# Protoplanetary disks and planet formation
# A computational perspective

Isaac Backus

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Marcel den Nijs, Chair

Tom Quinn, Chair

Jason Detwiler

Eric Agol

Program Authorized to Offer Degree:
Physics

University of Washington

**Abstract**

Protoplanetary disks and planet formation
A computational perspective

Isaac Backus

Co-Chairs of the Supervisory Committee:
Professor Marcel den Nijs
Physics

Professor Tom Quinn
Astronomy

In this thesis I present my research on the early stages of planet formation. Using advanced computational modeling techniques, I study global gas and gravitational dynamics in protoplanetary disks (PPDs) on length scales from the radius of Jupiter to the size of the solar system. In that environment, I investigate the formation of gas giants and the migration, enhancement, and distribution of small solids—the precursors to planetesimals and gas giant cores.

I examine numerical techniques used in planet formation and PPD modeling, especially methods for generating initial conditions (ICs) in these unstable, chaotic systems. Disk simulation outcomes may depend strongly on ICs, which may explain results in the literature. I present the largest suite of high resolution PPD simulations to-date and argue that direct fragmentations of PPDs around M-Dwarfs is a plausible path to rapidly forming gas giants.

I implement dust physics to track the migration of centimeter and smaller dust grains in very high resolution PPD simulations. While current dust methods are slow, with strict resolution and/or time-stepping requirements, and have some serious numerical issues, we

can still demonstrate that dust does not concentrate at the pressure maxima of spiral arms, an indication that spiral features observed in the dust component are at least as well resolved in the gas. Additionally, coherent spiral arms do not limit dust settling. We suggest a novel mechanism for disk fragmentation at large radii driven by dust accretion from the surrounding nebula. We also investigate self induced dust traps, a mechanism which may help explain the growth of solids beyond meter sizes. We argue that current apparent demonstrations of this mechanism may be due to numerical artifacts and require further investigation.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# GLOSSARY

## *Variables*

$M_*$: Stellar mass

$Q$: Toomre stability parameter for PPDs. $Q \equiv \frac{c_s \kappa}{\pi G \Sigma}$

$Q_{\rm eff}$: $\equiv \beta Q (H/R)^\alpha$, the effective Toomre $Q$ with a correction for disk height which tends to stabilize the disk. $\beta = 2.1$ and $\alpha = 0.18$ are fit such that disks with a minimum $Q_{\rm eff} < 1$ will fragment

$R$: Radius in cylindrical coordinates. Unless otherwise specified, assume cylindrical coordinates with $\hat{z}$ along $\mathbf{\Omega}$

$R_d$: PPD radius

$M_d$: PPD mass

$\Sigma$: Disk surface density, defined as $\Sigma \equiv \int_{-\infty}^{+\infty} \rho(z) dz$

$\alpha_{\rm S}$: Shakura & Sunyaev viscosity parameter, measures the strength of viscosity in a disk

$\Omega$: Orbital angular velocity

$\Omega_K$: $\equiv \sqrt{GM_*/R^3}$, Keplerian orbital angular velocity

$\kappa$: Unless otherwise noted, epicyclic frequency. $\kappa \equiv \frac{2\Omega}{R} \frac{\rm d}{{\rm d}R}(R^2\Omega)$, which for a Keplerian disk reduces to $\Omega$

$c_s$: $= \sqrt{\gamma k_B T/m}$, ideal gas sound speed. $\gamma = 1$ is used for an isothermal EOS

$\gamma$: Adiabatic index. Note, in many applications we take $\gamma \to 1$ for an isothermal EOS

$\rho$: Volume density (gas or gas + dust)

$\rho_g, \rho_d$: Gas and dust volume densities

$\rho_{\text{grain}}$: Intrinsic grain density, e.g. 1 g/cm$^3$ for ice or $\sim 3$ g/cm$^3$ for silica

$s_{\text{grain}}$ or $s$: Dust grain size (unless otherwise noted). Radius for spherical grains

$\epsilon$: Dust mass fraction, $= \rho_d/\rho$

$\epsilon_s$: Gravitational softening length

$h$: Smoothing length, or dimensionless height, context dependent

$W(|\mathbf{r} - \mathbf{r}_b|, h)$, $W_{ab}(h)$: SPH kernel

***Abbreviations***

ALMA: Atacama Large Millimeter/submillimeter Array, currently the premier telescope for PPD observation

AV: Artificial viscosity–a viscosity-like term used in SPH

CA: Core accretion

EOS: Equation of state, relating pressure and energy for an ideal gas, e.g. adiabatic or isothermal

GI: Gravitational instability

ICs: Initial conditions

MHD: Magnetohydrodynamics

ORP: Outer Rotational Period for a disk, typically take at the disk radius $R_d$

PPD: Protoplanetary disk

SPH: Smoothed particle hydrodynamics

# ACKNOWLEDGMENTS

First, I would like to thank my advisor over the last few years, Tom Quinn. Through tough times and easy, a person could not ask for a better advisor: so helpful, kind, caring, knowledgeable, and so able to push me to do my best.

I would like to thank Aaron Boley for many fruitful discussions, so much feedback, and for all his help on this work. I thank Joey Rucska for his collaboration on dusty-gas physics and grain-growth, and would also like to thank James Wadsley for so much help. I'd like to David Fleming for help on my work and on DISKPY.

Of course, I would like to thank all my friends and family for years of support. And most of all, I thank my wife Juliette for her endless strength and help as I finished my studies, even as she was so endlessly busy with immigration, with architecture licensing, and with work.

This work was performed as part of the NASA Astrobiology Institute's Virtual Planetary Laboratory, supported by the National Aeronautics and Space Administration through the NASA Astrobiology Institute under solicitation NNH12ZDA002C and Cooperative Agreement Number NNA13AA93A. The authors were also supported by NASA grant NNX15AE18G. This work used the Extreme Science and Engineering Discovery Environment (XSEDE)[**?**], which is supported by National Science Foundation grant number ACI-1053575. This work was facilitated through the use of advanced computational, storage, and networking infrastructure provided by the Hyak and Mox supercomputer systems at the University of Washington, and through use of the Stampede and Stampede2 supercomputers at the Texas Advanced Computing Center.

# DEDICATION

To my dear wife, Juliette, without whom my grains would not grow, my cores would not accrete, my disks would not fragment, and my planets would never form

# Chapter 1

# INTRODUCTION

Increased computing power, advances in computational techniques, and continually developing observational advances over the last several decades have allowed astronomers and physicists to develop increasingly sophisticated models of planet formation. Due to the mathematically and computationally difficult nature of the problem, and due to limited observational evidence, there remain many unanswered questions. The dynamics encountered in modeling planet formation are highly non-linear, span many orders of magnitude in length- and time-scales, and are difficult to probe with telescopes or in the lab. Despite this, a general picture has emerged.

The basic stages of planet formation are as follows. Gas and dust collapse in a molecular cloud core to form a young star. Due to stability and angular momentum considerations, the gas and dust surrounding the star form a thin disk orbiting at approximately Keplerian angular velocity. As this protoplanetary disk (PPD) is formed, the dust will tend to settle to the disk midplane [**?**]. The dust is expected to coagulate and eventually form planetesimals. At this stage, there are three main mechanisms for planet formation: (1) terrestrial planet formation through collisions of planetesimals; (2) giant planet formation through core accretion (CA); and (3) giant planet formation via gravitational instability (GI). Here I review the current understanding of PPDs and these mechanisms for planet formation. I then outline the general computational approach used here.

## *1.1 Protoplanetary disks*

Protoplanetary disks may exist around stellar binaries or with external binary companions, but for this simple picture let us assume a single central star of mass $M_*$. PPDs are mainly

gaseous, with abundances assumed to be similar to those of the progenitor stellar nebula, mainly hydrogen and helium, with disk masses somewhere around a few percent of $M_*$. Disks may extend out to radii of tens to hundreds of AU. The scale-height can be estimated as

$$H = c_s \Omega \tag{1.1}$$

where $\Omega$ is the orbital angular velocity and $c_s = \sqrt{\gamma k_B T/m}$ is the adiabatic sound speed for an ideal gas at temperature $T$ with a mean molecular weight of $m$. This approximation assumes $H \ll R$ and a negligible disk self-gravity compared to the star's gravity. Typical disk aspect ratios are around $H/R \approx 0.05$ (with a large scatter). To within around 10%, orbital angular velocities are approximately Keplerian: $\Omega \approx \Omega_K \equiv \sqrt{GM_*/R^3}$, where $R$ is the cylindrical radius.

While the viscosity of an ideal gas in typical PPD conditions is expected to be very small, disks are thought to have an effective viscosity due to small scale turbulence driven largely by growth of the magnetorotational instability (MRI) in the presence of magnetic fields in a slightly ionized disk [**???**]. The exact details of the magnetohydrodynamics (MHD) and turbulence are an area of ongoing research [**???**], but to low order the general scaling of the viscosity $\nu$ is usually parameterized according to the method of **?** as:

$$\nu = \alpha_{\rm S} \frac{c_s^2}{\Omega} \tag{1.2}$$

Where $\alpha_{\rm S}$ is a dimensionless parameter, often assumed to be approximately constant. Typical values used are of order $\alpha_{\rm S} \approx 0.01$, although much smaller or larger values may be reasonable. In a realistic disk, $\alpha_{\rm S}$ is probably not constant. Disks may possess "dead-zones": regions at intermediate radial locations, centered on the midplane $z = 0$ with a scale-height comparable to the disk scale-height where ionization and magnetic fields are suppressed, thus dropping the turbulence and effective viscosity dramatically [**?**].

Viscosity transports angular momentum, causing the disk to spread radially on time-scales of order the disk lifetime of Myrs. This angular momentum transport drives the inner regions of the disk to migrate radially inward, driving accretion onto the star. Outer regions

of the disk migrate radially outward. These viscous interactions can be a main source of heating in the disk.

Disk temperatures are thought to be approximately uniform along much of the vertical direction, with a layer with a different temperature at the large $|z|$ boundaries. By convention, we orient the disk in the $x-y$ plane such that $\mathbf{\Omega}$ points along $+z$. Along the radial direction, viscous heating and stellar irradiation dominate the disk temperature profile. Generally speaking, at the earliest stages of disk evolution, viscous heating may dominate, and later on stellar irradiation will, although this depends on the region of the disk under consideration. A stellar irradiation dominated profile for standard disks may be approximated by a powerlaw [**?**]:

$$T(R) = T_0(R/R_0)^{-q} \tag{1.3}$$

Where $q$ is between 3/7 for a fully flared disk and 3/4 for a flat disk. $T_0$ is the disk temperature at $R_0$. A viscous dominated disk profile will have the functional form [**?**]:

$$T^4(R > R_{in}) \propto \frac{1}{R^3}\left(1 - \sqrt{\frac{R_{in}}{R}}\right) \tag{1.4}$$

where $R_{in}$ is some inner radius of the disk with vanishing viscosity, where presumably the temperature will be increased by stellar irradiation. Taken together, these models indicate temperature profiles will decrease away from the star.

The surface density $\Sigma = \int_{-\infty}^{\infty} \rho(z)dz$ is thought to decrease away from the star for most of the disk but with a maximum somewhere in the inner disk. For simplicity, $\Sigma$ is often parameterized as a powerlaw:

$$\Sigma(R) = \Sigma_0(R/R_0)^{-p} \tag{1.5}$$

$\Sigma_0$ is the value at $R_0$, and typically $p = 1$. Another common parameterization, with more physical motivation, is the similarity solution for a viscous disk [**?**]:

$$\Sigma(R) \propto (R/R_c)^{-l} \exp[-(R/R_c)^{2-l}] \tag{1.6}$$

With $R_c$ being a length-scale parameter and $l$ some constant.

The $\Sigma$ and $T$ profiles create a disk with an outwardly decreasing pressure. This pressure gradient generates a force on the gas directed radially outward, thus decreasing the centripetal force, causing the gas to orbit more slowly than the solid components. Disk self gravity will cause further deviation from Keplerian orbit.

Eventually, photoevaporation and viscous spreading/accretion will cause disks to dissipate on time-scales of Myrs [**?**].

### *1.2 Planetesimal formation*

Planetesimals are compact, gravitationally bound bodies bigger than ∼100–1000m, formed by the aggregation of dust in the PPD. They are thought to be precursors to terrestrial planets and to the cores of giant planets. In the molecular cloud core precursor of the young star and the PPD there are expected to be micrometer sized dust grains [**?**]. At such a small size, the dust is strongly coupled to the gas, but as the dust grains grow through collisions in the PPD to mm sizes they will tend to settle to the mid-plane of the disk [**???**].

These small dust particles, coupled strongly to the gas, will undergo collisions with small relative velocities, due to Brownian motion. They will tend to aggregate and grow to centimeter sized objects. To continue growing to the size of planetesimals and larger, these pebbles must overcome two challenges, for which there is no generally accepted solution: the meter barrier, and sticking during collisions [**?**].

The meter barrier [**?**] refers to the phenomenon where meter sized objects will rapidly drift toward the central star, on timescales much shorter than those relevant to planet formation. As objects grow, they begin to decouple from the gas flow as their Reynolds number increases. Due to radial pressure gradients, the orbital velocity of the gas is lower than that of the solids. The solids feel a headwind, causing them to lose energy and fall inward. For very small and very large objects, this will be a minor effect, but for meter sized objects migration timescales can be of order 100 years, significantly shorter than any planet formation timescales.

The second major barrier is the problem of sticking during collisions. As has been seen in laboratory experiments, small particles, coupled to the gas, with small velocity disper-

sions must avoid bouncing in order to grow—the so called "bouncing barrier" [**???**]. Larger particles with higher velocity dispersions must avoid destruction during collisions [**?**]. Both of these processes will tend to stall grain growth at a maximum grain size.

I discuss some possible solutions to these problems in §1.6. If these barriers can be overcome, planetesimals 100-1000m and bigger are expected to form. These planetesimals will eventually form terrestrial planets and giant planet cores.

### *1.3 Core Accretion*

One method for giant planet formation with widespread acceptance is core accretion [see **?**; **?** for reviews]. Under the core accretion (CA) paradigm, a large rocky embryo is formed from planetesimals, which will become the core of a gaseous planet. Once the core is sufficiently large, typically some fraction of an Earth mass, it will rapidly accrete nearby planetesimals [**?**]. For illustration, Safronov's equation gives an estimate of the rate of planetesimal accretion onto the embryo:

$$\frac{\mathrm{d}M}{\mathrm{d}t} = F\pi R^2 \Sigma_s \Omega \tag{1.7}$$

Where $F$ is a gravitational focusing factor, $\pi R^2$ is the effective cross section of the embryo for capturing planetesimals, and $\Sigma_s$ is the surface density of solids in the PPD. $F$ depends strongly on the velocity dispersion of the solids, and therefore the state of the gas in the disk [**?**]. Stirring by the embryo can excite random motions of the planetesimal disk, decreasing $F$ and thereby stifling embryo growth [**?**]. Note that the accretion rate depends directly on the available mass of solids $\Sigma_s$, and therefore is sensitive to enhancement or depletion of solids.

Eventually, the embryo will grow large enough that the escape velocity will exceed the local thermal velocity ($v_{rms}$) of the gas. The protoplanet will then rapidly accrete gas and planetesimals until it clears a gap in the disk through gravitational torques. After this, the planet will continue to slowly accrete gas, and may undergo inward or outward migration [for example, see **?**)]. If the protoplanet can cool efficiently, it will contract and the gaseous envelope will become permanently bound, and may survive tidal interactions. Planetary

collapse can be further accelerated when full 3D models are considered as compared to 1D spherical collapse models [**?**].

Under the CA model, gas giants are naturally expected to preferentially form outside of the inner disk. The annulus cleared out by the protoplanet should be of order 1 Hill radius in width, so the total available gas and solids will increase with disk radius, so long as the surface density decreases sufficiently slowly. Additionally, beyond the snow line ice will form, increasing the available solids. The snow line occurs where the temperature drops below 170 $K$, somewhere in the vicinity of $1 - 5$ AU for a solar type star [**?**].

### *1.4 Gravitational Instability*

In contrast to the bottom up formation scenario of core accretion, the gravitational instability model (GI) proposes giant planet formation seeded by global disk instabilities. As demonstrated by **?**, a differentially rotating, 2D disk orbiting a central point mass will be unstable to axisymmetric perturbations provided:

$$Q \equiv \frac{c_s \kappa}{\pi G \Sigma} > 1 \tag{1.8}$$

where $c_s$ is the local sound speed in the gas and $\kappa$ is the epicyclic frequency ($\kappa \approx \Omega$ for very light disks). $Q$ is called the Toomre parameter. A simple physical interpretation of the above Toomre instability criterion can be outlined by considering under what conditions a small clump of gas will remain stable under compression. Compressing the gas increases the outward pressure and the inward gravitational forces. Additionally, because of the shearing in the disk, the compressed gas will spin faster to conserve angular momentum, thus creating an outward centrifugal force. If the pressure and centrifugal forces are larger than the gravitational force, the disk will remain stable. The above relation can be derived more rigorously from the dispersion relation for axisymmetric perturbations to a 2D disk [**?**]. While strictly speaking this is a 2D analysis, $Q$ has proven surprisingly effective at predicting stability in fully 3D numerical simulations—with the caveat that the minimum stable $Q$ value appears to differ from unity [**?????**].

For sufficiently unstable disks, perturbations will grow and form spiral arms, which may eventually collapse/break apart, forming dense, gaseous fragments. Figure 1.1 shows an example of a fully 3D simulation of such a gravitationally unstable PPD.

### *1.5 Observations*

Recent years have seen a surge in our understanding and knowledge of extra-solar planetary systems, driven in large part by advances in observation. Astronomers have found planets to be plentiful in our galaxy, and our improved estimates of planet populations have made the discovery of extraterrestrial life seem increasingly plausible. The Kepler mission has expanded our catalog of known planets by around 2800, with another ∼4700 candidates awaiting confirmation [**?**]. Radial velocity surveys performed with instruments such as HARPS having found and confirmed nearly 300 planets [**?**]. High-contrast direct imaging techniques have observed giant planets at large separations from their host stars [**?**]. Even gravitational microlensing measurements have proven sensitive enough to detect exoplanets [**?**].

Observations of PPDs have seen major improvements as well. Imaging of disks at millimeter and submillimeter wavelengths has allowed disk parameters such as mass and temperature to be probed, dating back over a decade [**???**]. Over 100 pre-main sequence disks have been resolved [**?**]. Next generation instruments, such as the Atacama Large Millimeter/submillimeter Array (ALMA) have greatly enhanced our knowledge of disk compositions, dynamics, and life-times. Stunning images of circumstellar disks (for an example, see figure 1.2) are now being released from campaigns that include observing disks in the Taurus molecular cloud [**?**], evidence of dead-zones [**?**], improved observation and modeling of the gas component in regions cleared of dust [**?**], and direct observations of disk turbulence [**?**].

These observational advances greatly inform and direct research into theoretical models of planet formation. Knowledge of disk masses, temperatures, sizes, and compositions, greatly constrain our models. Additionally, the final configurations of planetary systems dictate what results must be predicted by planet formation models. This marks an exciting time for

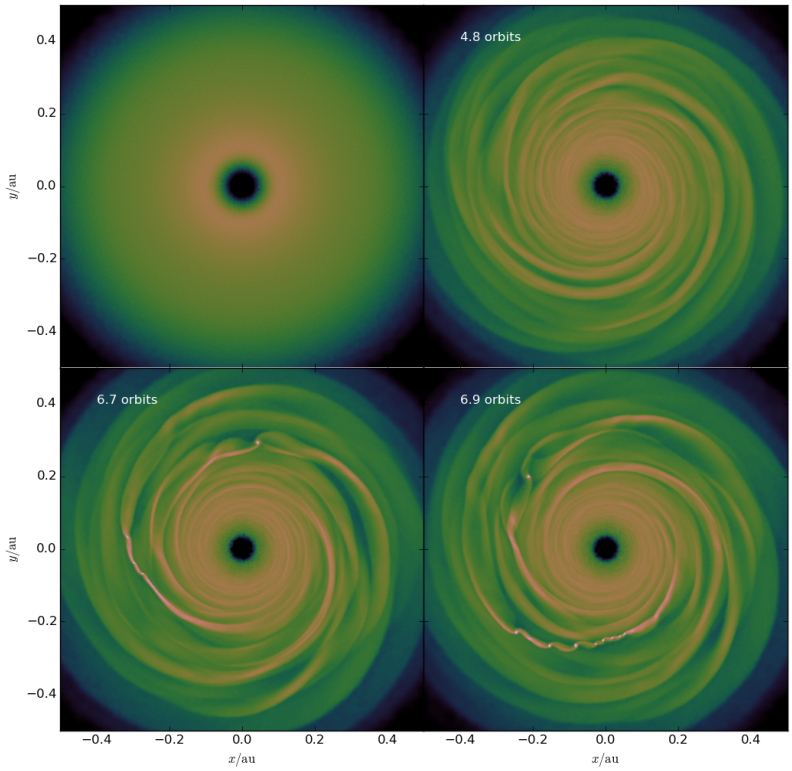



Figure 1.1: An example simulation of a gravitationally unstable disk. Pictured are four snapshots of density (with a logarithmic color scale). The initially smooth disk (top-left) rapidly develops spiral arms (top-right) within several orbits from gravitational instabilities seeded by Poisson noise in the spatial distribution of particles. Several orbits later, dense gravitationally bound clumps form from over-dense spiral arms (bottom-left), followed by rapid formation of more clumps in the now highly unstable disk (bottom-right). The disk orbits a central star of mass $M_* = 0.33 M_\odot$ and has a minimum Toomre parameter of $Q = 0.75$ at the disk radius $R_d = \frac{1}{3}$ AU. This $10^6$-particle SPH simulation, run with ChaNGa, models the disk as locally isothermal molecular hydrogen.

