \tilde{z} = z/(c_{s,0}^2/\pi G \Sigma_0), Q_0 = c_{s,0}^2/\pi G \Sigma_0, \) and \( h = H/(c_{s,0}^2/\pi G \Sigma_0) \). The approximation of the vertically stratified initial density creates some small oscillations in as
the simulation equilibrates to the true solution. This results in a small flux of mass away from the regions between $1 < |z| < 2$ and toward the outer disk atmosphere, but these result in density deviations on the order of $10^{-4}$ and do not have a significant impact on the outcome of the simulation. Due to the large empty regions above and below the disk midplane, it is necessary to impose a minimum density value to avoid extreme density differences which would slow down the code immensely.

The initial temperature distribution is not stratified but instead isothermal throughout the entire vertical extent, equal to the background irradiation level. Since distant regions of the disk will be dominated by disk irradiation rather than accretion heating this is a reasonable assumption. As the disk settles to equilibrium, shocks are generated from self-gravitational collapse and generate significant heat, but cooling returns the vertical temperature profile to its original flat distribution.

2.4. Comparison with the Thin Disk Approximation

Introducing vertical structure affects the stability and dynamics of a simulation, and should be accounted for in comparison to a thin disk approximation. Mass density should be stratified so that it is as close to hydrostatic equilibrium as possible; otherwise the simulation will adjust itself and if the difference is significant, the simulation will be adversely affected. In a simulation on the border between stratification and stability, too much heating initial heating may suppress fragmentation and too much cooling may result in premature fragmentation.

The stratification of mass density also results in the weakening of self-gravity, meaning a simulation established with $Q_0 = 1$ will not become unstable. In the thin disk approximation, the gas density is uniformly distributed over two scale heights, but in our 3D simulations, it is instead concentrated in the midplane but distributed over twelve vertical scale heights. With more mass further away from the midplane, a disk needs to be more massive to become unstable. Theoretical expectations indicate a critical Toomre criterion closer to $Q_0 \approx 0.7$ (Behrendt et al. 2015; Kratter & Lodato 2016) and our simulations reflect that.

3. Numerical Methods

All simulations were conducted using the open-source Pencil code\(^6\) (Brandenburg 2001) and no additional modifications were made in the implementation of these runs.

3.1. Artificial Viscosity

An important question regarding the stability of self-gravity is the role of artificial viscosity on the onset of fragmentation. We use a hyperdiffusion method, implemented in Pencil according to Appendix B of Yang & Krumholz (2012):

$$ f_D(\Sigma) = \zeta_D(\nabla^6 \Sigma), $$

which smooths perturbations on small scales without affecting the power on larger scales. It has been suggested that such viscosities are the reason for the non-convergence of the 2D simulations (Rice et al. 2014). The choice of a stronger dissipation scheme, such as hyperviscosity, is justified because of the effects of numerical oversteepening (Klee et al. 2017), which may lead to the overestimation of overdense perturbations and contribute to non-convergent behavior.

3.2. Boundary Conditions

The shearing box is set up with periodic $y$ and $z$ boundaries, corresponding to the azimuthal and radial physical extents, respectively. This means that all quantities that exit one boundary are reproduced at the other side of the box, including the self-gravitational potential in the vertical direction.

$$ f(x, y, z, t) = f(x, y + L, z, t) $$

The boundaries $x = 0$ and $x = L$ require a different boundary condition on account of the shear velocity $u_y = v_y + \frac{3}{2} \Omega x$ (Hawley et al. 1995).

$$ f(x, y, z, t) = f(x + L, y - \frac{3}{2} \Omega L z, t) $$

When calculating this over a shear periodic $x$-boundary, the displacement due to the shear is taken into account by shifting the entire $y$-direction to make the $x$-direction periodic before proceeding with the transform in the $z$-direction. After the calculation of the potential in Fourier space the process is reversed to get back to real space.