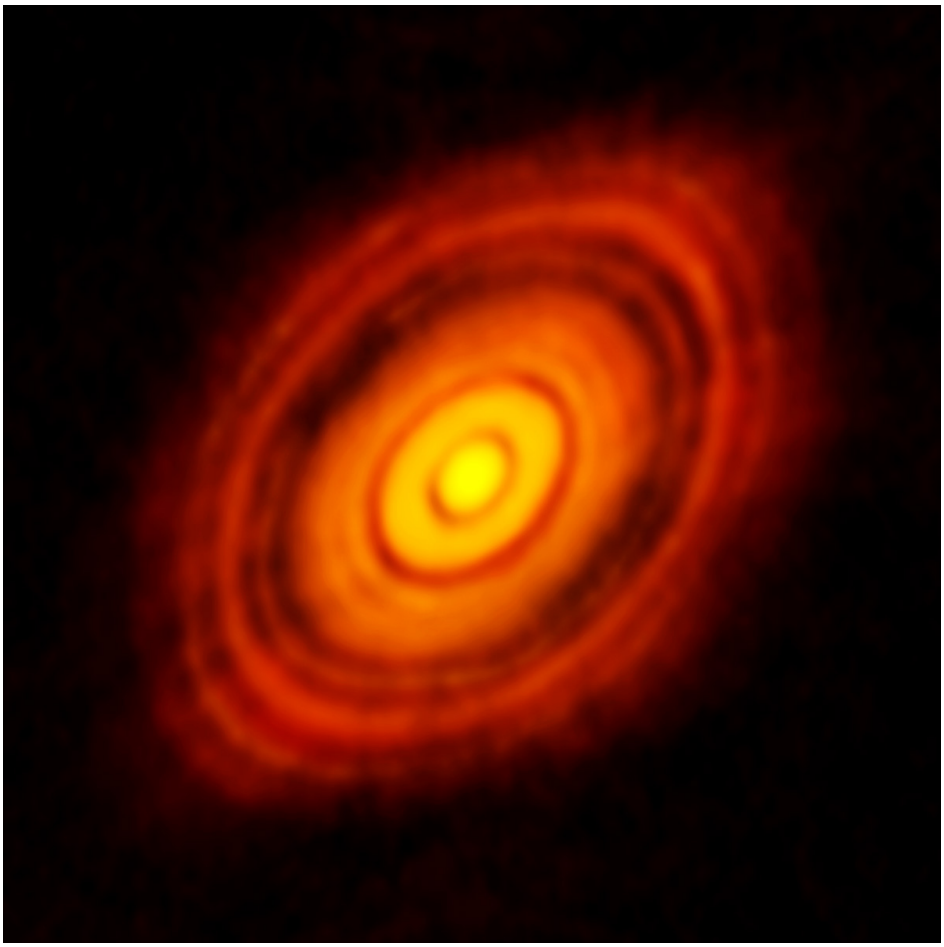


Figure 1.2: A remarkable image of the PPD around HL Tau. Taken from the ALMA website. See **?**

expanding our understanding of planet formation.

### *1.6 Growth barriers*

As discussed in §1.2, one of the largest gaps in our models of planet formation is the meter barrier. There is no universally agreed upon solution to this problem, though given the existence of planets a solution must exist. To cross the meter barrier and avoid migrating into the host star, either solids must grow extremely rapidly, or gas must be cleared out of some region of the disk. There are several proposed solutions to the meter barrier worth investigating, which may also be related to the bouncing and destruction barriers.

To accelerate the growth from dust to objects much larger than a meter, there will need to be enhancement of the solid to gas ratio. Dust is expected to settle to the midplane of the disk; however, the details of settling in a full, gravitationally active 3D disk are unknown. The amount of concentration due to settling is still open for investigation.

A proposed mechanism for concentration is dust enhancement seeded by gravitational instabilities. As shown in figure 1.1, large, non-axisymmetric density gradients can be driven by GI. As shown numerically in **?**, solids are expected to accumulate at pressure maxima due to gas drag and pressure gradients. However, the highly non-axisymmetric nature of spiral arms and of clumps due to fragmentation, along with dust migration, require global numerical simulations to probe.

Aaron Boley has proposed another mechanism for enhancing the concentration of solids. As mentioned in §1.2, objects approaching centimeter size and above must cross the destruction barrier. As particles approach the meter barrier, they will migrate inward rapidly and they will have larger velocity dispersions. High velocity collisions will tend to fragment and destroy these objects, thus returning their mass to the population of smaller solids. These small solids migrate much less rapidly and may pile-up in a specific disk region. This will enhance the concentration of small solids available for growth past the meter barrier.

This mechanism may operate in conjunction with the self induced dust-traps model proposed by **?**. Outer regions of the disk are expected to have lower small scale turbulent

velocities, due to decreased disk shearing and lowered sound speed. This, along with decreased gas densities, will stir small solids less, which can decrease their average relative velocities during collisions. This may allow them to avoid the destruction barrier and to grow efficiently to larger sizes. Solids that grow rapidly enough, and at sufficiently large radii, will migrate inward under gas drag and enter regions of high gas densities which will slow their migration. If enough dust can pile-up at the inner disk, gas will be cleared from some annulus by drag interactions with the dust. This in turn can create a local pressure maximum. The gas component feels pressure forces, while the dust does not. Ignoring magnetic fields, in the presence of gravity and hydrodynamics the specific forces on the dust and gas are related by:

$$\mathbf{f}_g = \mathbf{f}_d - \frac{1}{\rho_g}\nabla P \tag{1.9}$$

where $\mathbf{f}_g, \mathbf{f}_d$ are the specific forces (accelerations) on the gas and dust components, respectively. The gas density is given by $\rho_g$. $P$ is the gas pressure. Relative to the gas, the dust will therefore move in the direction of pressure gradients, toward pressure maxima. Dust will then accumulate at radial pressure maxima. This will clear more gas, deepening the "dust-trap", creating a region with low gas density where grains can grow beyond the meter barrier without migrating toward the star.

The formation of centimeter sized objects is also of interest to the growth of large cores. Typical estimates for the core growth time scales, especially in outer regions of the disk where giant planets are observed to exist, are as long or longer than the timescales of gas dissipation around the disk (1-10 million years). A possible solution to this problem [**?**] is that a disk with a high concentration of solids at the midplane, dominated by mm to cm sized grains, can rapidly grow cores. Key to this mechanism is that cm grains decouple sufficiently to migrate rapidly, creating a "rain" of solids onto young cores. This can cut core formation timescales by orders of magnitude.

## *1.7 Computational approach*

### *1.7.1 Challenges*

The physics involved in modeling protoplanetary disks provides a large computational challenge. Gravity is non-local and non-linear. Being non-local means all parts of a simulation are coupled to each other and any naive approach will therefore be $\mathcal{O}(N^2)$, where $N$ is the number of points sampled in space. Being non-linear and often unstable can drive complex, chaotic, and often counter-intuitive dynamics which may depend strongly on the initial conditions, the numerical methods, and the numerical parameters such as resolution.

There is an enormous range of variation in density, length, energy, and time-scales in PPDs, all of which may vary by orders of magnitude at different regions in the disk and at different stages of disk evolution. Accretion disks in general can be very difficult due to supersonic bulk orbital flow and sub-sonic local flow which must be simultaneously captured in a highly shearing environment. This poses serious challenges for handling advection, conservation, and viscosity. Accurately handling heating and cooling involves radiation physics, which is long range and naively $\mathcal{O}(N^2)$, especially in PPDs which typically include optically thin, thick, and intermediate regimes. Modeling solids also involves enormously different scales and can involve dynamical time-scales many orders of magnitude shorter than orbital timescales.

### *1.7.2 Smoothed Particle Hydrodynamics*

First introduced by **?** and **?**, Smoothed Particle Hydrodynamics (SPH) is a Lagrangian technique for integrating hydrodynamic equations which uses kernel interpolation to estimate fields, with widespread use especially in astrophysics. In contrast to common Eulerian grid methods which discretize the fields and hydrodynamic equations at fixed locations in space, Lagrangian methods integrate the hydrodynamic equations at discrete points which move with the flow. There are many good reviews on SPH [**????**], so here I will restrict myself to a very brief overview. In their simplest forms, the continuum hydrodynamic equations take

on the following forms:

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho(\nabla \cdot \mathbf{v}) \qquad \text{(continuity)} \tag{1.10}$$

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\frac{\nabla P}{\rho} + \mathbf{f} \qquad \text{(momentum)} \tag{1.11}$$

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -\frac{P}{\rho}(\nabla \cdot \mathbf{v}) + \Lambda \qquad \text{(energy)} \tag{1.12}$$

$\rho$ is the fluid density, $\mathbf{v}$ is the flow velocity, and $P$ is the fluid pressure. $\mathbf{f}$ is the total external force per mass acting on the fluid (the specific force) due to, e.g. gravity or artificial viscosity. Artificial viscosity terms are required by SPH for numerical stability and to handle shocks. $u$ is the internal energy and $\Lambda$ is the net heating/cooling, which may involve terms due to radiation, artificial viscosity, or other processes. The time derivatives are taken at points which follow the flow. Therefore, the convective derivative is used:

$$\frac{\mathrm{d}}{\mathrm{d}t} \equiv \frac{\partial}{\partial t} + (\mathbf{v} \cdot \nabla) \tag{1.13}$$

SPH "particles" move under the standard equations of motion with the fluid flow velocity $\mathbf{v}$. Each particle is assigned a mass. The fluid discretization begins by calculating the SPH density estimate as a weighted sum over nearby particles:

$$\rho(\mathbf{r}) = \sum_b m_b W(|\mathbf{r} - \mathbf{r}_b|, h) \tag{1.14}$$

where $m_b$ is the mass of neighboring particle $b$, $W$ is the *SPH kernel* which serves as a weighting function, and $h$ is the *smoothing length*, a length scale-factor. The kernel $W$ must be normalized, i.e. the integral of $W$ over all space must equal 1. There are many kernels which find use in practice, and the best choice of kernel is debated and depends on the environment being modeled. Typically, kernels should decrease monotonically with distance, be fairly smooth, be flat at zero separation, and should be computationally cheap (which rules out e.g. a gaussian) [**??**]. For illustration, the simplest kernel in common use is the $M_4$ cubic spline kernel, which is a piecewise polynomial of order 3 which is defined to be zero for $|\mathbf{r} - \mathbf{r}_b| > 2h$. Fig. 1.3 shows this kernel and its derivative, for particles along the

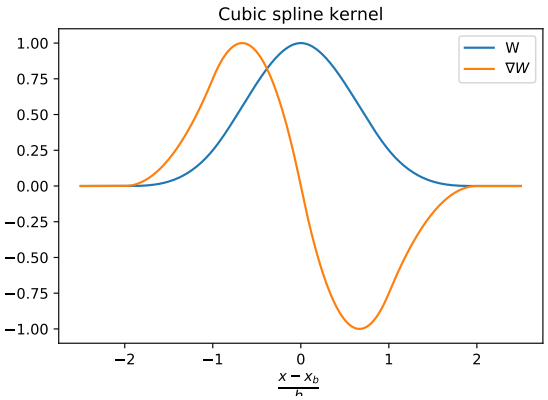



Figure 1.3: The $M_4$ cubic spline kernel and its derivative for particles along the $x$ axis, in arbitrary units.

$x$ axis. A standard method is to choose $h$ for a particle such that there are a fixed number of particles $N_{\text{smooth}}$ within a distance $2h$ of that particle. All hydrodynamic interactions are then calculated as interactions with a particle's $N_{\text{smooth}}$ nearest neighbors.

Fields are then estimated at the particle positions and can be interpolated, or smoothed, with a kernel weighted sum over neighbors. A general field or vector field can be interpolated at a given location as weighted sum:

$$\mathbf{A}(\mathbf{r}) \approx \sum_b \frac{m_b \mathbf{A}_b}{\rho_b} W(|\mathbf{r} - \mathbf{r}_b|, h) \tag{1.15}$$

which serves as the basis for estimating spatial derivatives as well. In general, there are many approximations to spatial derivatives which are used in the literature, but for illustration I

reproduce the following here:

$$\nabla \cdot \mathbf{A}(\mathbf{r}) \approx \sum_b \frac{m_b \mathbf{A}_b}{\rho_b} \cdot \nabla W(|\mathbf{r} - \mathbf{r}_b|, h) \tag{1.16}$$

$$\nabla \times \mathbf{A}(\mathbf{r}) \approx \sum_b \frac{m_b \mathbf{A}_b}{\rho_b} \times \nabla W(|\mathbf{r} - \mathbf{r}_b|, h) \tag{1.17}$$

$$\nabla A(\mathbf{r}_a) \approx \rho_a \sum_b m_b \left( \frac{A_a}{\rho_a^2} + \frac{A_b}{\rho_b^2} \right) \nabla W(|\mathbf{r}_a - \mathbf{r}_b|, h) \tag{1.18}$$

$$\tag{1.19}$$

Unlike with grid codes, finite differences are not used to estimate spatial derivatives. These estimates are functions only of fields at the particle locations and kernel gradients which are analytically known for standard kernels. Since kernels and their gradients are typically zero for $|\mathbf{r} - \mathbf{r}_b| > 2h_a$, these sums involve loops over the nearest $N_{\text{smooth}}$ neighbors. These SPH gradient estimates can then be plugged into the hydrodynamic equations to discretize them.

There are several main benefits to using SPH over an Eulerian method in the environment of a PPD, including:

- Conservation. Mass, momentum, angular momentum, and energy are all simultaneously conserved in SPH, which can be difficult to maintain with grid codes, especially in PPD environments.

- Resolution follows mass. Density and length scales in PPDs span orders of magnitude, and as structure such as spiral arms and clumps begin to form this, range increases rapidly and dynamically.

- Advection. Advection is handled exactly in SPH, which is particularly important in PPDs which have strong, non-uniform advection due to orbital velocities.

- Vacuum boundary conditions. Boundary conditions can be very difficult to handle for grid codes, in particular for the vacuum boundaries which are natural to PPDs. These are handled automatically in SPH.

There are also several main drawbacks to using SPH in a PPD simulation:

- Artificial viscosity (AV). This is perhaps the largest drawback to SPH in a PPD. AV is necessary in SPH to handle the shocks which naturally arise in PPDs which have highly supersonic orbital flow. In a highly shearing environment, SPH AV can generate strong numerical heating.

- Implementation. In general, it tends to be significantly more difficult to implement new physics in SPH codes, due to computational and algorithmic challenges inherent to SPH.

- Generating initial conditions (ICs). In general, it is more difficult to generate ICs for SPH, since the density is a derived parameter (eq. 1.14) and cannot be directly specified.

- Noise. SPH is prone to noise, especially in the estimate of spatial derivatives. This noise typically worsens with higher order derivatives.

- Magnetic fields. Magnetohydrodynamics is difficult to implement in SPH for many reasons, including that it is not easy to enforce $\nabla \cdot \mathbf{B} = 0$ in SPH or to handle diffusion with non-ideal MHD, although advances are being made [**??**].

### *1.7.3* ChaNGa *and supercomputing*

The challenges outlined above place strict resolution and time-stepping constraints on PPD simulations. In the simplest cases, PPDs typically require $N = 10^6$ particles, and in some cases may require as many as $N = 10^8$. Therefore, any $\mathcal{O}(N^2)$ algorithms must be immediately discarded. To achieve $\mathcal{O}(N \log N)$ scaling requires intelligent approaches to gravity and SPH, and to achieve decent wall-times requires sophisticated, highly parallel approaches.

The simulation code used and developed in part by the N-Body Shop at the University of Washington is called ChaNGa. It achieves good scaling through tree based SPH and

gravity. Physical space is hierarchically partitioned into sub-domains with approximately equal numbers of particles using the "oct-tree" decomposition. This tree can be built in $\mathcal{O}(N \log N)$ time and allows rapid look-up. Equal numbers of particles in sub-domains allow efficient load balancing. The gravitational potential is then calculated exactly for interactions with nearby particles and approximately for distant particles by using a multi-pole expansion. This reduces the complexity from the naive $\mathcal{O}(N^2)$ to the desired $N \log N$ scaling.

The same tree can be used to efficiently look-up particle neighbors, allowing rapid calculation of SPH sums. SPH interactions are local, so with the tree already built this calculation is $\mathcal{O}(N_{\mathrm{smooth}} N)$.

The resolution requirements of PPDs mean it is necessary to run in a massively parallel environment, with up to $10,000$ cores being used in some simulations. ChaNGa was developed originally for cosmological simulations and has been tested to scale well up to half a million cores [**?**]. To achieve this, ChaNGa is written in C++ and uses the Charm++ parallel system. The simulation is dynamically broken into TreePieces–small sections of the tree containing sub-groups of particles. Communication with other TreePieces is handled by Charm++ to allow interaction among TreePieces. Overlapping asynchronous computation of SPH and gravity is built into ChaNGa. Dynamic load balancing is required and implemented using Charm++ to distribute computation among the cores on a node and among the nodes in the network. Overlapping communication and computation are also managed dynamically, which is especially important in a supercomputing cluster where communication occurs between nodes on a cluster, processors in a node, cores within a processor, and sometimes virtual cores within a core, all with very different associated latencies and bandwidths. Developing in the Charm++ environment also allows ChaNGa to make efficient use of many different computing architectures, from personal laptops to massive supercomputing clusters.

### *1.8 Outline*

In this dissertation I present my doctoral research in computational modeling of protoplanetary disks. In Chapter 2 I discuss the role of initial conditions (ICs) in PPD simulations and present a fast method for generating equilibrium ICs for SPH simulations of PPDs. I explore direct formation of gas giants around M-Dwarfs via fragmentation under GI in Chapter 3. Chapter 4 presents my research on dust migration and concentration in PPDs. I investigate grain growth and self induced dust traps in Chapter 5. The results of my research and suggestions for future research are discussed in Chapter 6.

# Chapter 2

# INITIAL CONDITIONS

## *2.1 Overview*

Along with the numerical techniques and the governing equations, initial conditions define the system and the model for any simulation. Depending on the problem and the computational approach, choosing and generating ICs can be straightforward, or it can be as difficult and computationally costly as running the simulation itself. In the case of PPD simulations, while the choice of disk parameters and the methods for generating ICs have historically received little attention in the literature, they may be of primary importance in dictating the outcomes of PPD studies.

The general difficulties are twofold: (i) the exact properties of young PPDs are uncertain, both theoretically and observationally; (ii) the systems we are investigating are often very sensitive to their ICs due to their complexity and, particularly in the case of gravitational instability studies, their unstable nature. Improving our understanding of (i) the properties of young PPDs will require involved studies to make further progress. On the theoretical side, work has been done on modeling proto-star and PPD formation [**???**], but that is an entire field of research unto itself and well beyond the purview of my research. Current simulation techniques are unable to model the entire life time of a nebular cloud through the stages of star, proto-star, and PPD formation, due to the vast differences in scales and required physics. In particular, magnetic fields and radiation physics play an important role in nebular collapse and disk formation. Additionally, observations of PPDs remain very limited: very few high resolution images exist, and those are mainly of the dust component.

We therefore attempt to use explicitly simple models, making as few assumptions as possible. We choose to begin with disks as near to equilibrium as possible, which have

temperature and surface density profiles which decrease away from the star. Since the systems we are modeling are very sensitive to ICs, it is important that any deviations from equilibrium in the ICs are well understood and intentional. This introduces the second general difficulty.

Generating ICs close to equilibrium has been approached through many different methods. With grid codes, this is typically straightforward: all state/field quantities such as density, pressure, temperature, and velocity ($\rho$, $P$, $T$, $\mathbf{v}$) can be directly specified. All forces can then be calculated using the simulation code and quantities such as $\rho$ or $\mathbf{v}$ can be directly updated to reach equilibrium. In the case of PPDs, equilibrium can be achieved so precisely that even very unstable systems must be artificially perturbed to evolve away from equilibrium in any computationally feasible time-frame. This is often done by applying spiral arm structure (as tested in §2.3) or by applying small, random density perturbations.

Generating ICs with SPH is not so straightforward. Directly specifying $\rho$ is not possible with precision in SPH. Density is calculated from particle masses and positions, according to 1.14, and is therefore a derived quantity and not a specified one. One could generate particle positions, calculate density, adjust particle masses to approach the desired density, and iteratively repeat, hoping succesive density estimates would approach the desired value. However, for computational and algorithmic reasons, SPH implementations tend to behave much better when SPH particles all have identical masses.

Therefore, the desired $\rho(x, y, z)$ must be approximated by assigning particle locations. The simplest method for this is to treat $\rho$ as a probability distribution function (PDF) up to some normalization constant and use that to randomly generate particle positions drawn from the PDF. This will not exactly reproduce $\rho$, and in general vacuum boundries or regions of rapidly varying $\rho$ will be very poorly approximated since the density estimate (on average) will be a convolution of the SPH kernel with the PDF.

A bigger issue is that this method suffers from "poisson noise". For simplicity, consider a uniform density which we discretize using SPH by randomly drawing particle positions from a uniform PDF. Density estimates are taken over $N_{smooth}$ nearest neighbors, with $\rho$

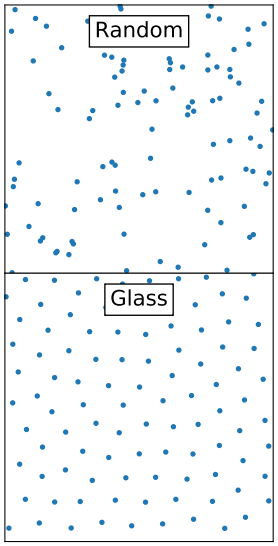
scaling approximately as the inverse of the volume enclosed by a sphere extending out to the furthest of the neighbors. For Poisson distributed particles, the standard deviation in this volume obeys $\sigma_V \propto \sqrt{N_{smooth}}$, giving us a density estimate which will have a standard deviation of:

$$\sigma_\rho = \frac{\rho}{\sqrt{N_{smooth}}} \tag{2.1}$$

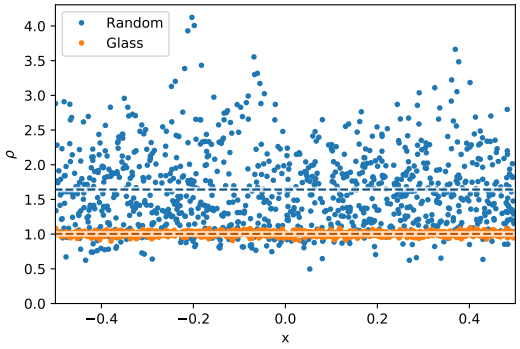
This can introduce very large, pseudo-random variations in the density estimate, which do not decrease in amplitude as resolution (i.e. total particle count) is increased. Poisson noise can be mitigated by increasing $N_{smooth}$, but this increases the computational cost. Although spatial resolution cannot be as precisely defined in SPH as it can be for grid codes, in general for a fixed spatial resolution $N/N_{smooth}$ = constant. Increasing $N_{\rm smooth}$ therefore forces $N$ to be increased rapidly as: $N \propto 1/\sigma_\rho^2$. Historically, SPH simulations were run with $N_{smooth} = 32$. Contemporary runs may use 64 or 128 neighbors, and some authors have argued for 200 neighbors or more [**?**]. However, getting Poisson noise to decrease to the 1% level would require $N_{smooth} = 10,000$.

The noise in density estimates will also be reflected in hydrodynamic force calculations, and even magnified since force estimates typically involve gradient estimates which may be very sensitive to noise. Additionally, poisson distributed particles will tend over-estimate the density. Particles are more likely to be in higher density regions than low density regions, since high density by definition requires more particles. Therefore, random poisson "clumps" tend to be over-sampled. See figure 2.1b for a demonstration of this.

Density over-estimation, coupled with random, "noisy" variations in $\rho$ mean a poisson distributed system will be in an excited energy state. Approaching equilibrium will tend to smooth out density variations. This leads naturally to the idea of an SPH "glass". In this context, a glass can be loosely defined as the minimum energy state which reproduces a desired density and which has no preferential ordering. This keeps particles as evenly spaced as possible with particle separations being randomly oriented. Glasses produce much better density estimates than poisson distributed particles and do not impose an artificial geometry as grids do. Figure 2.1 shows a comparison of random vs glass distributed particles.




(a) Random vs. glass particle placement, for a 2D periodic box. The glass was formed by running the random ICs with a damping force proportional to the velocity to settle into a glass-like state. Notice the "poisson clumping" in the random ICs. Particles in the glass state are nearly evenly spaced without using a grid. $N_{smooth} = 32$ was used.




(b) Random vs. glass density estimate for 1000 particles in a periodic cube with a target density of $\rho = 1$. The $\rho$ estimates for all particles are plotted as dots. The mean densities for particles for the two snapshots are plotted as dashed horizontal lines. The $\rho$ estimate for randomly placed particles is much noisier and biased significantly above $\rho = 1$. $N_{smooth} = 32$ was used.

Figure 2.1: Random vs. Glass

Grids produce very evenly spaced systems at the cost of introducing preferred directions, i.e. long range order. Such order can strongly influence the behavior of some systems, and is one of the reasons to choose an SPH approach over a grid discretization.

Stable SPH systems tend to evolve towards glass states naturally, especially in the presence of velocity damping forces. Unstable systems may not, however. Given the amount of energy that poisson noise deposits into a system, it is therefore very desirous to have ICs for unstable systems be placed in as glass-like a state as possible to prevent numerical effects from dominating the dynamics.

We therefore desire to generate ICs as close to equilibrium and as glass-like as possible. Getting a complex SPH system into equilibrium is not usually straightforward. For PPDs, estimating the required equilibrium $\rho$ involves calculating disk self gravity and pressure gradients, both of which depend on $\rho$ and for which there are no direct analytic solutions. Numerical effects such as artificial viscosity, force approximations, and SPH noise, all must be accounted for self consistently. SPH noise in particular means that force estimates must be averaged over many particles to estimate equilibrium state values.

In this chapter, I begin by describing my method for generating equilibrium PPD ICs (§2.2). While direct comparison with other authors' methods is not in general possible, since they are not publicly available and published research can be sparse on details with inaccuracies, I argue that my approach rapidly generates state of the art equilibrium ICs. In §2.3 I explore the dependence of simulation results on the state of unstable ICs in the context of fragmentation studies by perturbing unstable disks near equilibrium. I find that minor errors in the vertical density profile estimate, poor radial boundary conditions, or small density perturbations are all unlikely to lead to spurious fragmentation. Minor errors in velocity calculations can have a large impact and cause disks to fragment. In §2.4 I explore the effect of resolution on unstable PPDs. For these isothermal SPH simulations, decreasing resolution tends to drive these disks to fragment.

## *2.2 Generating ICs*

A major emphasis of this work was to ensure that ICs were as close to equilibrium as possible. Axisymmetric disks very near equilibrium may not realistically model actual PPDs, but we wish to make as few assumptions as possible about disks and have attempted to minimize numerical artifacts. One worry is that disks too far from equilibrium may artificially enter the non-linear regime, possibly initiating fragmentation in an otherwise stable disk. This possibility is explored in §2.3. In this section, I present my method for generating ICs.[1]

Many methods have been used in previous work for generating initial conditions. As discussed by **?**, apparently contradictory results in fragmentation studies may be due to differences in ICs used. Most published research does not detail IC generation in sufficient detail to be reproducible, but we can sketch out a few different approaches used. **?** generated ICs by defining the midplane density $\rho(R, z = 0)$, analytically estimating $\rho(z)$, using an approximate circular velocity, and iteratively adjusting the temperature profile to create a steady state solution.

Such an approach makes defining disk profiles difficult. As with us, some other authors [**??**] were able to define the surface density and temperature profiles. They then estimated vertical hydrostatic equilibrium to calculate density. **?** estimated the gas velocity required for circular orbits ($v_{circ}$) from gravitational forces and adjusted for hydrodynamic forces. As with others (e.g. **?**), they approached low $Q$ values by running the simulation and slowly growing the disk mass. Similarly, **?** used low mass, high-$Q$ models in his grid code simulations and accreted mass from the $z$ boundaries gradually to grow simulations towards instability.

**?** placed great care in developing equilibrium ICs. In contrast to our method, they also modeled the central, accreting star. They generated a stable disk ($Q = 1.8$) using a field equilibrium code [**?**]. They specified the specific angular momentum $j(R)$ of the gas (which forces velocity to be solely a function of radius) then iteratively used a self consistent

[1] My code for generating ICs is freely available on github at `https://github.com/ibackus/diskpy` as a part of my PPD python package DISKPY

field method to solve the Poisson gravity equation and balance the hydrodynamic forces to approach equilibrium. A shooting method for $j(R)$ was used to reach a desired $\Sigma(R)$. For low-$Q$ simulations, they ran the simulation and cooled the disk with an external cooling term until it reached $Q_{min} = 0.9$.

For our simulations, we desired to scan parameter space by defining the surface density ($\Sigma$) and temperature ($T$) radial profiles, along with the star mass. From these, the gas density ($\rho$) can be estimated to ensure vertical hydrostatic equilibrium in the disk. SPH particles are then semi-randomly seeded and their equilibrium circular velocities are estimated using the NBody/SPH simulation code ChaNGa to calculate the forces. Our method allows us to directly and quickly generate equilibrium ICs for low $Q$ values and arbitrary $\Sigma$ and $T$ profiles.

### *2.2.1 Estimating $\rho(R, z)$*

To estimate $\rho(R, z)$, we first select $M_*$, $\Sigma(R)$, and $T(R)$. Hydrostatic equilibrium is solved by adjusting the vertical density profile to maintain vertical hydrostatic equilibrium and adjusting the gas orbital velocity to ensure radial equilibrium.

To be in equilibrium along the vertical direction, the vertical component of gravity from the star and from the disk's self-gravity should balance the vertical pressure gradient in the gas. All quantities are axisymmetric and symmetric about the midplane $z = 0$. For the disk self-gravity term, we assume a thin disk approximation where $\rho$ is locally assumed to be only a function of $z$ (equivalently, this is the infinite sheet approximation). For most regions, the disk is slowly varying and this approximation is valid. The disk is much lighter than the star and in most regions the star's gravity dominates.

Under this approximation, the disk self-gravity force can be estimated by treating the disk as thin, infinite sheets of mass, each of which creates a uniform gravitational force along the $z$ axis given by:

$$f = 2\pi G\rho\Delta z \tag{2.2}$$

where $f$ is the specific force and $\Delta z$ is the sheet's thickness. The gravitational forces of all sheets at $|z| > |z_0|$ cancel out. In the limit $\Delta z \to 0$, the specific force due to disk self gravity as a function of $z$ is:

$$f_{disk}(z) = -2\pi G \int_{-z}^{z} \rho(z')\, dz' = 4\pi G \int_{0}^{z} \rho(z')\, dz' \tag{2.3}$$

$T$ is set to be independent of $z$, which is reasonable for the locally isothermal equation of state used in these simulations. Given the ideal gas equation, this gives the specific force due to vertical pressure gradients as:

$$\mathbf{f}_{pressure}(z) = -\frac{\nabla P}{\rho} = -\left(\frac{k_B T}{m}\right) \frac{1}{\rho(z)} \frac{\mathrm{d}\rho(z)}{\mathrm{d}z} \hat{z} \tag{2.4}$$

The specific force due to the central star's gravity is just the vertical component of the gravitational force:

$$\mathbf{f}_{star}(z) = -GM_* \frac{z}{(z^2 + R^2)^{3/2}} \hat{z} \tag{2.5}$$

Putting together eqs.(2.3)-(2.5) gives the condition for vertical hydrostatic equilibrium:

$$\frac{k_B T}{m} \frac{\mathrm{d}\rho}{\mathrm{d}z} + GM_* \frac{z\rho(z)}{(z^2 + R^2)^{3/2}} + 4\pi G \rho(z) \int_{0}^{z} \rho(z')\, dz' = 0 \tag{2.6}$$

where $m$ is the mean molecular weight of the gas, $M_*$ is the star's mass and $R$ is the cylindrical radius. The boundary conditions for this equation are:

$$\int_{0}^{\infty} \rho(z') dz' = \Sigma/2 \tag{2.7a}$$

$$\left.\frac{d\rho}{dz}\right|_{z=0} = 0 \tag{2.7b}$$

By physical reasoning, we can note that $\rho$ should also be maximum at $z = 0$, therefore we can also state:

$$\left.\frac{\mathrm{d}^2\rho}{\mathrm{d}z^2}\right|_{z=0} < 0 \tag{2.8}$$

Eq.(2.6) for vertical hydrostatic equilibrium can then be solved numerically, which turns out to be a non-trivial task. The boundary condition on $\frac{\mathrm{d}\rho}{\mathrm{d}z}$ is trivially satisfied. At $z = 0$ eq.(2.6) reduces to $\frac{k_B T}{m} \frac{\mathrm{d}\rho}{\mathrm{d}z} = 0$. A simple integration beginning at the boundaries at $z = 0$

or $|z| = \infty$ is also not possible. At $z = 0$ every term in eq.(2.6) is trivially zero, making it impossible to apply the boundary conditions. As $|z| \to \infty$, all 3 terms also trivially go to zero for a normalizable $\rho$. Shooting techniques were explored, where an initial guess $\rho(0)$ is used to integrate the equation. $\rho(0)$ is then scaled to attempt to satisfy eq.(2.7a). However, these techniques all displayed poor convergence for a large range of physical parameters. Ultimately, many techniques were attempted. Here I detail the two methods that worked. The latter method is currently the preferred one.

The first numerical method[2] was used robustly for the set of parameters explored in my initial work for young, gravitationally unstable PPDs around M-Dwarfs **?** (see chapter 3). Given that direct numerical integration is not possible, numerical root finding methods were investigated. Unfortunately, integro-differential equations can behave poorly with such methods, and performing the integral is computationally costly. Additionally, applying the boundary condition on the integral of $\rho$ is non-trivial, since this is a non-linear equation and re-scaling $\rho$ to satisfy eq.(2.7b) will no longer be a solution to the equation.

To apply the boundary conditions and avoid performing integrals, eq.(2.6) can be transformed to be a differential equation for the quantity $I(z) \equiv \int_z^\infty \rho(z')\, dz'$, which gives an equation of the form:

$$\frac{d^2 I}{dz^2} + \frac{dI}{dz}\left[\frac{c_1 z}{(z^2 + R^2)^{3/2}} + c_2\left(\frac{\Sigma}{2} - I\right)\right] = 0 \tag{2.9}$$

with the constants set by eq.(2.6). Given the definition of $I(z)$, $\rho$ can be calculated directly as:

$$\rho(z) = -\frac{\mathrm{d}I}{\mathrm{d}z} \tag{2.10}$$

and the boundary conditions are now all on $I(z)$ and its derivatives:

$$I(0) = \Sigma/2 \tag{2.11a}$$

$$\left.\frac{\mathrm{d}^2 I}{\mathrm{d}z^2}\right|_{z=0} = 0 \tag{2.11b}$$

[2] `ICgen` is my publicly available python package which implements this method for solving vertical hydrostatic equilibrium, along with generating PPD ICs. It is available at `https://github.com/ibackus/ICgen`. Note that `ICgen` is now deprecated in favor of `diskpy` (`https://github.com/ibackus/diskpy`)

By the definition of $I(z)$, as $|z| \to \infty$, $I(z) \to 0$ (along with its derivatives). This transformed eq.(2.9) still cannot be integrated from the boundaries trivially. Minimizing the norm of the residuals of eq.(2.9) was found to be ineffective—converging to a good solution is rarely achieved and it is even more rare to converge to the solution which satisfies all the boundary conditions. A root finding algorithm therefore is used to solve hydrostatic equilibrium as follows.

The dimensionality of the root finding problem is the number of data discretization points $N_z$ chosen along $z$. Typically, 1000 points were chosen from $z = 0$ to about 6 scale heights. This generates a very large space and therefore a good initial guess is necessary, both to ensure convergence to the proper solution and to ameliorate computational costs. An error function was used as an initial guess: $I_0(z) = \frac{\Sigma}{2}\left[1 - \text{erf}\left(\frac{z}{\sqrt{2}H}\right)\right]$. This is the exact solution in the limit of a massless, thin disk such that $|z| \ll R$. In that limit, the scale height is:

$$H = \sqrt{\left(\frac{R^3}{GM_*}\right)\left(\frac{k_B T}{m}\right)} \tag{2.12}$$

regardless of equation of state (EOS), so long as $T$ is locally independent of $z$. For an isothermal disk where $c_s^2 = \frac{k_B T}{m}$, this is identical to the typical approximation used in studying PPDs: $H = c_s/\Omega$.

An erf of the form above is then fit by minimizing L1 norm of the residuals, with the scale height $H$ as the free parameter. Residuals are estimated by using second order central finite differences to calculate the LHS of eq.(2.9). This solution is then used as the initial guess for the root finding algorithm. The roots of the residuals are found using a Newton-Krylov solver [**?**] as implemented in `scipy` [**?**]. Through much experimentation, it was found that this can often produce an inadequate solution to the equation governing $\rho$, but it can provide a good initial guess. From the solution to $I(z)$, the initial guess for $\rho$ is calculated from eq.(2.10). The same Newton-Krylov solver is used to find the roots of the residuals to eq.(2.6). To ensure that $\rho$ is properly normalized according to eq.(2.7a), the integral of $\rho$ is numerically calculated and $\rho$ is rescaled. This solution can be used as an initial guess and the root finding procedure can be repeated. The $\rho$ fitting procedure is then iteratively

repeated until numerical convergence is detected. For most of the PPDs around MDwarfs analyzed in my first study (see table 3.1), convergence was achieved after no more than several iterations.

There are two main draw-backs to this method: (i) it is slow and (ii) there are a large range of physical parameters of potential interest for which it does not converge. While the method is fairly slow, it can be easily parallelized by solving for $\rho(z)$ at several radii simultaneously. Even so, for a typical PPD, this method can take an hour to generate the full solution $\rho(R, z)$. This makes exploring parameter space difficult.

The second (ii) problem motivated developing another approach: solving the hydrostatic equilibrium equation iteratively. The iterative approach uses a generator function by rewriting an equation for $\rho$ as an equation of the form $\rho(z) = g(\rho(z), z)$. Approximations to the solution are then iteratively generated as:

$$\rho_n(z) = g(\rho_{n-1}(z), z) \tag{2.13}$$

Iterative approaches have been used by other authors [**?**] to calculate vertical hydrostatic equilibrium, although the details of their use are unclear and may differ significantly from the approach presented here.