3.3. Initial Conditions

Initial values of density and sound speed are set by establishing an overall Toomre $Q$ (Equation (1)) that is borderline unstable. To this end, we use $\Omega = G = 1$ and must consider how we calculate the sound speed and surface density. By setting the sound speed to $c_{s,0} = \pi$, we are left with the surface density $\Sigma$, which while normally one would ensure $\Sigma_0 = \int \rho_0(z) dz = 1$ to achieve $Q_0 = 1$, we find it necessary to also consider the case where $Q_0 \approx 0.676$ and initialize the surface density accordingly. To account for the gravitational field of the central star, we include a sinusoidal acceleration profile, such that mass settles toward the midplane but acceleration approaches 0 at the vertical boundaries. Such a profile is not as realistic as a linear profile, but it improves stability.

A reference temperature proportional to $c^2_{s,0} = c^2_{s,0} = \pi^2$ is used in Equation (14) because the sound speed is initialized as $c_{s,0} = \pi$. The shearing box needs to be large enough to contain a critical Toomre wavelength, $\lambda = 2\pi/k_{\lambda,0} = 2\pi H$, while also resolving this critical wavelength with at least four grid cells and avoid artificial fragmentation (Nelson 2006). We meet these conditions using $L_x = L_y = L_z = (40/\pi) H \approx 12.7 H$; see Figure 1 for an example of the box scale. These boxes are smaller in radial and azimuthal extent than other comparable simulations (Hirose & Shi 2017; Riols et al. 2017), allowing for higher resolution per scale height comparatively.

3.4. Resolution Requirements

Artificial fragmentation due to poor resolution is a significant concern with self-gravitating simulations, as insufficient resolution can lead to non-negligible errors which form spurious fragments (Truelove et al. 1997). Nelson (2006)
found that for gravitationally unstable disk simulations, the critical Toomre wavelength needs to be resolved by at least four resolution elements to avoid such fragmentation, described by a maximum Toomre number of

\[
T = \frac{\delta x}{\lambda_T} = 0.25. \tag{24}
\]

Our simulations contain almost exactly one critical Toomre wavelength and with a resolution of \(N^3 = 256^3\) cells in each direction, our minimum resolution easily meets this resolution criterion with \(T = 0.008\), decreasing by a factor of one-half when increasing the resolution to \(N^3 = 512^3\) and \(N^3 = 1024^3\).

The simulations of Boss (2017) only consider fragments at or near this resolution criterion, and are thus potentially susceptible to fragmentation due to numerical errors. Additionally, the Toomre length only roughly describes the feeding zone of an unstable clump which may or may not fragment further (Kratter & Murray-Clay 2011), and is not necessarily a length indicative of a fragment.

### 4. Results

As mentioned above, preliminary simulations of our 3D self-gravitating setup were conducted at our lowest resolution of \(N^3 = 256^3\), which is the same effective resolution as the 2D simulations with \(N^2 = 1024^2\) of Baehr & Klahr (2015). When initialized with \(Q_0 = 1\) and \(\beta < 3\), these simulations show no gravitational instability, remaining largely static for tens to hundreds of orbital timescales. Self-gravity is not strong enough to excite gravitationally unstable modes. At our medium resolution of \(N^3 = 512^3\), we found that \(Q_0 = 1\) still shows no evidence for gravitational instabilities, suggesting that this is not a resolution-dependent effect. For this reason, we use a new value of \(Q_0 = 0.676\), a value close to what is typically assumed for disks of finite thickness (Kim et al. 2002; Wang et al. 2010).

We then conducted our simulations with this new \(Q_0\) at three resolutions and at four cooling timescales as summarized in Table 1.

| Simulation   | Grid Cells \((N^3)\) | \(Q_0\) | \(\beta\) | Fragmentation |
|--------------|----------------------|--------|----------|--------------|
| Q256e2       | 256\(^3\)            | 1      | 2        | No           |
| Q256e10      | 256\(^3\)            | 1      | 10       | No           |
| Q512e2       | 512\(^3\)            | 1      | 2        | No           |
| Q512e10      | 512\(^3\)            | 1      | 10       | No           |
| Q512e5       | 512\(^3\)            | 0.676  | 0.5      | No           |
| G256e1       | 256\(^3\)            | 0.676  | 1        | No           |
| G256e2       | 256\(^3\)            | 0.676  | 2        | No           |
| G256e10      | 256\(^3\)            | 0.676  | 10       | No           |
| G512e2       | 512\(^3\)            | 0.676  | 2        | Yes          |
| G512e5       | 512\(^3\)            | 0.676  | 5        | No           |
| G512e10      | 512\(^3\)            | 0.676  | 10       | No           |
| G512e20      | 512\(^3\)            | 0.676  | 20       | No           |
| G1024e2      | 1024\(^3\)           | 0.676  | 2        | Yes          |
| G1024e5      | 1024\(^3\)           | 0.676  | 5        | No           |
| G1024e10     | 1024\(^3\)           | 0.676  | 10       | No           |
| G1024e20     | 1024\(^3\)           | 0.676  | 20       | No           |