If $\rho_n(z)$ approaches a fixed point, it is guaranteed to be a solution to the equation for $\rho$, since $\rho_n = \rho_{n-1}$ implies $\rho_n(z) = g(\rho_n(z), z)$ which is the full equation for $\rho$. However, $\rho$ is not guaranteed to approach a fixed point. It is important to rewrite the equation and choose a well behaved generator $g$, as well as a well behaved initial guess $\rho_0$. Several generators and initial guess were tried that did not converge to a fixed point for disk parameters of interest, however one was found. Given $\frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}z} = \frac{\mathrm{d}\ln\rho}{\mathrm{d}z}$, eq.(2.6) can be re-arranged as:

$$\frac{\mathrm{d}\ln\rho}{\mathrm{d}z} = -\frac{GM_*m}{k_BT}\frac{z}{(z^2+R^2)^{3/2}} - 4\pi\frac{Gm}{k_BT}\int_0^z \rho(z')dz' \tag{2.14}$$

which can be integrated from 0 to $z$ and rearranged to solve for $\rho$, which gives our generator

function:

$$\rho(z) = \rho(0)\exp\left[-\frac{GM_*m}{k_BT}\left(\frac{1}{R}-\frac{1}{\sqrt{R^2+z^2}}\right) - 4\pi\frac{Gm}{k_BT}\int_0^z\int_0^{z'}\rho(z'')dz''dz'\right] \tag{2.15}$$
$$= g(\rho(z), z)$$

The initial guess used is the solution for a thin disk without self-gravity, $\rho_0(z) = \frac{\Sigma}{\sqrt{2\pi}H}\exp(-z^2/2H^2)$ with $H$ given by eq.(2.12). The second integral of this can be evaluated analytically:

$$\int_0^z\int_0^{z'}\rho_0(z'')dz''dz' = z\frac{\Sigma}{2}\mathrm{erf}\left(z/\sqrt{2}H\right) + H^2\rho_0(z) - \frac{\Sigma H}{\sqrt{2\pi}} \tag{2.16}$$

Which can be placed into a discretized version $g$ (eq. 2.15) to generate the next estimate. Subsequent estimates of $\rho$ are calculated by plugging previous estimates into $g$ and normalizing to enforce eq.(2.7a):

$$\tilde{\rho}_n(z) = g(\rho_{n-1}(z), z) \tag{2.17}$$

$$\rho_n(z) = \left(\frac{\Sigma}{2}\right)\frac{\tilde{\rho}_n(z)}{\int_0^\infty \tilde{\rho}_n(z')dz'} \tag{2.18}$$

Note that this is still an iterative approach as the result of the two steps in eqs.(2.17) and (2.18) are together of the form of eq.(2.13). For all sets of physical parameters tested, this method converges rapidly to a fixed point solution. Typically, convergence to within double floating point precision is achieved in fewer than 20 iterations. Figure 2.2 shows an example of this convergence for a standard powerlaw profile 20 AU disk around and M-Dwarf at $R = 10$ AU.

The root finding method detailed above was initially preferred, because it displayed much better and more robust convergence than the iterative approaches initially attempted. Many generator functions of the form of eq.(2.13) can be easily derived from eq.(2.6), however many of them display very poor numerical convergence. The generator function chosen here typically behaves well, but for some parameters of interest it will not converge if only the gaussian initial guess is used, instead of the 1st order analytic estimate of eq.(2.16).

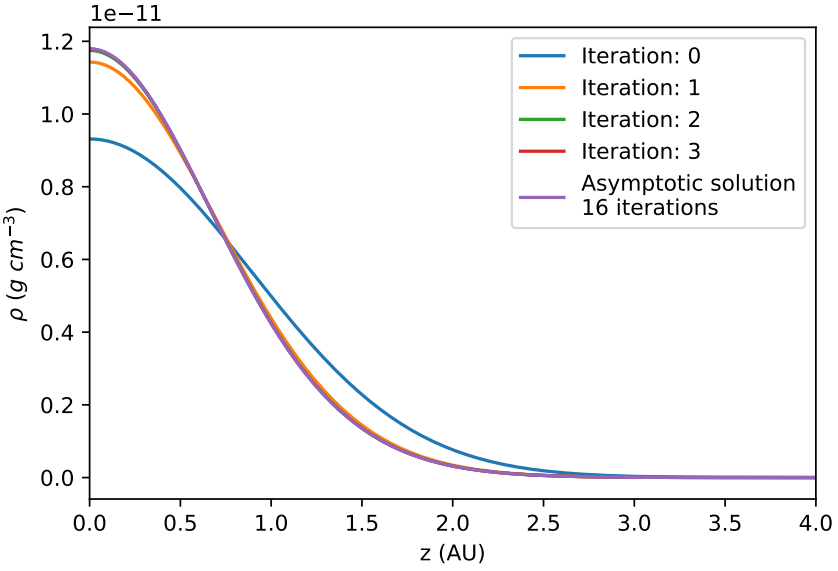



Figure 2.2: Iterative solutions $\rho_n(z)$ to the vertical hydrodynamic equilibrium equation (eq. 2.6), starting with initial guess $\rho_0(z) \propto \exp(-z^2/2H^2)$. After 3 iterations, the solution is nearly indistinguishable from the asymptotic solution. Convergence to within double floating point precision is achieved after only 16 iterations. This is for a standard powerlaw profile disk around an M-Dwarf at $R = 10$ AU.

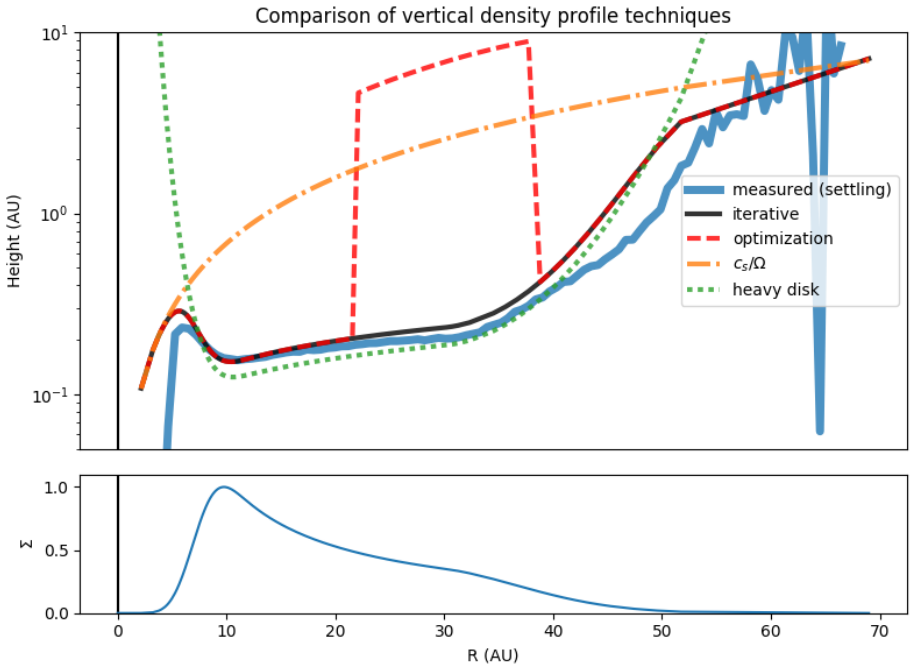



Figure 2.3: A comparison of methods for estimating vertical hydrostatic equilibrium. The bottom panel shows the surface density profile. The top panel shows the disk height for various methods of estimating $\rho(z)$. Disk height is defined here as the standard deviation of the vertical density profile.

Figure 2.3 compares the solutions for various methods of approximating disk vertical density profiles for an example heavy, unstable disk. Disk height (defined as the standard deviation of the vertical density profile) is plotted as a function of radius for an example disk around a solar mass star. The ICs are for a heavy (unstable) PPD around a star of mass $M_* = 1 M_\odot$ with a standard powerlaw temperature profile of $T \propto R^{-q}$ and a powerlaw surface density profile of $\Sigma \propto R^{-p}$, where $(q, p) = (0.59, 1)$ were used.

The heavy disk approximation assumes disk self-gravity dominates such that the second (star) term in the vertical hydrostatic equilibrium equation (eq. 2.6) is zero. Following **?**

(eq. 42), the solution becomes:

$$\begin{aligned} \rho(z) &= \rho_0 \,\mathrm{sech}(z\sqrt{2}/H) \\ \rho_0 &= \frac{\Sigma}{\sqrt{2}H} \\ H &= \frac{k_B T}{\sqrt{2}\pi G m \Sigma} \end{aligned} \tag{2.19}$$

The $c_s/\Omega$ curve in fig. 2.3 is the scale-height of eq. 2.12, ignoring disk self gravity and assuming $|z| \ll R$. The optimization curve is the solution using the root finding method described above. The "measured" curve was calculated by generating ICs according to the method presented here and then running those ICs in ChaNGa for several outer rotational periods (ORPs) defined at the disk radius $R_d$, with all forces not along the $z$ direction set to zero and with a damping force of the form $\vec{f}_{damp} = -\alpha\vec{v}$ with $\alpha > 0$ being a constant. This settling method guarantees ICs very close to vertical equilibrium, however it is very slow, taking of order a typical simulation walltime.

Clearly, the best approximation method for this disk is the iterative approach detailed above, which converges to a valid solution everywhere in the disk and is fast, allowing rapid generation of ICs. The root finding (optimization) method fails to converge to the correct solution between about 20 to 40 AU. For this heavy disk, the $H = c_s\Omega$ approximation overestimates the disk height except at the interior and exterior of the disk. The heavy disk approximation fairs well in the high density regions of the disk, but overestimates the disk height at the interior and exterior.

### *2.2.2 Generating particle positions*

The solution to $\rho(z, R)$ is then used to semi-randomly seed SPH particles. Particles are assigned a position in the $x-y$ plane and then a $z$ position. Rather than randomly assigning particle positions by drawing from a radial distribution defined by $\Sigma$ and a uniform angular distribution, to mitigate Poisson noise, we have implemented two methods:(i) the method of **?** which places the particles along a spiral in the $x - y$ plane to keep them more evenly spaced, and (ii) tiling and deforming a glass cube.

The first method, originally presented in **?**, is as follows. The spiral is made in such a way that $\Sigma(R)$ is reproduced. Following their method, the radial cumulative distribution function (CDF) can be calculated as:

$$f_R(R) = \frac{\int_0^R \Sigma 2\pi R' dR'}{\int_0^\infty \Sigma 2\pi R' dR'} \tag{2.20}$$

where the integrals are performed numerically. For $N$ particles, radial positions can be generated from the CDF inverse as: $R_i = f_R^{-1}(m_i)$ where $m_i$ are $N$ evenly spaced numbers between 0 and 1. The inverse is calculated numerically using spline interpolation. Particle spacings can be made approximately equal along the radial and azimuthal directions by increasing successive particle angles by:

$$\Delta\theta_i = -\sqrt{2\pi(1 - \frac{R_{i-1}}{R_i})} \tag{2.21}$$

Note that the sign of $\Delta\theta$ is important. The tight spirals should be trailing to avoid swing amplification and unwinding effects. For disks orbiting counter-clockwise when seen from above (i.e. in the $+\hat{\theta}$ direction), this requires $\Delta\theta < 0$. For a Keplerian disk, ignoring self-gravity and pressure, $\theta(t) = \Omega t$ where $\Omega = \sqrt{GM_*/R^3}$ is the Kepelerian angular velocity. A "kink" in spirals will develop when $\Delta\theta_i$ changes sign. This does not happen for trailing spirals, but does happen for leading spirals at a radius given by: $R(t) = (3t)^{1/2} (GM_*\Delta R/2\pi)^{1/4}$ where $\Delta R$ is the spacing of sequential particles at a radius $R$. Figure 2.4 shows a spiral "kink" development. Plotted are 1000 points generated radially according to a powerlaw profile $\Sigma \propto R^{-1}$ from $R = 0.1$ to 1. While gas pressure, artificial viscosity, and SPH noise will tend to disrupt the kink formation, such a well defined leading spiral arm can also be susceptible to swing amplification **?** which can drastically amplify the effect and drive a disk to instability.

For the $z$ positions, the vertical CDF at a given radius is calculated as:

$$f_z(z, R) = \frac{\int_0^z \rho(z', R) dz'}{\Sigma/2} \tag{2.22}$$

From which $z$ can be assigned randomly by $|z_i| = f_z^{-1}(m_i, R_i)$ where $m_i$ are $N$ random

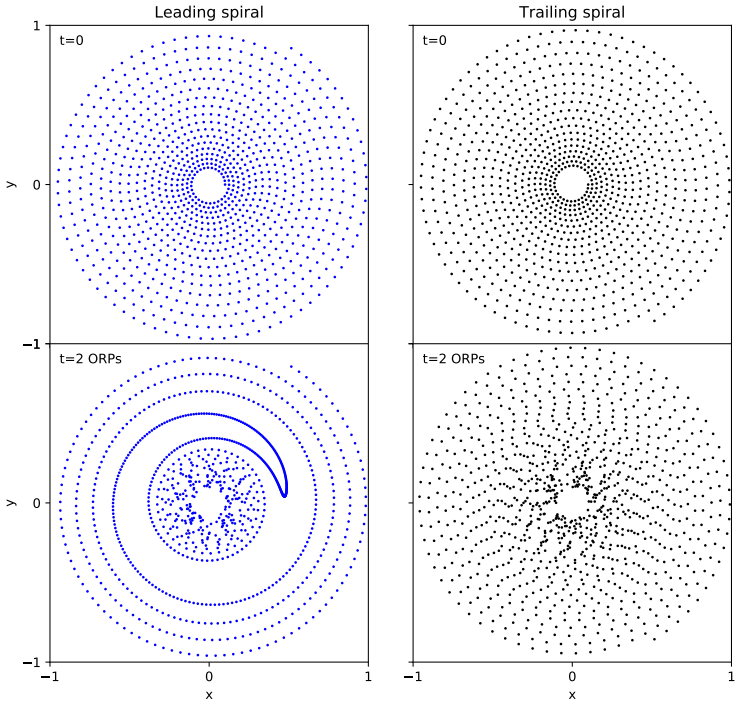



Figure 2.4: Spiral "kink" development. The left column is for leading spiral ICs ($\Delta\theta > 0$), the right column is for trailing spiral ICs ($\Delta\theta < 0$). The disk is orbiting in a counter-clockwise direction with angular velocity $\Omega \propto R^{-3/2}$. Plotted are 1000 points generated radially according to a powerlaw profile $\Sigma \propto R^{-1}$ from $R = 0.1$ to 1. A "kink" develops in the leading spiral which propagates outward as $R(t) \propto t^{1/2}(\Delta R)^{1/4}$ as the leading spiral unwinds and transitions to a trailing spiral.

numbers distributed uniformly on $[0,1]$ and the inverse of $f_z$ is taken with respect to $z$. The sign of $z$ is then randomly chosen.

The second method avoids using random numbers along $z$ by using a small glass cube in a periodic box as a starting point (with $N_c = 4096$ particles by default). Since the cube is periodic we can tile it to generate an arbitrarily large glass. Any structure created by this glassy lattice is minimal. We begin by generating a cylinder, in what I call the CDF space, of arbitrary radius $\tilde{r}_{max}$ with a height $2\tilde{z}_{max}$, centered on the origin with $|\tilde{z}| < \tilde{z}_{max}$, by tiling the cubic glass. The cube length is scaled to generate the proper number density according to: $\dfrac{N_c}{\tilde{L}^3} = \dfrac{N}{2\pi\tilde{r}_{max}\tilde{z}_{max}}$, where $N$ is the total number of particles in the simulation and $\tilde{L}$ is the cube length in CDF space. This can be done by tiling cubes in a region that includes the cylinder and removing particles outside the cylinder. This will generate semi-random, glassy positions, uniformly distributed in a cylinder CDF space. Positions can be mapped from CDF space to real space with the following mappings:

$$\begin{aligned} \theta &= \tilde{\theta} && \equiv g_\theta(\tilde{\theta}) \\ R &= f_R^{-1}\left((\tilde{r}/\tilde{r}_{max})^2\right) && \equiv g_R(\tilde{r}) \\ z &= f_z^{-1}\left(\tilde{z}/\tilde{z}_{max}, R\right) && \equiv g_z(\tilde{z}) \end{aligned} \tag{2.23}$$

where the square in the radial mapping is due to the fact that positions are uniformly distributed in $\tilde{x} - \tilde{y}$ but are distributed along $\tilde{r}$ with a PDF proportional to $\tilde{r}$.

With $\tilde{r}_{max}$ as an arbitrary parameter, we need to determine a reasonable value for $\tilde{z}_{max}$. The mass weighted mean of $H/R$ over the disk provides a decent value for $\tilde{z}_{max}/\tilde{r}_{max}$, but we have implemented slightly more sophisticated approach. The goal is to have the particles remain as glassy as possible when mapped to physical space. One way to achieve this is to attempt to keep the stretching along all 3 directions as uniform as possible. Stretching along the $\hat{\theta}$ direction will be equal to that along the $\hat{R}$ direction, which will differ from stretching along $\hat{z}$. Using $\Delta R = \frac{\mathrm{d}g_R}{\mathrm{d}\tilde{r}}\Delta\tilde{r}$ and $\Delta z = \frac{\mathrm{d}g_z}{\mathrm{d}\tilde{z}}\Delta\tilde{z}$. For $\Delta\tilde{z} = \Delta\tilde{r}$, our goal is to achieve a stretch factor $S \equiv \frac{\mathrm{d}g_R}{\mathrm{d}\tilde{r}}/\frac{\mathrm{d}g_z}{\mathrm{d}\tilde{z}}$ as close to unity as possible.

In principle, we could use a variable $\tilde{z}_{max}$ and remove particles with $|\tilde{z}| > \tilde{z}_{max}$ to get

$S \approx 1$. However, the radial mapping in eq. (2.23) assumes a distribution in CDF space of $P_{\tilde{r}} \propto \tilde{r}$. A variable $\tilde{z}_{max}$ would change the required mapping, which would in turn change the required $\tilde{z}_{max}$. This might be solvable iteratively, but using a fixed $\tilde{z}_{max}$ is much simpler. From the mappings in eq. (2.23) we get $S = \frac{2\tilde{z}_{max}}{\tilde{r}_{max}^2} \frac{P_z(z,R)\tilde{r}}{P_R(R)}$, where $P_z$ and $P_R$ are the PDFs for $z$ and $R$. Performing a mass weighted mean over the disk, $\langle S \rangle = \int\int SP_z(z,R)P_R(R)\, dz\, dR$. Solving for $\tilde{z}_{max}$ gives:

$$\tilde{z}_{max} = \frac{\tilde{r}_{max}}{2} \left[ \int_0^{\infty} \sqrt{f_R(R)} \int_{-\infty}^{+\infty} P_z^2(z,R) \mathrm{d}z\, \mathrm{d}R \right]^{-1} \tag{2.24}$$

where $P_z(z,R) = \rho(z,R)/\Sigma(R)$.

### *2.2.3 Circular velocity calculation*

A major concern with avoiding artifacts in the ICs is calculating the particle velocities required for circular orbits ($v_{circ}$). For typical disks, $v_{circ}$ is one to two orders of magnitude larger than $c_s$. The ratio of kinetic to thermal energies is proportional to $(v_{circ}/c_s)^2$ with an order unity constant of proportionality. Thus, particle velocities which deviate from $v_{circ}$ on the percent level will deposit a significant amount of power into the disk. The disk may enter a highly non-linear regime and fragment for unrealistically high $Q$. We include a demonstration of this in §2.3.3.

Figure 2.5 shows the ratio $(v_{circ}/c_s)^2$ as a function of radius for a typical disk. Everywhere in the disk, the kinetic energy is orders of magnitude larger than the thermal energy and increases with decreasing radius. For a temperature profile of $T \propto R^{-q}$ and assuming the keplerian circular velocity $v_{circ} = \sqrt{GM_*/R}$, we have: $(v_{circ}/c_s)^2 \propto 1/R^{1-q}$. For almost any disk, $q < 1$. The keplerian circular velocity approximation is typically valid within 10%. Therefore, for almost any disk, this ratio will grow towards the interior of the disk, meaning it is especially important to estimate the circular velocity well at the interior of the disk.

For a massless disk, pressureless disk $v_{circ} = \sqrt{GM_*/R}$ is the Kepelerian velocity. Pressure gradients change the net radial force and tend to decrease $v_{circ}$. Disk self gravity also

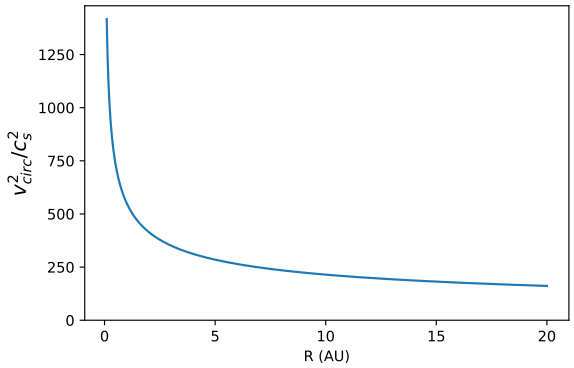



Figure 2.5: $v^2_{circ}/c^2_s$ for the fiducial M-Dwarf disk orbiting a $1/3M_\odot$ central star with a powerlaw temperature profile $T = T_0(R/R_0)^{-q}$ with $q = 0.59$. Throughout the disk, the available kinetic energy is orders of magnitude larger than the available thermal energy, varying as $\propto 1/R^{1-q}$.

changes $v_{circ}$ and is non-trivial to estimate. Numerical forces due to e.g. artificial viscosity can also be significant. To further complicate things, all these forces have $z$ dependence. Some studies use the keplerian velocity. Others attempt to adiabatically evolve the disk towards an equilibrium state by cooling the disk, growing the disk mass artificially, or by continuously loading mass onto the disk from the boundaries. However, such methods are slow, make precisely defining disk profiles difficult, and often do a poor job reaching an equilibrium state.

To calculate $v_{circ}$, we employ our simulation code ChaNGa to calculate the radial gravitational and hydrodynamical forces self-consistently on all particles. To deal with SPH noise, the forces must be averaged over many particles. The gravitational and SPH forces are averaged separately because of a different spatial dependence. In both cases, 50 radial bins are used.

For the gravitational forces, linearly spaced bins are used. To fit the $R$ and $z$ dependence of the gravitational forces, particles are binned radially, and for each bin a line is fit to the

radial force per mass due to gravity, as a function of $\cos\theta$:

$$f_{g,i}(\cos\theta) = m_i \cos\theta + b_i \tag{2.25}$$

where $\theta$ is the angle above the plane of the disk. The radial gravitational force from the central star is proportional to $\cos\theta$. The fit parameters $(m_i, b_i)$ are then linearly interpolated as a function of $R$ to create a model which allows us to estimate the radial gravitational specific force as

$$f_g(R, z) = m(R)\frac{R}{\sqrt{z^2 + R^2}} + b(R) \tag{2.26}$$

The radial hydrodynamic forces display very little $z$ dependence, as is expected for a temperature profile independent of $z$. The SPH forces are averaged over logarithmically spaced radial bins and then interpolated with a linear spline. From the total force per mass ($a = a_{grav} + a_{SPH}$) we can calculate $v_{circ}$ as:

$$v_{circ} = \sqrt{aR} \tag{2.27}$$

Following **?**, we set the gravitational softening length to $\epsilon_s = 0.5\,\langle h \rangle$, where $\langle h \rangle$ is the SPH smoothing length, calculated over the $N_{smooth}$ nearest neighbors (typically $N_{smooth} = 32$ or 128), and averaged over all particles in the simulation. The softening length $\epsilon_s$ for the star is by default set as the distance to the nearest gas particle in the ICs, although this can be safely decreased. ChaNGa is then used to estimate the gravitational and SPH forces separately.

### *2.3 Sensitivity to ICs*

It is important to understand the causes of disk fragmentation in our simulations. We have paid particular care to characterizing numerical issues which may drive fragmentation. In §2.4 we present a simple convergence test which demonstrates that for our locally isothermal SPH treatment, lower resolution disks fragment more easily.

To assess the sensitivity of disk fragmentation to the state of the ICs, we applied a series of small perturbations to disks close to the fragmentation boundary, with a $Q_{\rm eff}$ slightly greater

than 1. $Q_{\rm eff}$ is an effective Toomre $Q$ which includes a scale-height correction (see §3.3). Disks with $Q_{\rm eff} < 1$ are expected to fragment. We used two of the ICs in our suite of M-Dwarf runs which didn't fragment: one with $Q_{\rm eff} = 1.01$ and one with $Q_{\rm eff} = 1.12$ (simulations 2 and 5 in Table 3.1). We find that perturbing orbital velocities slightly out of equilibrium can cause a disk to fragment, but that results are fairly insensitive to perturbing the disk height or to applying small spiral density perturbations. These results are summarized in Table 2.1.

We also tested the importance of the inner and outer boundaries by taking a disk near the fragmentation boundary with $Q_{\rm eff}$ slightly above 1 and applying a step function cut-off to $\Sigma$ on the inner edge, the outer edge, and on both edges. Waves were seen to reflect off these hard boundaries, but the effect was not sufficiently strong to drive the disk to fragmentation. We tested runs 2, 10, 12, 39, 40, 45, 48, 49, 52 from table 3.1. Figure 2.6 shows the first two ORPs of run 2 ($Q_{\rm eff} = 1.01$) with a hard exterior cut applied. A strong wave propagates inward from the ICs, reflects, and travels outward. After several ORPs, the disk settles into a typical spiral arm dominated state. The wave is insufficient to cause the disk to fragment.

### *2.3.1 Density perturbations*

We applied $m = 2, 3, 4$ spiral density perturbations by multiplying the particle masses $M_0$ by:

$$M_1 = M_0 \left(1 + \delta \cos(m\theta)\right) \tag{2.28}$$

where $M_1$ is the perturbed particle mass and $\delta$ is the depth of the perturbation, here chosen to be $\delta = 0.01$. These perturbations are similar to those done by e.g. **?** to seed instabilities in a grid code which does not have SPH particle noise. These perturbations were performed on the $Q_{\rm eff} = 1.01$ simulation and were not sufficient to force the disk to fragment.

Figure 2.7 shows the results of the density perturbation runs. The midplane gas density is plotted on a logarithmic color scale for the original ("control", top-panel) and the spiral density perturbation runs. Time increases to the right. By 0.4 ORPs (first column), the spiral structure due to the perturbations is evident, with the number of arms corresponding

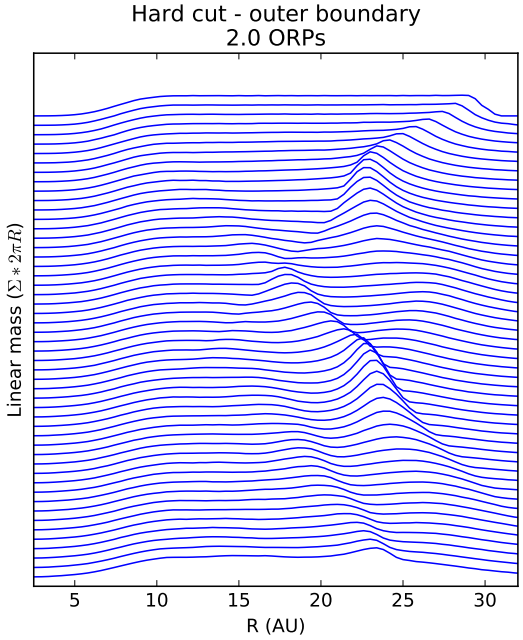



Figure 2.6: A stack of plots of the linear density ($\Sigma 2\pi R$) vs radius for a $Q_{\text{eff}} = 1.01$ disk with a step function cut-off applied to $\Sigma$ at the disk radius $R_d = 30$ AU. Lines are stacked for ease of viewing, earliest to latest from top to bottom. A strong wave propagates inward from the ICs, reflects, and travels outward. The wave is still insufficient to cause the disk to fragment. The first 2 ORPs are plotted.

to the order of the perturbation, $m$. Shortly after, the original mode no longer dominates. By 6.3 ORPs, the $m = 2$ perturbed disk shows the strongest spiral structure, but by the end of the simulation after 25 ORPs, all disks have approached a similar, non-fragmenting state.

### *2.3.2 Height perturbations*

We also wished to probe how sensitive disks are to perturbing the height. When generating ICs, there are many different methods for estimating the vertical density profile. We therefore wished to see whether perturbing a disk's scale height away from equilibrium could cause it to fragment. We ran the $Q_{\rm eff} = 1.01$ simulation twice, multiplying the particle $z$ positions by 0.98 and 0.9, decreasing the scale heights by 2% and 10%, respectively. Neither of these runs fragmented.

Figure 2.8 shows the results for these runs. The vertically integrated density is shown on a logarithmic color scale for the original ("control", top-panel) and the 2 height perturbation runs. The end results (rightmost column) are all similar. The disks have all approached a stable $Q_{\rm eff}$ value and do not fragment. The disk were run for 25 ORPs (not pictured). At 0.3 ORPs, oscillations in the height are visible in the 10% perturbation run as rings, but they are insufficient to cause the disk to fragment.

### *2.3.3 Velocity perturbations*

Disks do appear to be much more sensitive to particle velocities. We applied small, axisymmetric velocity perturbations to disks which otherwise did not fragment. The perturbed velocity $v_1$ was calculated as:

$$v_1 = v_0 \left(1 + \delta \frac{R_0}{R}\right) \tag{2.29}$$

where $v_0$ is the original velocity, $R_0$ is the inner edge of the disk (where the $\Sigma$ reaches a maximum) and $\delta << 1$ is the depth of the perturbation. This applies a fractional perturbation of $\delta$ at $R_0$ which decays as $1/R$. Nearly all the disk mass lies outside of $R_0$, meaning most of the disk has a perturbation applied which is much less than $\delta$. The general functional form

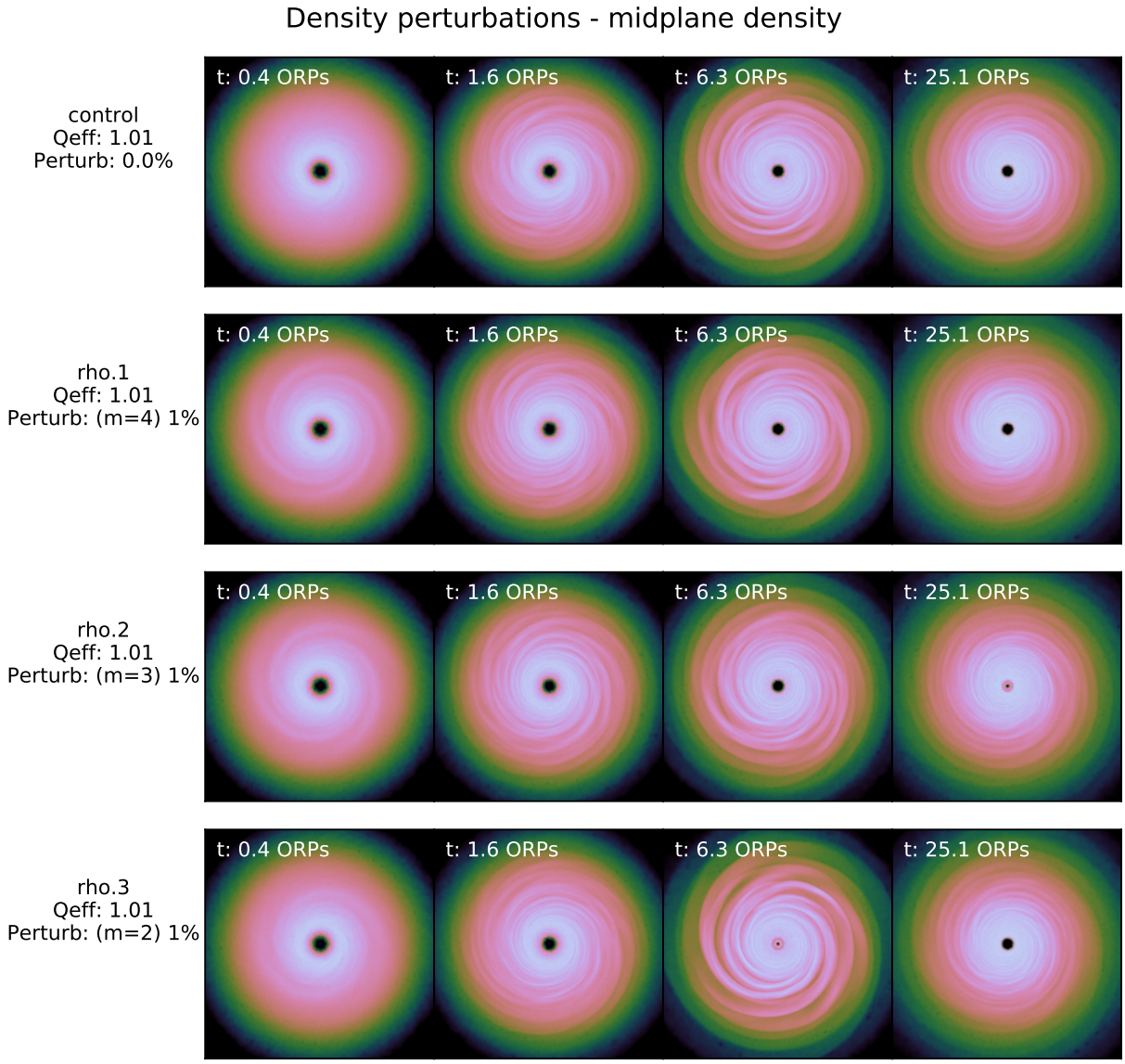



Figure 2.7: A comparison of simulations near the $Q_{eff} = 1$ boundary with density perturbations of the form $\Delta\rho \propto \cos m\theta$ (see eq. 2.28), listed in table 2.1. Pictured is the midplane gas density on a logarithmic color scale. Time increases to the right. The original disk ("control", top row) approaches a stable $Q_{eff}$ and does not fragment. The perturbations are insufficient to cause the disk to fragment.

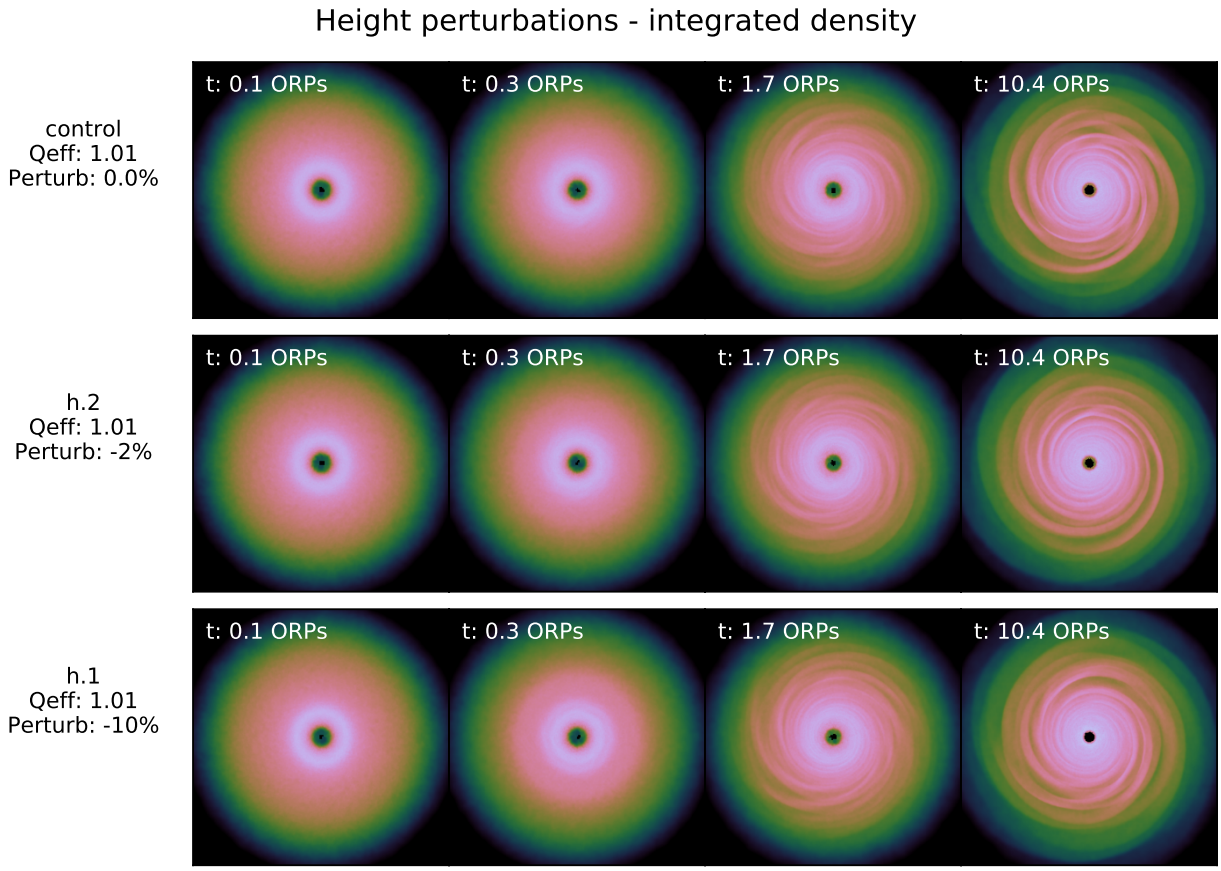



Figure 2.8: A comparison of simulations near the $Q_{eff} = 1$ boundary with height perturbations applied by scaling particle $z$ positions as $z' = (1 - \delta)z$, listed in Table 2.1. Pictured is the vertically integrated gas density on a logarithmic color scale. Time increases to the right. By 0.3 ORPs, oscillations in the height of the 10% perturbed disk are visible as rings. The original disk (control, top row) approaches a stable $Q_{eff}$ and does not fragment. The perturbations are insufficient to cause the disk to fragment. All disks were run longer (not pictured) and approached stable $Q_{eff}$ values.

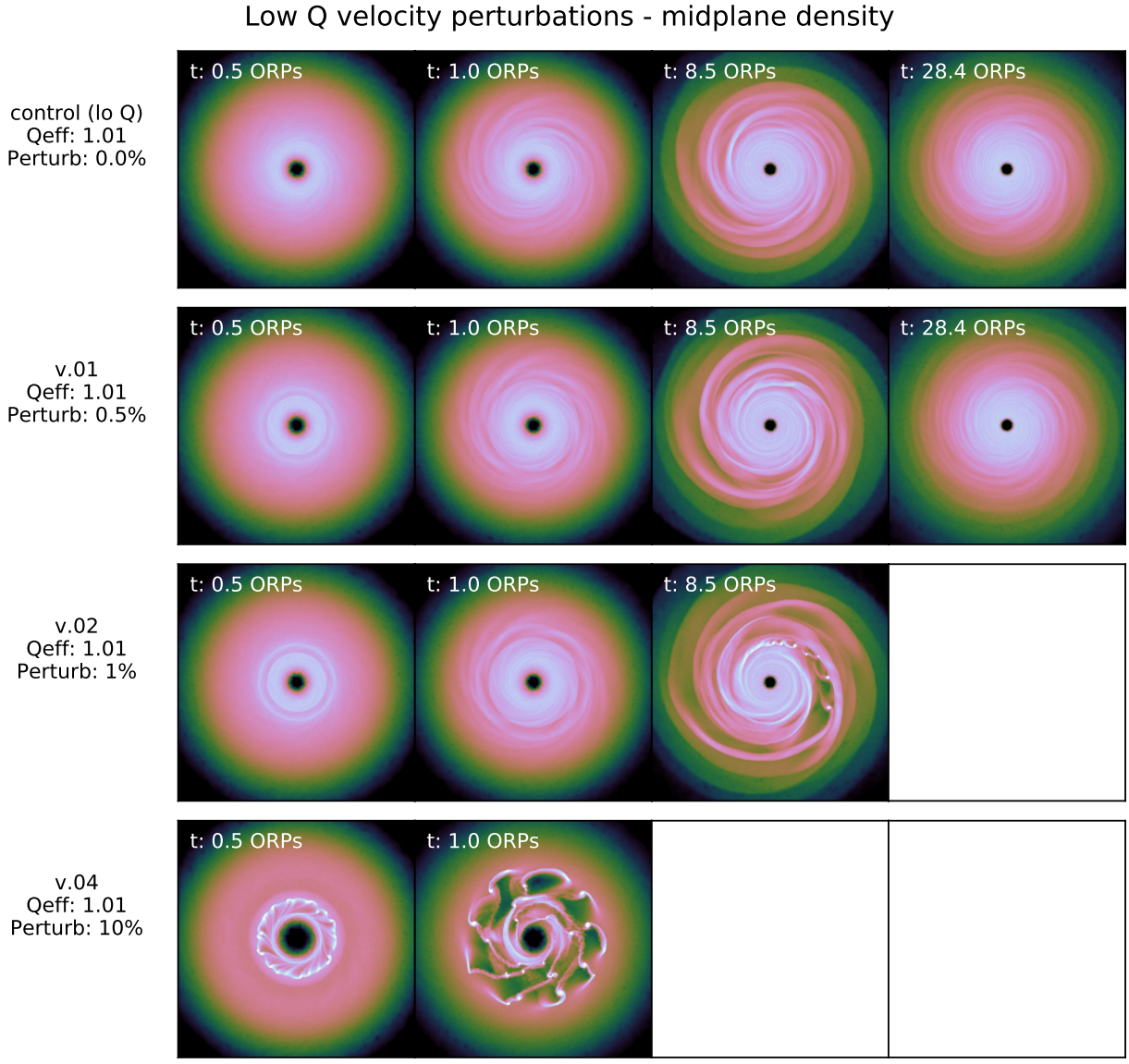



Figure 2.9: A comparison of simulations near the $Q_{eff} = 1$ boundary with velocity perturbations of the form of eq.(2.29), listed in table 2.1. These are the $Q_{eff} = 1.01$ runs. Pictured is the midplane gas density on a logarithmic color scale. Time increases to the right. The original disk (control, top row) approaches a stable $Q_{eff}$ and does not fragment. The 0.5% perturbation run also does not fragment, but all runs with larger perturbations fragment. Blank panels are for runs that were halted after fragmenting.