Table 1. The three different cooling timescales are selected assuming a critical fragmentation value of \(\beta \sim 3\). For \(\beta = 2\), we expect fragmentation in the classical scenario of 2D disk fragmentation (Gammie 2001) and for \(\beta = 5, 10, 20\), we expect a stable gravitoturbulent disk, but are also concerned with the possibility of non-convergent effects, where fragmentation is possible for cooling timescales longer than \(\beta = 3\).

Using this lower value of \(Q_0 = 0.676\) immediately resulted in the unstable behavior expected from local disk instability simulations. While \(Q_0 = 0.676\) is necessary to initiate GI due to the vertically stratified density (see Figure 2), it eventually settles to \(Q = 1\) as GI develops because the vertical structure of the disk becomes much less stratified as the disk warms up and \(Q\) rises. This is shown in Figure 3, where the box-averaged Toomre stability parameter starts at the initial unstable \(Q_0 = 0.676\) before stabilizing around \(Q = 1\). The initial oscillations correspond to the adjustment to hydrostatic equilibrium and do not affect the outcome. Figure 2 shows...
the average vertical density and temperature profiles over time in a fragmenting simulation at our medium resolution. Similar to the fiducial $Q = 1$ simulations, the vertical density profile makes a slight adjustment to reach hydrostatic equilibrium and once gravitational instability sets in, mass concentrates in the midplane with time.

Fragmentation was observed in our high-resolution simulations for critical cooling timescales of $2 < \beta_{\text{crit}} < 5$. However, in contrast to Gammie (2001), our $N^3 = 256^3$ simulations never fragmented in either 2D or 3D, even at $\beta = 0.5$. That the high-resolution cases agree with earlier results suggests the problem is with the resolution rather than with the introduction of the third dimension. We believe the lack of fragmentation in our low-resolution runs might arise from higher numerical dissipation in Pencil as compared with ZEUS.

These low-resolution runs also show that our models are not susceptible to premature or spurious fragmentation, an issue that affects simulations with poor resolution or inadequate error suppression. Spurious errors may produce density perturbations that may result in undesired fragmentation (Truelove et al. 1997; Nelson 2006). Inability to maintain a fragment for more than a few orbits does not necessarily mean that we cannot say something about their ability to fragment, however. As we will show in Section 4.1, one can derive a density threshold that designates the formation of a fragment, even if it is soon disrupted.

Figure 4 shows the maximum density and average temperature for our $N^3 = 512^3$ (left) and $N^3 = 1024^3$ (right) simulations and demonstrates the anticipated fragmentation behavior of our 3D self-gravitating setups. At our medium resolution, $N^3 = 512^3$, fragmentation occurs in the case where the cooling timescale is short enough to remove pressure support and the clump collapses before being torn apart by shear. For longer cooling timescales, the fragment cannot collapse fast enough to remain bound and is soon disrupted; the difference between the two states is illustrated in Figure 1. Also shown in Figure 4 are the box-averaged temperatures over time, showing that the fragmenting simulations are able to withstand the rising temperatures as the fragment collapses and becomes denser.

We ran identical initial conditions with at highest resolution, $N^3 = 1024^3$. The right plot of Figure 4 shows the results of these simulations and they also show a fragmentation boundary of $\beta_{\text{crit}} = 3$, showing that our models converge with resolution.

Viscous stresses are only well parametrized in the case the disk is stable/gravitoturbulent, as is shown with the runs in Figures 5 and 6. Broken down into the hydrodynamic and gravitational components of the $\alpha$ stress, the total in Figure 5 averages to $\alpha = 0.21$ after settling at around $t = 50 \Omega^{-1}$. However, at the medium resolution in Figure 6, the $\alpha$ stress agrees with Equation (12) to within 10%, suggesting that the inability of the $\alpha$ stress to reach expected levels at low resolutions is related to the inability to fragment and the higher numerical dissipation at lower resolution.