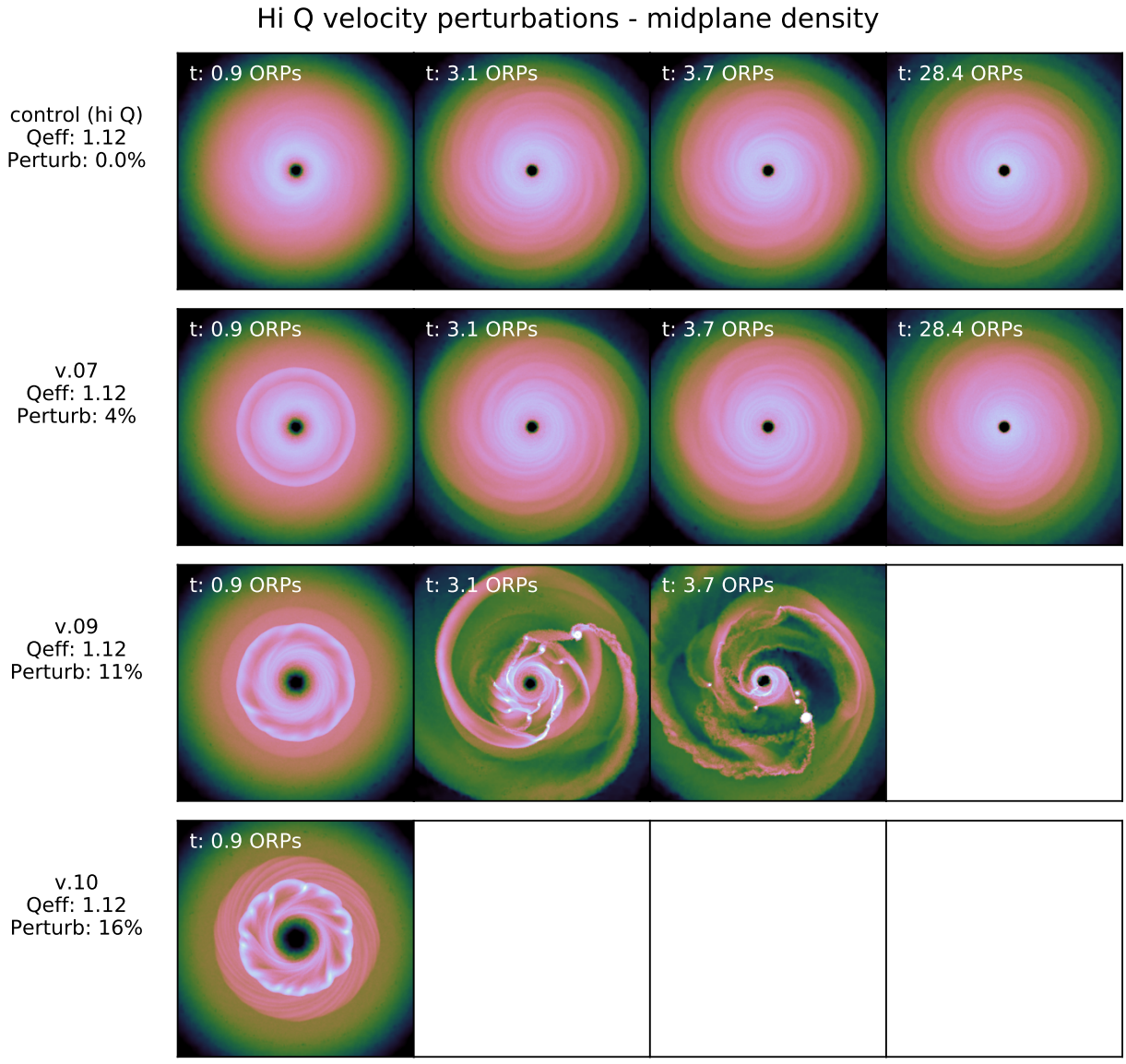



Figure 2.10: A comparison of simulations with $Q_{eff} = 1.12$ with velocity perturbations of the form of eq.(2.29), listed in table 2.1. Pictured is the midplane gas density on a logarithmic color scale. Time increases to the right. The original disk (control, top row) is at a stable $Q_{\rm eff}$ and does not fragment. The 4% perturbation run also does not fragment, but the 11% perturbation is sufficient to cause fragmentation. Blank panels are for runs that were halted after fragmenting. In particular, the 16% run fragments violently early in the simulation and was stopped after about 1 ORP.

was chosen to mimic behavior observed in many runs with ICs potentially out of equilibrium: a density wave spreading out from the inner disk. Such behavior indicates a disk with poorly estimated $v_{circ}$, and has been observed in early tests of mine and which I have observed in other tests.

For the $Q_{\rm eff} = 1.01$ disk, a 1% perturbation ($\delta = 0.01$) was sufficient to cause fragmentation. Figure 2.11 shows the spiral power as a function of time for this run. Spiral power is calculated by binning $\Sigma$ in $(R, \theta)$ and calculating the standard deviation. This is equivalent to summing the non-DC components of the fourier transform. The perturbed and original simulations initially develop in a similar manner for the first orbital period (defined at the most unstable radius). During this stage an axisymmetric ($m = 0$) density wave moves outward from the disk center. After about an orbital period, the perturbed disk develops significantly more pronounced spiral density waves. After 9-10 orbits the disk fragments. Fragmentation is accompanied by a rapid spike in spiral power as the disk becomes highly non-axisymmetric.

A similar test was performed with a much more stable disk ($Q_{\rm eff} = 1.12$). A series of perturbations were applied ($\delta = .01, .02, .04, .08, .11, .16$). An 11% perturbation was sufficient to force the disk to fragment. Depending on the approximations used to estimate circular velocities, discrepancies on this order can happen for disks sufficiently massive to be close to the fragmentation boundary. In particular, the $\Omega = \sqrt{GM_*/R^3}$ approximation can be wrong at the 10% level, especially for the massive disks of interest to gravitational instability studies.

Figures 2.9 and 2.10 show the results of these velocity perturbations for the $Q_{\rm eff} = 1.01$ and $Q_{\rm eff} = 1.12$ runs, respectively. The panels show the midplane gas density plotted on a logarithmic scale with time increasing to the right. Early in the runs, the $m = 0$ axisymmetric waves are clearly visible for the perturbed runs as rings. For sufficiently strong perturbations, these rings can seed strong spiral structure which ultimate leads the disk to fragment. Note that increased perturbations cause the disks to fragment more rapidly.

This demonstrates the care which must be taken in developing equilibrium models of

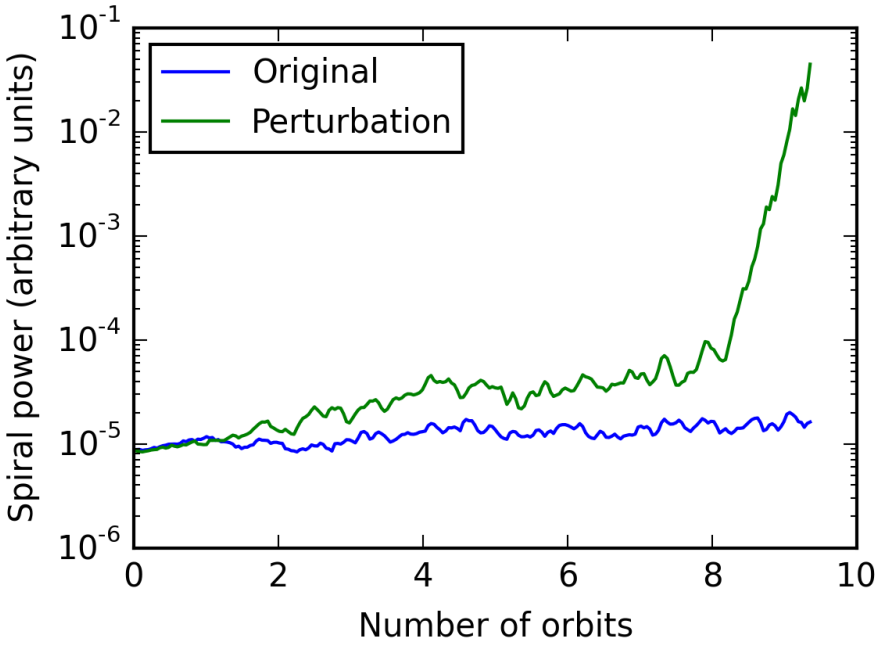



Figure 2.11: Total non-axisymmetric power versus time for a disk with and without a 1% velocity perturbation (simulation number 2). The original simulation ($Q_{\rm eff} = 1.01$) did not fragment but the simulation with a small, axisymmetric velocity perturbation does. Spiral structure initially develops similarly for both simulations for about the first ORP. For the next 8 ORPs the perturbed simulation develops deeper spiral structure until it fragments after 9-10 ORPs.

unstable disks near the fragmentation boundary, especially with regard to the velocity calculation. An apparently small perturbation can deposit large amounts of energy in a disk, forcing it sufficiently far of equilibrium to fragment.

### *2.4 Resolution test*

Previous work has indicated that the results of SPH simulations of PPDs can be resolution dependent [**???**]. **?** laid out several resolution requirements for SPH simulations of PPDs.

Table 2.1: Results for a series of tests to perturb disks near equilibrium.

| name | type | depth | fragment? | $Q_{eff}$ | notes |
|---|---|---|---|---|---|
| $\rho.1$ | density | 1% | no | 1.01 | m=4 |
| $\rho.2$ | density | 1% | no | 1.01 | m=3 |
| $\rho.3$ | density | 1% | no | 1.01 | m=2 |
| $h.1$ | height | -10% | no | 1.01 | |
| $h.2$ | height | -2% | no | 1.01 | |
| $v.01$ | velocity | 0.50% | no | 1.01 | |
| $v.02$ | velocity | 1% | yes | 1.01 | |
| $v.03$ | velocity | 2% | yes | 1.01 | |
| $v.04$ | velocity | 10% | yes | 1.01 | |
| $v.05$ | velocity | 1% | no | 1.12 | |
| $v.06$ | velocity | 2% | no | 1.12 | |
| $v.07$ | velocity | 4% | no | 1.12 | |
| $v.08$ | velocity | 8% | no | 1.12 | |
| $v.09$ | velocity | 11% | yes | 1.12 | |
| $v.10$ | velocity | 16% | yes | 1.12 | |

For our suite of $10^6$ particle M-Dwarf runs (§3), we exceed the mass resolution requirement by a minimum factor of 7, and most simulations exceed it by a factor of 40. We also easily meet their scale height resolution requirement. At the midplane of the most unstable disk radius, the ratio of the smoothing length to the scale height is between $4 - 12$ for all our runs (**?** finds this ratio should be at least $\sim 4$).

As part of our analysis, we ran a basic convergence test to verify our simulation code and to investigate the effects of resolution on disk fragmentation in SPH simulations. The resolution tests presented here are done without dust. Dust resolution requirements are discussed in Chapter 4. We ran a simulation close to the fragmentation boundary with a minimum $Q_{\rm eff} = 1.06$ (simulation 48 in table 3.1) at 6 different resolutions from 50k to 10M particles. These runs are summarized in Table 2.2.

We find (i) that low resolution simulations are more susceptible to disk fragmentation and (ii) that simulations appear to converge reasonably well, with our chosen resolution of $10^6$-particles being sufficient for the analysis presented in §3. However, these convergence tests are only preliminary and future work may reveal that higher particle count is required for fully believable results. Convergence may also depend on EOS, cooling prescriptions, and whether a grid code or an SPH code is used.

Figure 2.12 shows logarithmic surface density plots of all 6 runs after 4.0 ORPs. The 50k- and 100k-particle runs have already fragmented violently. The 500k-particle run has developed somewhat stronger spiral power than the higher resolution runs and eventually fragments after about 12 ORPs. The other runs have developed some spiral power that is insufficient to drive the disks to fragmentation. Figure 2.13 shows the logarithmic midplane density at several times for these runs.

Figure 2.14 shows what we call the normalized spiral power. We calculate spiral power by binning the surface density in $R, \theta$ (we used $128 \times 128$ bins here), calculating the standard deviation along the angular direction, and summing along the radial direction. We then normalize by multiplying the spiral power by $\sqrt{N}$, where $N$ is the number of SPH particles in the run. This is done to account for noise in the number of particles per bin which scales

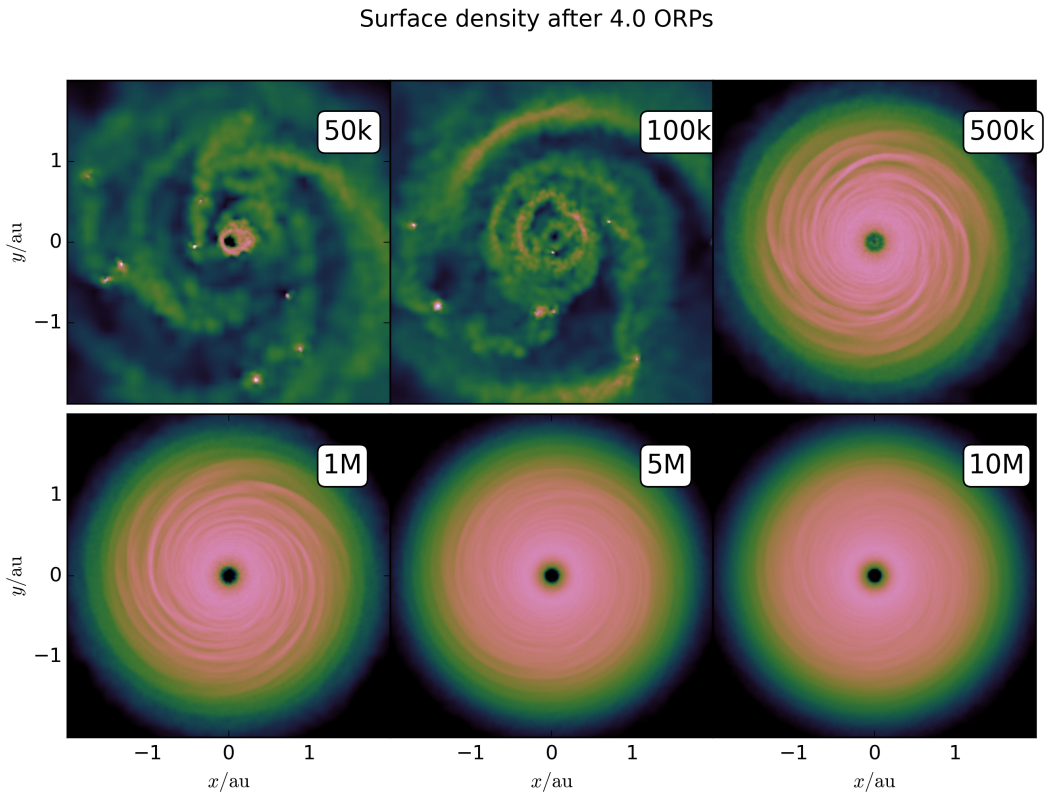



Figure 2.12: Logarithmic surface density plots for the simulations listed in Table 2.2 of a $Q_{\text{eff}} = 1.06$ run after 4.0 ORPs. The low resolution runs have already fragmented. The 500k-particle run has developed strong spiral power and will eventually fragment (see fig. 2.13). The remaining runs do not fragment and approach a value of $Q_{\text{eff}} > 1.1$ within the course of the simulations.

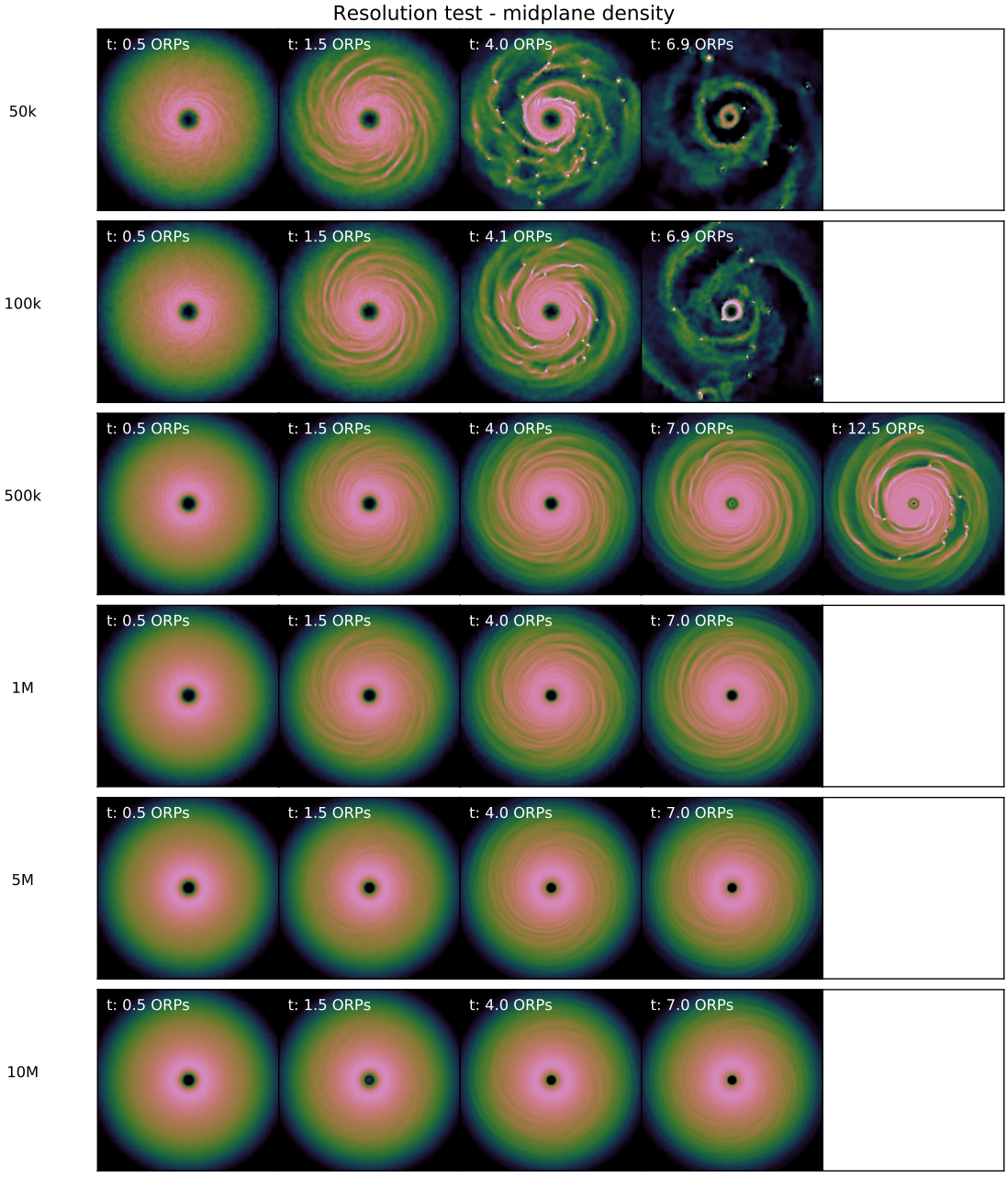



Figure 2.13: A comparison of the convergence test runs (Table 2.2) of a $Q_{\rm eff} = 1.06$ disk near the fragmentation boundary at 6 different resolutions. Pictured is the midplane gas density on a logarithmic color scale. Time increases to the right. Disks become asymptotically more stable with increasing resolution. The 50k, 100k, and 500k runs all fragment. Higher resolution runs stabilize as $Q_{\rm eff}$ increases. The 500k disk was run longer as it did not tend towards a stable $Q_{\rm eff}$. See also figure 2.12.