4.1. Fragmentation

In the 2D case, fragmentation can be easily established with a Roche surface density threshold (See Section 2.4 of Baehr & Klahr 2015), but this threshold does not translate to 3D simulations and a volume density threshold. Instead, we find the pressure-modified Hill condition of Kratter & Murray-Clay.
to result in a suitable threshold density. This threshold considers that pressure support stabilizes a fragment against collapse and leaves it more susceptible to tidal shear. Starting with the condition that a fragment has a radius that is a fraction of its Hill radius,

$$r_{\text{frag}} = \left( \frac{1}{f} \right)^{1/3} \left( \frac{M_{\text{frag}}}{3M_*} \right)^{1/3} R,$$

(25)

where $M_*$ is the mass of the central star, $R$ is the location away from the center, and $f$ is factor determined by the choice of specific heat ratio $\gamma$. Defining the fragment mass as the integration of density over radial shells,

$$M_{\text{frag}} = 4\pi \int_{h_{50}}^{h_{\text{frag}}} \rho_h \left( \frac{r}{r_h} \right)^k r^2 dr,$$

(26)

with the density at the center of the fragment $\rho_h$ at a size one-fiftieth the scale height $h_{50}$, a fragment outer radius of $r_{\text{frag}}$ and a power-law index of the fragment density distribution $k$.

Then, inserting the fragment mass into Equation (25) and assuming $h_{50} \ll r_{\text{frag}}$, we then solve for the central density $\rho_h$:

$$\rho_h = \frac{3f(3-k)M_*}{4\pi R^3} \left( \frac{r_{\text{frag}}}{h_{50}} \right)^k.$$

(27)

We use $\gamma = 5/3$ which makes $f = 4$, whereas a softer choice of $\gamma = 7/5$ would mean $f = 16$, resulting in a slightly higher central density. Equation (27) will serve as our threshold density for our simulations, provided that the maximum density is above this threshold for several timescales, thus separating fragments from overdensities that have not yet collapsed enough.

Finally, we wish to apply our fragment threshold, Equation (27), to our simulations, from which we need to extract the density profile of a fragment. We find that a fragment scales with $k = 0.7$ to a fragment radius of...
which, for a protoplanetary disk around a solar-mass star, results in a density of $1.2 \times 10^{-10}$ g cm$^{-3}$ at 50 au and $1.5 \times 10^{-11}$ g cm$^{-3}$ at 100 au. Comparing this with our simulations and with the Roche density in Figure 7, we see that this density threshold more adequately delineates between fragments and non-fragments. A disk with the temperatures as in these simulations will have an aspect ratio $H/R \approx 0.11$ at 50 au and $H/R \approx 0.16$ at 100 au. With these densities and aspect ratios, we approximate threshold fragment masses of $5 M_J$ and 15 $M_J$ at 50 and 100 au, respectively.

Applying the density threshold to our low-resolution simulations, which failed to form sustained fragments as can be seen in Figure 8, simulations with very short cooling times ($\beta = 1, 0.5$) surpass this threshold. If the maximum density of a fragment can surpass this critical density for more than an orbit, it can be considered a fragment even if it is subsequently disrupted. Thus, this threshold may provide a useful sink cell formation criterion for simulations with resolution too low to fragment.

5. Discussion

5.1. Convergence

Following the results of Young & Clarke (2015), which demonstrated the need for a smoothing length to achieve convergent disk fragmentation in thin disk simulations, we expected our 3D simulations would converge, as a smoothing length is no longer required for a disk with finite thickness. There might be additional effects which affect fragmentation at different resolutions including, but not limited to, cooling, viscosity and numerical effects, but investigating the extension of local self-gravitating disks to include a vertical component is the most natural extension of that work and our results support their findings in this regard.

Recent work has focused on numerical effects and their influence on non-convergence. Klee et al. (2017) found that numerical oversteepening could be responsible for the over-estimation of overdense peaks resulting in fragmentation. Through our use of sixth-order hyperdiffusion, we do not expect oversteepening to be a significant factor. Deng et al. (2017) find that dissipation of angular momentum in SPH methods removes shear support from regions with high flow velocities and as this depends on resolution, fragmentation becomes easier with increasing resolution.