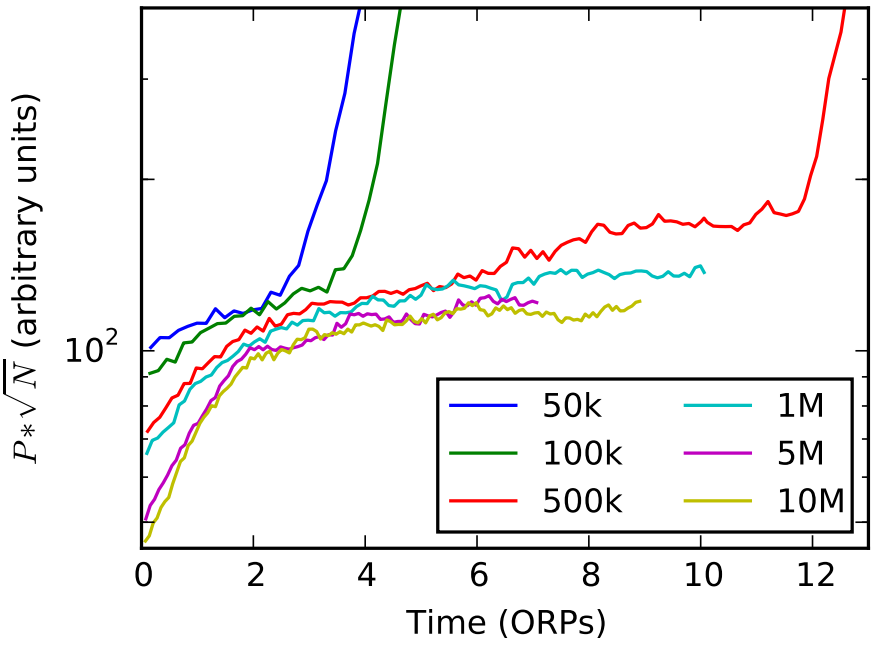



Figure 2.14: Normalized spiral power vs. time for the 6 simulations in the convergence test. Power is calculated by binning $\Sigma$ in $(R, \theta)$, calculating the standard deviation along $\theta$ and summing along $R$. Power is then normalized by multiplying by $\sqrt{N}$ to adjust for the power in particle noise. Fragmentation is seen for the simulations with 50k-, 100k-, and 500k-particles as a rapid increase in spiral power at the end of the simulation. As expected, the higher resolution simulations do not fragment.

as $\propto 1/\sqrt{n_{per\ bin}}$ where $n_{per\ bin}$ is approximately proportional to $N$.

The normalized spiral power then represents how much larger than the particle noise the spiral power is. Fragmentation is visible in figure 2.14 as a sharp rise in power at the end of the simulation. As can be seen, the normalized spiral power decreases with increasing particle count and converges for the higher-resolution simulations.

As expected for $Q_{\rm eff} = 1.06$, the higher resolution simulations do not fragment. As shown in Table 2.2, the simulations with $N \geq 10^6$ particles approach a stable value of $Q_{\rm eff} \approx 1.11$

Table 2.2: Convergence tests runs. ICs are identical to simulation 48 in table 3.1 but with a different number of particles. The resolution is the number of SPH particles in the run. Runs that fragment are highly non-axisymmetric, so the quoted $Q_{\rm eff}$ values are only illustrative. Runs that don't fragment approach a stable value significantly above 1.

| Run | Resolution | Fragment? | Final $Q_{eff}$ |
|---|---|---|---|
| 0 | 50k | yes | 0.67 |
| 1 | 100k | yes | 0.69 |
| 2 | 500k | yes | 0.89 |
| 3 | 1M | — | 1.11 |
| 4 | 5M | — | 1.11 |
| 5 | 10M | — | 1.12 |

and therefore would not fragment if run for longer. Simulations with $N \leq 500$k particles approached very low $Q_{\rm eff}$ minimum values, although $Q_{\rm eff}$ is strictly speaking not well defined for highly non-axisymmetric disks.

## *2.5 Conclusions*

Particularly in unstable systems, generating ICs is of the utmost importance, and in the field of PPDs it is often glossed over. Systems as chaotic as gravitationally active PPDs are very sensitive to ICs. Fortunately, we have demonstrated that SPH runs are not terribly sensitive to density estimates, in particular vertical density profiles or azimuthal perturbations. When disk heights are perturbed by 10%, even disks on the verge of fragmentation are not driven to fragment. However, PPDs are highly sensitive to velocity estimates, and percent level errors in estimating them may drive disks to fragment.

Here we presented a method which is robust and generates ICs very near equilibrium. My ICgen package is available on GitHub as a sub-package of my DISKPY python package.

Importantly, this method is rapid and independent of particle count beyond the $10^6$ or so maximum number of particles required for velocity estimates. Typical ICs take about 1 minute to generate on a modern desktop. Many methods in the literature involve difficult computations which may take of order the simulation walltime to generate, such as artificially growing the disk mass or cooling the disk.

We have also seen here just how important resolution can be in interpreting simulation results. Very few convergence tests have been reported in the literature, and very little has been published on the effect of resolution on SPH simulations of PPDs. We find that decreasing resolution in SPH can lead to artificially unstable and fragmenting disks. This is different for grid codes, which often suppress fragmentation at low resolution since the maximum density can be effectively limited by the grid size.

Much future work remains to be done on understanding the role of resolution in SPH. Some work has been done on characterizing artificial viscosity as a resolution dependent turbulence [**?**], but in general AV is resolution dependent and its effects are poorly characterized. If realistic SPH simulations extend beyond the isothermal approximation, resolution may play an entirely different role. The choice of $N_{\rm smooth}$, the choice of gravitational softening length, and the growth and nature of SPH noise remain poorly characterized in the literature.

# Chapter 3

# FRAGMENTATION OF PROTOPLANETARY DISCS AROUND M-DWARFS[1]

## *3.1 Introduction*

The importance of gravitational instabilities (GI) in the evolution of protoplanetary disks (PPDs) and in planet formation remains hotly debated [**?????**]. In recent years, the core accretion (CA) plus gas capture model of giant planet formation has received much attention [**??**], but GI is still seen as a candidate for the direct formation of giant planets, especially at large orbital radii [**?**]. While CA gives a more natural explanation of terrestrial planet formation and small bodies, GI may be important for the formation of these objects via solid enhancement within spiral arms or fragments [**?**]. GI may also play an important role during the embedded phase of star formation [**?**].

Understanding the role of GI in planet formation will require continued observation of PPDs [**???**] and further theoretical work. Of primary importance are (i) disk cooling times [**????**], which must be sufficiently short to allow density perturbations to grow against pressure support, and (ii) the Toomre $Q$ parameter [**?**]:

$$Q \equiv \frac{c_s \kappa}{\pi G \Sigma} \tag{3.1}$$

where $c_s = \sqrt{\gamma k_B T/m^2}$ is the gas sound speed, $\kappa$ is the epicyclic frequency ($\kappa = \Omega$ for a massless disk), and $\Sigma$ is the disk surface density. As $Q$ decreases toward unity, PPDs become increasingly unstable, and if $Q$ becomes sufficiently small, disks will undergo fragmentation.

The parameters required for fragmentation, such as disk mass ($M_d$), disk radius ($R_d$), and disk temperature ($T$), are constrained by the critical $Q$ required for fragmentation. Some

[1] This chapter has been published in **?**

[2] $\gamma \to 1$ for the isothermal sound speed

previous studies have found values of $Q_{crit} = 1.3 - 1.5$ [???], although it has been noted that $Q$ can drop below unity and the disk may still trend toward a self-regulating state [?].

Determining the parameters required for fragmentation is complicated by issues of resolution. The constant ($\beta$) cooling simulations of ? demonstrated non-convergence of SPH simulations. Further work [??] suggested artificial viscosity is to blame. Work is underway to investigate this problem; however, resolution dependent effects are still poorly understood in SPH simulations of PPDs [?].

Previous work has tended to focus on PPDs around solar mass stars. Motivated by the large population of low mass stars, we study GI around M-dwarfs with mass $M_* = M_\odot/3$. Around 10% of known exoplanets are around M-dwarfs [?]. Due to selection effects of current surveys such as Kepler [?], this is expected to be a large underestimate of the actual population. Recent discoveries show that disks around M and brown dwarfs are different from those around solar analogs: the mass distribution falls of more slowly with radius, and is denser at the midplane. These differences change disk chemistry and the condensation sequence. M-dwarf disks are also less massive and survive longer [??]. Core accretion timescales, which scale as the orbital period, are long around M-dwarfs. Because the stars are much lower in luminosity, their disks are substantially cooler. Additionally, planets orbiting nearby M-dwarfs are likely to be the first smaller planets spectroscopically characterized [?].

The simulations of ? indicate that GI is able to form gas giants around M-dwarfs. ? and ? even argue that super earths around M-dwarfs can be explained as gas giants, formed via GI, and stripped of their gaseous envelopes by photoevaporation.

In this chapter I explore the conditions required for disk fragmentation under GI around M-dwarfs. Previous studies have found a range of values of the $Q_{crit}$ required for disk fragmentation [????]. Discrepancies may be due to different equations of state (EOS), cooling algorithms, numerical issues such as artificial viscosity prescriptions, Eulerian vs. Lagrangian codes, and initial conditions (ICs). ICs close to equilibrium are non-trivial to produce and so in Chapter 2 I explored the dependence on ICs of simulations of gravitationally unstable

disks.

Here, I focus on probing disk fragmentation around M-Dwarfs, which remains poorly studied. This warrants a simple, well understood isothermal EOS. We therefore probe the Toomre $Q$ required for fragmentation while leaving the question of the cooling required for fragmentation for future work.

§3.2 describes the suite of simulations presented here and discusses the theoretical and observational motivations behind our disk profiles. §3.3 presents our analysis of disk fragmentation around M-dwarfs. §3.4 presents our method for finding and tracking gravitationally bound clumps and discusses clump formation in our simulations. We present our discussion in §3.5. We consider the effects of thermodynamics and ICs on fragmentation in PPD simulations and argue that GI should play an important role in PPDs around M-dwarfs and that we expect disk fragmentation at large radii to occur around many M-dwarfs.

## *3.2 Set of runs*

### *3.2.1 Disk Profiles*

Figures 3.1 & 3.2 show radial profiles for example simulations. Pictured are the surface density $\Sigma(R)$ (top panel), temperature $T(R)$ (middle panel), and $Q(R)$ (bottom panel). The exact disk structure of PPDs, especially young ones, is poorly constrained. Therefore we adopt simple, easy to interpret profiles, and consider a range of values of disk mass ($M_d$) and temperature ($T$) in order to bracket plausible disk parameters. Below we describe our choice of disk temperature and surface density profiles, along with the theoretical and observational motivations for them.

*Temperature*

For a blackbody disk with heating dominated by solar radiation, **?** showed that the temperature profile should be a power law. The exponent is $q = 3/7$ for a fully flared disk and $q = 3/4$ for a flat disk. We therefore adopt a temperature profile of the form:

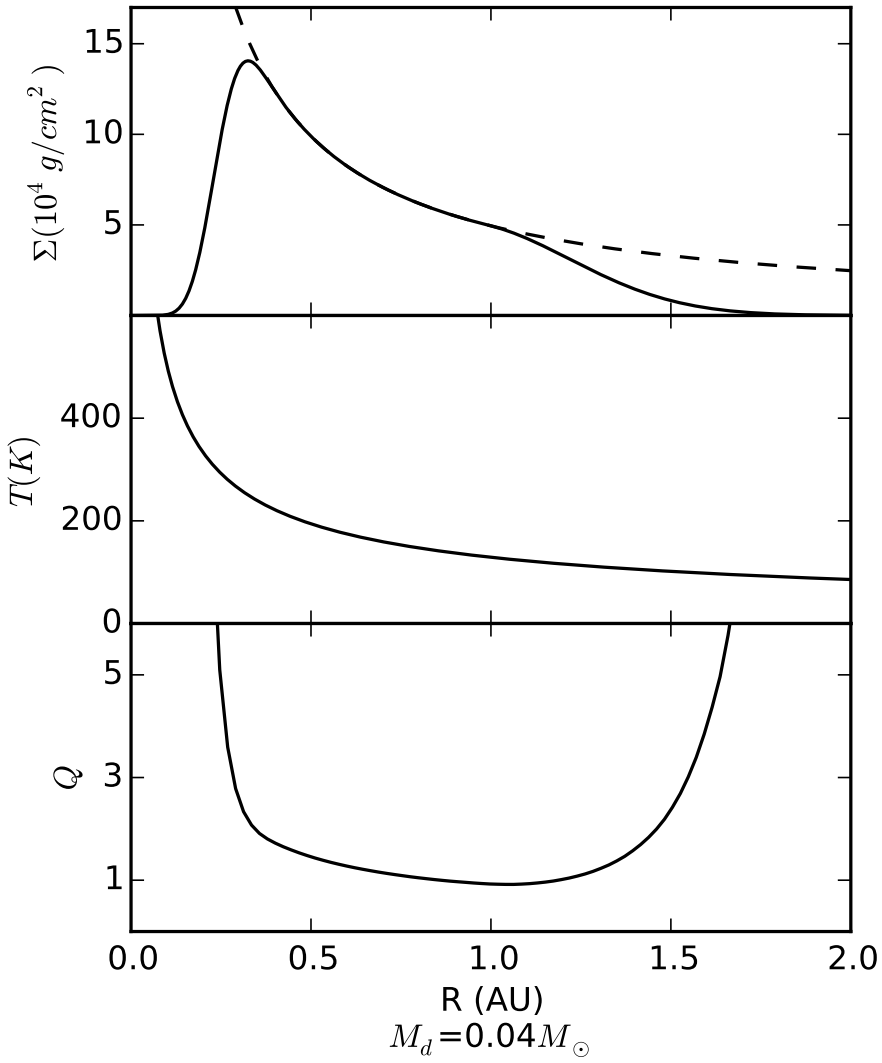



Figure 3.1: Example radial profiles for powerlaw surface density $\Sigma \propto 1/R$, with $R_{in} = 0.3$ AU and $R_d = 1$ AU. **Top:** Surface density profile including cutoffs (solid) and excluding cutoffs (dashed). **Middle:** Disk temperature $T \propto R^{-0.59}$ **Bottom:** Toomre $Q$, calculated including full disk self gravity, SPH forces, and calculation of $\kappa$.

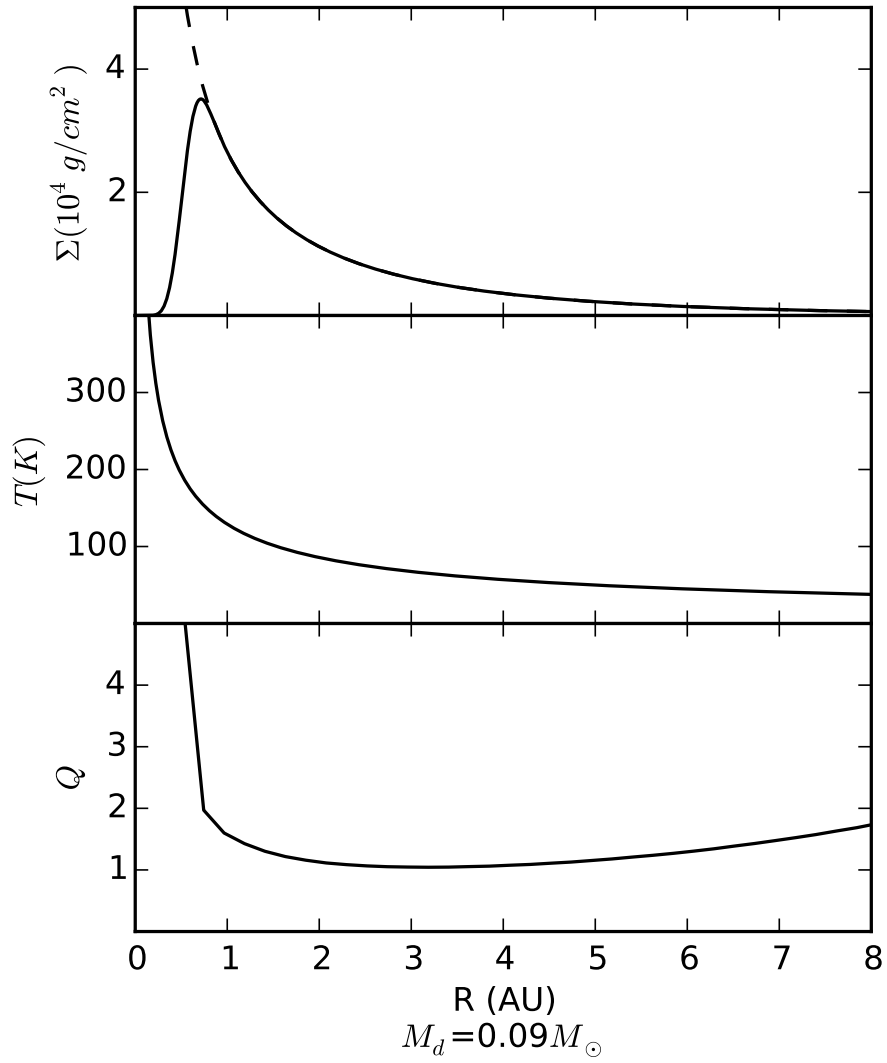



Figure 3.2: Example radial profiles for a viscous disk surface density $\Sigma(r) = \Sigma_0 r^{-\gamma} \exp(-r^{2-\gamma})$, where $r$ is a dimensionless radius and $\gamma = 0.9$. The radius containing 95% of the mass is $R_d = 11$ AU. **Top:** Surface density profile including cutoffs (solid) and excluding cutoffs (dashed). **Middle:** Disk temperature $T \propto R^{-0.59}$ **Bottom:** Toomre $Q$.

$$T(R) = T_0 \left(\frac{R}{R_0}\right)^{-q} \tag{3.2}$$

where $R_0$ was set to be 1 AU and $T_0$ is the temperature at 1 AU. The values we adopt for $T_0$ and $q$ come from the observations of **?**. They observed dust SEDs of circumstellar disks in the Taurus-Auriga star forming region. Fit results for 44 mainly solar type stars yielded median inferred parameters of $T_0 = 148K$ and $q = 0.58$ (see their Table 2). Averaging their results for M-Stars only, we adopt $q = 0.59$ for every simulation and $T_0 = 130K$ as our central fiducial values. This power law of $q = 0.59$ lies between a fully flared and completely flat disk. Given the uncertainty of these values we also ran simulations with $T_0 = 65K$ and $260K$ to bracket plausible disk temperatures.

*Surface density*

Two functional forms for $\Sigma$ were used: a power law (Fig. 3.1) and the similarity solution for a thin, viscous disk (Fig. 3.2). The power law used was:

$$\Sigma(R) = \Sigma_0 \left(\frac{R}{R_0}\right)^{-1} \tag{3.3}$$

Where the normalization $\Sigma_0$ is fixed by the desired disk mass. As shown in Fig. 3.1, interior and exterior cutoffs were applied. For $R > R_d$, an exponential cutoff was applied by multiplying $\Sigma(R)$ in eq.(3.3) by:

$$\Sigma_{exterior} = \Sigma(R) \mathrm{e}^{-(R-R_d)^2/L^2} \tag{3.4}$$

where the cutoff length was set to $L = 0.3R_d$. This form ensures that $\Sigma$ and $\frac{d\Sigma}{dR}$ are unchanged at $R = R_d$. The interior cutoff was applied by multiplying $\Sigma$ by a smooth high order polynomial approximation to a step function, defined to be $[0, 1]$ at $R = [0, R_{cut}]$ with the first 10 derivatives set to be 0 at $R = [0, R_{cut}]$. For these simulations $R_{cut} = 0.5R_d$ and $\Sigma$ differs significantly from a power law for $R \lesssim 0.3R_d$ (see Fig.3.1). This radius was chosen such that: (a) $Q \gg 1$ at $R_{cut}$ to ensure the disk is stable at $R_{cut}$, and (b) the cutoff is

applied far enough from the most unstable disk region and removes little enough mass that fragmentation should not be strongly affected by the cutoff.

The second functional form for $\Sigma$ (see Fig.3.2) comes from the similarity solution to a thin, light, viscous disk orbiting a star, as found in **?**. For a viscosity obeying a power law $\nu \propto R^{\gamma}$, at a given time $\Sigma$ can be written as:

$$\Sigma(r) = \Sigma_0 r^{-\gamma} \exp(-r^{2-\gamma}) \tag{3.5}$$

where $r$ is a dimensionless radius and the normalization $\Sigma_0$ is fixed by the disk mass. Note that the full similarity solution includes a time dependence which we fold into $\Sigma_0$ and $r$. From fits to observations of 9 circumstellar disks, **?** found a median value of $\gamma = 0.9$, which is the value we adopt here. For this profile, no exterior cutoff is required. The same interior cutoff as for the power law $\Sigma$ was applied at $r = 0.1$.

If we define $R_d$ to be the radius containing 95% of the disk mass (ignoring the interior cutoff), then:

$$r = \frac{R}{R_d} \ln(1/0.05)^{1/(2-\gamma)} \tag{3.6}$$

As has been noted before (e.g. **?**), there is no physical motivation for adopting a power law for $\Sigma$. It is also not clear how applicable the viscous profile is. These are just simple functional forms often adopted in previous work. We have examined the end state of stable high-$Q$ runs and they are better approximated by the viscous profile, but the fit is not perfect.