While our results strongly suggest convergence, at our lowest resolution fragmentation is suppressed due to insufficient resolution for both 2D and 3D models, similar to the results of Turk et al. (2012), which found that turbulent velocity fluctuations suppressed the formation of small scale magnetic dynamos when the Jeans length is not resolved by at least 64 grid cells. At higher resolutions, the results follow the standard fragmentation scenario, with clear fragmentation at $\beta = 2$ and a lack of fragmentation evident at $\beta = 5, 10, 20$. The $\beta$ cooling prescription is not a true substitute for realistic cooling, only a simple parametrization of the underlying physics. The use of a fixed background term may have an effect on convergence (Rice et al. 2012), thus one must be careful that convergence is not lost if this term is removed (Lin & Kratter 2016). However, as simulations using $\beta$ cooling and background irradiation converge, it supports our understanding of the physics.

5.2. Reconciling with the Thin Disk Approximation

A large part of this investigation has been about the differences between 2D and 3D simulations of self-gravitating disks, but it is also necessary to discuss how they are similar. Vertical structure and the accompanying gravitational dilution mean disk instability and fragmentation require even higher gas surface densities to collect locally before GI can set in. Once instabilities develop, the results of 2D simulations with smoothed self-gravity and 3D simulations show similar structures and features on similar scales (Figure 9).
Gravitoturbulent simulations in 3D show non-axisymmetric structures of the same scale as in 2D, but do not show as strong a contrast as 2D simulations due to the weaker self-gravity on scales smaller than $H$.

Ultimately, both 2D and 3D simulations converge to the same critical cooling timescale, but 3D simulations take significantly more resources. While 3D simulations are convergent and can capture effects at small scales, there is growing evidence that scales smaller than $\sim H$ contribute little to the growth of gravitationally unstable modes (Young & Clarke 2015; Riols et al. 2017). Thus, for the modeling of self-gravitating disks, 2D simulations may be sufficient so long as an appropriate smoothing length is used to account for the vertical scale of the disk.

5.3. Fragmentation and Planet Formation

Planet formation by gravitational instability is hindered by the large disk masses required. With the requisite $Q_0$ for fragmentation decreasing in these simulations, it is even more difficult to form planets by GI by requiring either more mass or a colder disk for fragmentation. Larger relative disk sizes required for fragmentation also suggest that the resulting fragments will draw from a larger mass reservoir to overcome stabilizing effects and form larger objects. Thus, it is more likely that fragments already have enough mass to be considered brown dwarfs or low-mass stars, without taking into account post formation accretion and growth.

The density threshold of Equation (28) and the subsequent fragment mass suggest objects formed in these simulations are formed with at least 4.9 $M_\odot$ at both 50 au and 100 au around a solar-like star. Since companions will generally form larger than these thresholds and subsequently accrete more mass, these fragments will likely form planets many Jupiter masses in size and brown dwarfs as well. Recent observations show fragments are indeed quite large, as is the case of L1448 IRS3B (Tobin et al. 2016) which is consistent with theoretical expectations (Stamatellos & Whitworth 2009).

5.4. Further Considerations

Global simulations can model the growth, migration, and potential destruction of fragments formed in a disk, but with our shearing box simulations are able to resolve the small scales of GI structure and development. For a resolution and convergence study this makes local simulations ideal for the relatively small patches of the disk where fragmentation occurs.

The Fourier solver in Pencil is limited to domains with the same number of grid cells in the vertical direction as in the $x$-direction, meaning that a simulation with a flattened box would have significantly higher resolution in $z$, resulting in different values of hyperviscosity in each direction, possibly leading to unphysical effects. For that reason, we conducted our simulations spanning roughly 12.7 scale heights in the vertical direction, much of which was not crucial to the simulation of fragmentation. Accommodating these largely empty spaces require carefully establishing a minimum density low enough to prevent undesired additional mass from settling onto the disk and affecting fragmentation. A more efficient approach would be to simulate the vertical direction in four to six scale heights while still retaining the radial and azimuthal extents. Additionally, the periodic vertical boundary conditions for the self-gravitational potential rather than typical outflow conditions and may contribute to the somewhat erratic $\alpha$ stresses calculated in Figure 6.

While the simple cooling with irradiation used here might result in convergence of the cooling criteria, a better description of collapse and fragmentation will ultimately come from using radiative transfer calculations, whether in the form of approximations such as flux-limited diffusion (Boley et al. 2006; Mayer et al. 2007) or more sophisticated but intensive methods like ray tracing. These methods take into account the gas opacities and optical depths to more appropriately model variations of cooling from one region of the disk to the next.

One feature of protoplanetary disks that is typically ignored in self-gravitating contexts but might have an effect on fragmentation is the disk magnetic field. While ionization is
expected to be low in the cold outer regions of circumstellar disks, the effect of magnetic fields on fragmentation is uncertain. Only a pair of studies have looked into the effects of magnetic fields on the fragmentation criterion (Riols & Latter 2016; Forgan et al. 2017), and came to opposite conclusions about the importance of magnetic fields on disk fragmentation.

6. Conclusions

In this paper, we conducted a converge study of 3D shearing box self-gravitating disks with simple cooling and background irradiation. Starting with a marginally unstable disk, simulations were run at various cooling timescales through the initial burst phase where they have either formed a fragment or settled to gravitoturbulence. These simulations were run at several high resolutions to investigate fragmentation conditions and the convergence of the solution with respect to the cooling timescale and our results are as follows:

1. Consistent with analytic results, we find that 3D vertically stratified disk simulations fragment for two high-resolution simulations with a cooling timescale boundary of roughly $\beta_{\text{crit}} = 3$. Fragmentation only occurs when a scale height is resolved by at least 40 grid cells.

2. Convergence indicates that there is a firm criterion for fragmentation in a disk: where disks cool rapidly with $\beta < 3$. This limits fragmentation to the distant regions around a star where temperatures are relatively low and dominated by stellar irradiation.

3. By achieving convergence, we demonstrated that the standard $\beta$ cooling model with a fixed floor temperature may have no inherent flaws numerically when used in our 3D models. It is nevertheless desirable to use a more realistic model for disk cooling in future studies.

4. Using a pressure-supported Hill criterion, we find a suitable minimum density consistent with the formation of brown dwarfs and low-mass stars. This serves as a useful criterion for the inclusion of sink cells to aid in the efficient computation of fragmentation as it can be applied to simulations with low-resolution simulations which are unable to fragment.

The authors thank Chao-Chin Yang and Andreas Schreiber for useful discussions and assistance with the Pencil code. We also thank the anonymous referee for comments that improved the paper. K.M.K. received support for this research from the Max Planck Institute for Astronomy. Simulations shown were run on the Theo/Isaac clusters at the Rechenzentrum Garching (RZG) of the Max Planck Society and the JUQUEEN cluster of the Jülich Supercomputing Centre (Stephan & Docter 2015).

ORCID iDs

Hans Baehr @ https://orcid.org/0000-0002-0880-8296
Hubert Klahr @ https://orcid.org/0000-0002-8227-5467
Kaitlin M. Kratter @ https://orcid.org/0000-0001-5253-1338

References

Baehr, H., & Klahr, H. 2015, ApJ, 814, 155
Balbus, S. A., & Hawley, J. F. 1991, ApJ, 376, 214
Balbus, S. A., & Papaloizou, J. C. B. 1999, ApJ, 50, 650
Behrendt, M., Burkert, A., & Schartmann, M. 2015, MNRAS, 448, 1007
Boley, A. C., Mejía, A. C., Durisen, R. H., et al. 2006, ApJ, 651, 517
Boss, A. P. 1997, Sci, 276, 1836
Boss, A. P. 2017, ApJ, 836, 53
Brandenburg, A. 2001, in Advances in Non-linear Dynamos, ed. A. Ferriz-Mas & M. Nunez (London: Taylor and Francis), 269
D’Alessio, P., Canto, J., Hartmann, L., Calvet, N., & Lizano, S. 1999, ApJ, 511, 896
Deng, H., Mayer, L., & Mera, F. 2017, arXiv:1706.00417
Durisen, R. H., Boss, A. P., Mayer, L., et al. 2007, in Protostars and Planets V, ed. B. Reipurth, D. Jewitt, & K. Keil (Tucson, AZ: Univ. Arizona Press), 607
Forgan, D., Price, D. J., & Bonnell, I. A. 2017, MNRAS, 466, 3406
Gammie, C. F. 2001, ApJ, 533, 174
Goldreich, P., & Lynden-Bell, D. 1965, MNRAS, 130, 125
Hawley, J. F., Gammie, C. F., & Balbus, S. A. 1995, ApJ, 440, 742
Hirose, S., & Shi, J.-M. 2017, MNRAS, 469, 561
Hubeny, I. 1990, ApJ, 351, 632
Huré, I.-M., & Pierens, A. 2009, A&A, 507, 573
Kim, W.-T., Strickler, E. C., & Stone, J. M. 2002, ApJ, 581, 1080
Klee, J., Illensieer, T. F., Jung, M., & Duschl, W. J. 2017, arXiv:1704.02193
Kratter, K. M., & Lodato, G. 2016, ARA&A, 54, 271
Kratter, K. M., & Murray-Clay, R. A. 2011, ApJ, 740, 1
Kuiper, G. P. 1951, PNAS, 37, 1
Lichtenberg, T., & Schleicher, D. R. G. 2015, A&A, 579, A32
Lin, M.-K., & Kratter, K. M. 2016, ApJ, 824, 91
Lissauer, J. J. 1993, ARA&A, 31, 129
Lodato, G., & Clarke, C. J. 2011, MNRAS, 413, 2735
Mamatsashvili, G. R., & Rice, W. K. M. 2010, MNRAS, 406, L56
Mejia, A. C., Durisen, R. H., & Pickett, M. K. 2005, ApJ, 619, 1098
Michael, S., Steiman-Cameron, T. Y., Durisen, R. H., & Boley, A. C. 2012, ApJ, 746, 98
Mordasini, C., Alibert, Y., Benz, W., & Henning, T. 2012, A&A, 541, 97
Müller, T. W. A., Kley, W., & Mera, F. 2012, A&A, 541, A123
Nelson, A. F. 2006, MNRAS, 373, 1039
Paardekooper, S.-J. 2012, MNRAS, 421, 3286
Rafikov, R. R. 2004, ApJL, 621, L69
Rafikov, R. R. 2016, ApJ, 831, 122
Rice, W. K. M., Forgan, D. H., & Armitage, P. J. 2012, MNRAS, 420, 1640
Rice, W. K. M., Lodato, G., & Armitage, P. J. 2005, MNRAS, 364, L56
Rice, W. K. M., Paardekooper, S.-J., Forgan, D. H., & Armitage, P. J. 2014, MNRAS, 438, 1593
Riols, A., & Latter, H. N. 2016, MNRAS, 460, 2223
Riols, A., Latter, H. N., & Paardekooper, S.-J. 2017, MNRAS, 471, 317
Safronov, V. S. 1969, Evolution of the Protoplanetary Cloud and Formation of the Earth and Planets (Jerusalem: Israel Program for Scientific Translations), 212
Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337
Shi, J.-M., & Chiang, E. 2013, ApJ, 764, 20
Shi, J.-M., & Chiang, E. 2014, ApJ, 789, 34
Stamatellos, D., & Whitworth, A. P. 2009, MNRAS, 392, 413
Stephan, M., & Docter, J. 2015, Journal of large scale computing facilities, 1, A 1
Taylor, J., Kratter, K. M., Persson, M. V., et al. 2016, Natur, 538, 485
Toomre, A. 1964, ApJ, 139, 1217
Truelove, J. K., Klein, R. I., McKee, C. F., et al. 1997, ApJL, 489, L179
Tur, M. K., Oishi, J. S., Pudritz, R. E., & Bryan, G. L. 2012, ApJ, 745, 154
Vorobyov, E. I. 2013, A&A, 552, A129
Wang, H. H., Klessen, R. S., Dullemond, C. P., van den Bosch, F. C., & Fuchs, B. 2010, MNRAS, 407, 705
Yang, C.-C., & Knutson, M. R. 2012, ApJ, 758, 48
Young, M. D., & Clarke, C. J. 2015, MNRAS, 451, 3987