### *3.2.2 Run Parameters*

For a given functional forms of $\Sigma(R)$ and $T(R)$, three parameters must be set to define our ICs and thereby fix a value of $Q_{min}$: the temperature normalization ($T_0$), the disk mass ($M_d$), and the disk radius ($R_d$). The choices of $T_0$ are discussed in §3.2.1.

*Disk mass*

Under our scheme, setting $M_d$ fixes the surface density normalization. We selected plausible values to explore, ranging from 0.01 to $0.08 M_\odot$. **?** reported observations of 11 disks around pre-main-sequence stars, including 7 around M-stars. From their Tables 1 & 2, we find a median value of $M_d/M_* = 0.15$, which for our simulations ($M_* = 1/3 M_\odot$) gives a central value of $M_d = 0.05 M_\odot$. It should be stressed that these disk masses are inferred using an assumed gas to dust ratio of 100 and therefore may have large, uncharacterized uncertainties.

*Disk radius*

**?** argue that for a typical disk, $R_d$ increases from around 20 AU to 100 AU over the course of $\sim$5 Myr. This tends to stabilize disks by decreasing $\Sigma$ over time. Since we are interested in disks at their most unstable, we adopt 20 AU as our central fiducial value. To explore parameter space, we used values of $R_d$ ranging from 1/3 AU to 30 AU.

*Numerical Parameters*

For the analysis of fragmentation criteria, a set of 64 ICs were run with $10^6$ particles. For typical disks, this yields particles with a mass of $m_{particle} \approx 5 \times 10^{-5} M_{Jupiter}$. A locally isothermal EOS with a mean molecular weight of 2 was used for the gas. Following **?**, we set the gravitational softening length to $\epsilon_s = 0.5 \langle h \rangle$, where $\langle h \rangle$ is the SPH smoothing length calculated over the 32 nearest neighbors and averaged over all particles. Typical values are around $\langle h \rangle = 5 \times 10^{-3} R_d$.

The central star (mass $M_\odot/3$) was set to be a sink particle: when a gas particle approaches the star within a distance of $R_{sink}$, its mass and momentum are accreted onto the star. $R_{sink}$ was set as the distance to the closest gas particle in the ICs. This yielded $R_{sink} = [0.08 R_d, 0.02 R_d]$ for the power law and viscous $\Sigma$ profiles, respectively

Making the central star a sink serves two purposes worth mentioning. A gas particle near the star gains a large velocity, experiences strong forces, and is often captured in a tight

orbit around the star. This requires a very small time-step which can increase computation time by orders of magnitude. Secondly, the adaptive time-stepping used by ChaNGa can fail to conserve momentum when two interacting particles require time-steps which differ by more than a couple orders of magnitude. The effect is sufficiently strong that very stable disks ($Q > 2$) were seen to fragment when the central star was not treated as a sink.

We note that treating the star as a sink in this manner is unrealistic in that it forces accretion to happen at very large radii (of order an AU). When accretion happens, the star will jump to the center of mass of the star + particle system. In the limit of low accretion rates and in locations far from the star this will be a negligible effect, but in disks with high accretion a different scheme should be used. For our simulations there is very little accretion: less than $10^{-3} M_d$ over the duration of the simulations.

As shown in figure 3.5, non-fragmenting simulations were run for $\sim$30 outer rotational periods (ORP), where we define 1 ORP as the orbital period at the most unstable disk radius ($R_d$ in the case of the powerlaw surface density profiles). ORPs for our disks range from 0.3 yrs for our smallest disks to 300 yrs for the largest. Disks with parameters close to fragmentation were run longer (100–200 ORP) to ensure they reached a steady state which would not fragment. Since the main goal of this work is to investigate fragmentation of PPDs and since computation time increases drastically after clump formation, simulations which fragmented were run for around 1–2 ORP after fragmentation. Figure 3.5 shows the fragmentation timescales for these disks.

### *3.3 Fragmentation analysis*

The primary goal of this analysis is to investigate under what conditions we can expect a PPD surrounding an M-dwarf to fragment. The parameters explored by our model are $T$, $M_d$, and $\Sigma$. Although all simulations were run with a star of $M_* = M_\odot/3$, we also extend our analysis to stars of similar mass. As expected, we find that sufficiently heavy or cold disks will fragment under GI.

Gravitational instability in PPDs is typically parameterized by the Toomre $Q$ parameter

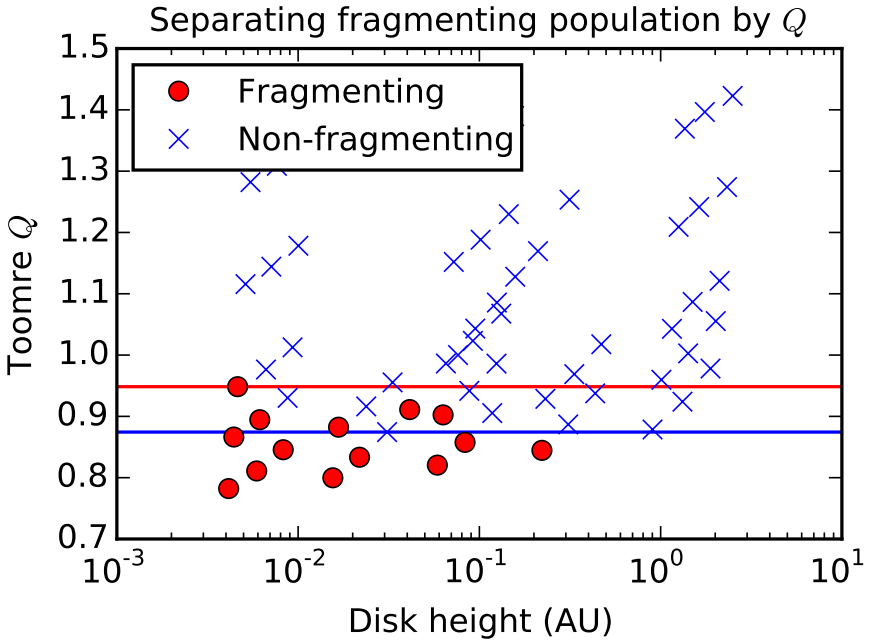



Figure 3.3: Minimum Toomre $Q$ for the fragmenting (clump–forming) and non-fragmenting simulations. The red (blue) lines mark the largest (smallest) values of the fragmenting (non-fragmenting) simulations. The two populations overlap around $Q \approx 0.9$.

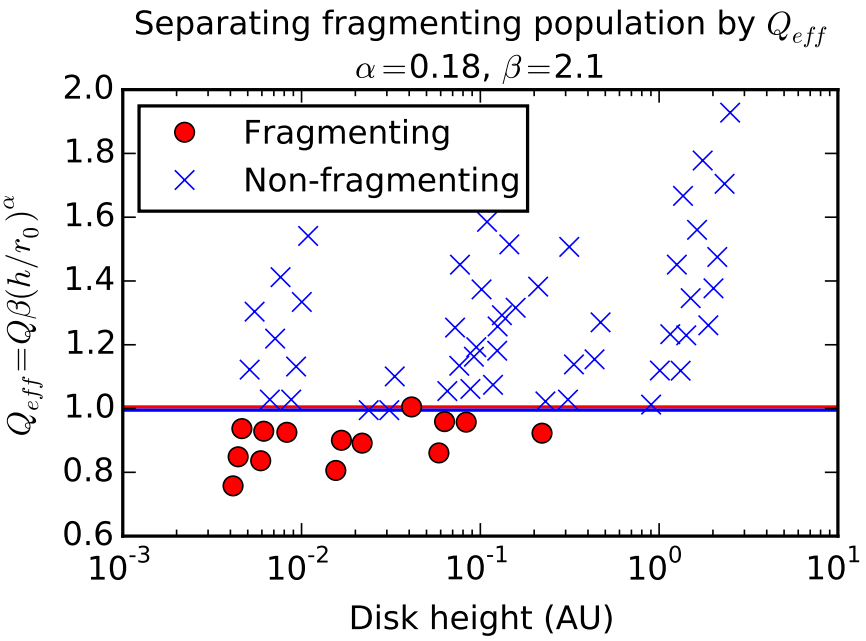



Figure 3.4: Minimum effective Toomre $Q$ ($Q_{\rm eff}$) for the fragmenting (clump–forming) and non-fragmenting simulations. Re–parameterizing the stability criterion for a protoplanetary disk as $Q_{\rm eff} = Q\beta(H/R)^{\alpha}$ is sufficient for predicting whether a protoplanetary disk will fragment. $\beta$ is a normalization factor chosen such that disks with a $Q_{\rm eff} < 1$ will fragment.

(eq. 3.1). For our isothermal simulations, $c_s = \sqrt{k_B T/m}$. Two dimensional disks will be unstable for $Q > 1$, but the instability and fragmentation criteria for 3D disks remain uncertain. Previous studies have found that 3D disks will fragment for $Q_{min}$ significantly greater than 1 [**????**]; however, we do not find this to be the case for our simulations. Figure 3.3 shows the fragmentation boundary for our simulations. Disks with $Q \lesssim 0.9$ fragment. Here, we define fragmentation to be when the first gravitationally bound clump is found (see §3.4).

We also considered the effects of swing amplification. If the parameter:

$$X_m = \frac{\kappa^2 R}{2\pi G\Sigma}\frac{1}{m} \tag{3.7}$$

for integer $m \geq 1$ is near or below unity, small leading disturbances will amplify upon becoming trailing disturbances and can drive disk dynamics, especially spiral arm growth [**?**]. Lower order modes are expected to dominate. For our disks, our lowest value is $X_1 = 5.7$, and most disks have $X_1 > 10$, so swing amplification should not be significant in these simulations.

Another proposed instability in gaseous disks is provided by the SLING mechanism [**??**], which is driven by $m = 1$ mode growth and is sensitive to the outer edge of the disk. For our simulations, we see little $m = 1$ power. Additionally, we tested the importance of the outer edge by applying a step-function cutoff to $\Sigma$ on the disk outer edge for marginally stable disks and the overall behavior was unaltered. It therefore seems unlikely that the SLING mechanism plays a major role in these simulations.

It should be noted that for our simulations we calculate $Q$ from the ICs. Since the equilibrium orbital velocity is known (see §2.2.3), we can directly calculate $\kappa^2 = \frac{2\Omega}{R}\frac{d}{dR}(R^2\Omega)$. This fully includes the effects of a 3D disk with self-gravity and pressure gradients. $\Sigma$ is calculated by binning SPH particles radially, summing their masses, and dividing by the annulus area. A common approximation for light disks is to ignore disk self gravity and pressure gradients and use $Q \approx c_s\Omega/\pi G\Sigma$, which for disks of $M_d/M_* \approx 0.1$ underestimates $Q$ at the 10% level. Different estimates of $Q$ may account for some of the discrepancy in

the literature regarding the critcal $Q$ required for fragmentation.

As can be seen in figure 3.3, there is some overlap in $Q$ for the fragmenting and non-fragmenting populations. Since the Toomre stability criterion strictly applies to a 2D disk, this is unsurprising. Following the work of **?**, to higher order a disk scale height correction enters into the dispersion relation for axisymmetric perturbations. Taller disks should be more stable than thinner ones.

We re-parameterized $Q$ to include disk height as:

$$Q_{\rm eff} \equiv Q\beta(H/R)^{\alpha} \tag{3.8}$$

where $H$ is the disk scale height and $\alpha$ is a free parameter. $\beta$ is normalization parameter that we set such that $Q_{\rm eff} < 1$ is the boundary for disk fragmentation (see below). $H$ is calculated as the standard deviation of the vertical density profile, rather than the first order approximation $c_s/\Omega$. We then fit the power law $\alpha$ to minimize the overlap of the boundaries in $Q_{\rm eff}$ for the fragmenting/non-fragmenting population. Figure 3.4 shows the separation of the two populations (compare to fig. 3.3). For $\alpha = 0.18$ and $\beta = 2.1$, simulations with $Q_{\rm eff} < 1$ fragment.

A power law was chosen because it is a simple functional form, and there is a great deal of self-similarity in PPDs, but other forms such as linear corrections might be suitable as well. As expected, taller disks have a larger $Q_{eff}$ and are therefore less prone to fragmentation, although the dependence on height is weak. A simple linear correction can be derived by Taylor expanding and demanding $Q_{\rm eff}(H = 0) = Q$. Doing so and defining $h \equiv H/R$ gives:

$$Q'_{\rm eff} = Q\left[1 + \frac{\alpha}{1-\alpha}\left(\frac{h}{h_0}\right)\right] \tag{3.9}$$

where $h_0 = [\beta(1-\alpha)]^{-1/\alpha} = 0.043$, which is a very reasonable value for typical disk aspect ratios.

We also found that $Q_{\rm eff}$ correlates strongly with time until fragmentation (see fig. 3.5). We found it to predict fragmentation time with less scatter than $Q$. To verify that $H/R$ is an important parameter in predicting disk fragmentation, we considered power law dependence

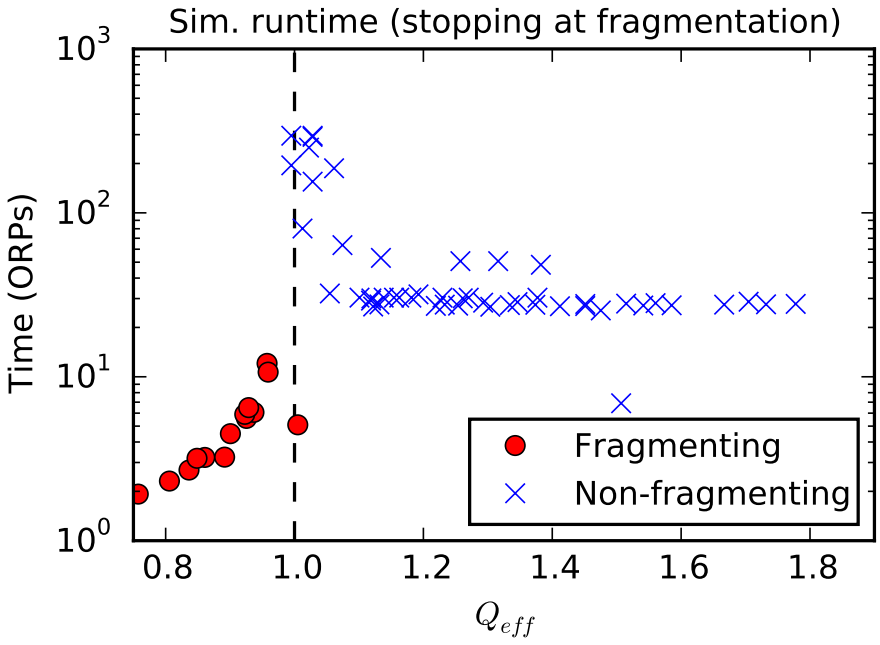



Figure 3.5: Total simulation time (non-fragmenting simulations) and time until fragmentation (fragmenting simulations), in units of the orbital period at the most unstable disk radius. Fragmentation is defined to occur when a gravitationally bound clump forms. The fragmentation timescale increases rapidly as $Q_{\rm eff}$ approaches 1. Simulations with $Q_{\rm eff} \gtrsim 1$ were run for longer to verify that they do not fragment.

of $Q_{\rm eff}$ on various dimensionless combinations of parameters, including the most unstable wavelength, $c_s$, $T$, and $\Omega$. We found $H/R$ to separate the populations most strongly.

These considerations indicate disk scale height is an important parameter in dictating stability and fragmentation. With the fragmentation boundary $Q_{\rm eff} = 1$, we are equipped to estimate disk parameters for which we may reasonably expect disks to fragmentation.

Figure 3.6 shows the boundaries for disk fragmentation for a star of mass $M_* = M_\odot/3$ as a function of $R_d$, $M_d$, and the temperature at 1 AU ($T_0$). The contour lines mark the boundary for various values of $T_0$. Disks to the right of the contour lines have a minimum

$Q_{\rm eff} < 1$ and will fragment.

The red boundary marks the fiducial value of $T_0 = 130K$ from the observations of **?** (see §3.2.1), with the surrounding red region marking the sample scatter in their observations of $25K$. The red point marks the fiducial values (for a young disk) of $M_d$ and $R_d$ from the observations of **?** (see §3.2.1).

Since the fiducial disk parameters lie to the left of the fiducial boundary, we expect the observed disks to not be susceptible to fragmentation by GI. This is to be expected, since the timescales for fragmentation are so much shorter than observed disk lifetimes/ages, it is unlikely to observe gravitationally highly unstable PPDs. However, we note that observed disk parameters lie close to the region of fragmentation.

### *3.4 Clumps*

The formation of tightly bound, dense clumps of gas marks the stage of disk evolution where our isothermal treatment begins to break down. Clumps can collapse rapidly, much faster than cooling timescales, and therefore are better approximated as adiabatic than isothermal. Since isothermal clumps do not heat up, they will be less pressure supported and tend to collapse more easily. We therefore restrict our analysis of clumps to the early stages of formation and accordingly limit the scope of our results.

To track the formation of clumps, we developed a simple clump finding/tracking software routine[3] built around the group finder SKID[4]. To find gravitationally bound clumps, we may factor in the disk geometry. Particle masses are scaled by $R^3$ and a density threshold is set such that at least $N$ particles lie within the Hill-sphere of particles under consideration, where $N = N_{\rm smooth}$ is chosen as the number of neighbors used for SPH smoothing. This gives a threshold of

$$\rho_{min} = \frac{3N_{\rm smooth}M_*}{R^3} \tag{3.10}$$

[3]Our clump finding code is freely available on github at `https://github.com/ibackus/diskpy` as a part of our PPD python package *diskpy*

[4]SKID is freely available at `https://github.com/N-BodyShop/skid`

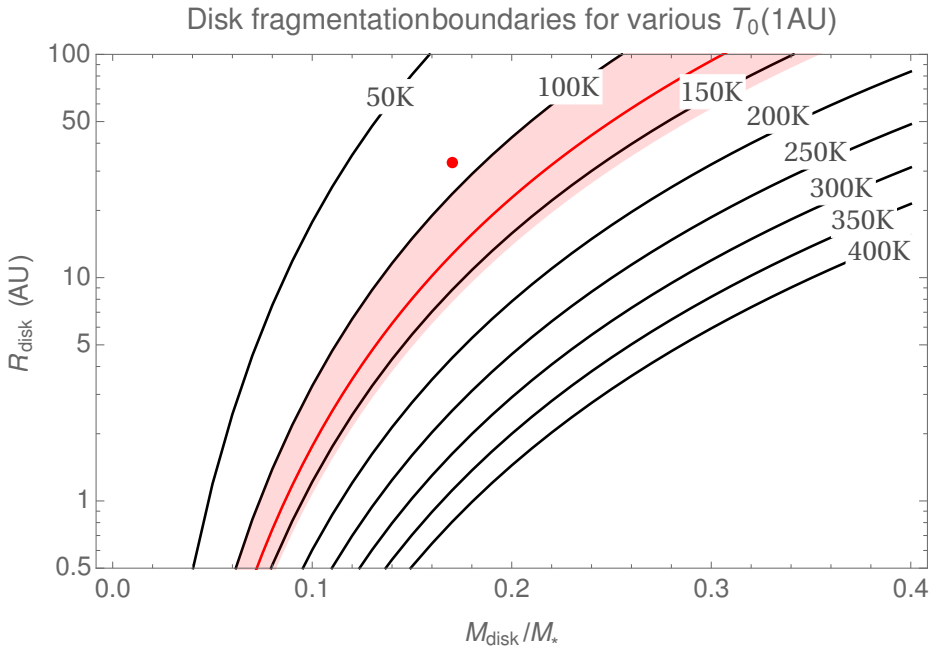



Figure 3.6: Fragmentation criteria for disk ICs. The curved lines define which disks will fragment for various disk temperatures at 1 AU, assuming a temperature profile of $T \propto r^{-0.59}$ and a surface density profile of $\Sigma \propto 1/R$. ICs which lie to the right of a line will fragment. The boundaries are $Q_{\text{eff}} = 1$ contours, where the $Q_{\text{eff}}$ estimates include approximations for disk height and disk self gravity. The red line marks the boundary defined by the fiducial temperature of $T_0(1\text{ AU}) = 130 \pm 25K$ from **?** and the red dot marks the fiducial disk mass and radius for a young disk from **?** (see §3.2.2 for a discussion of these values). These fiducial values likely have large uncertainties (not pictured).

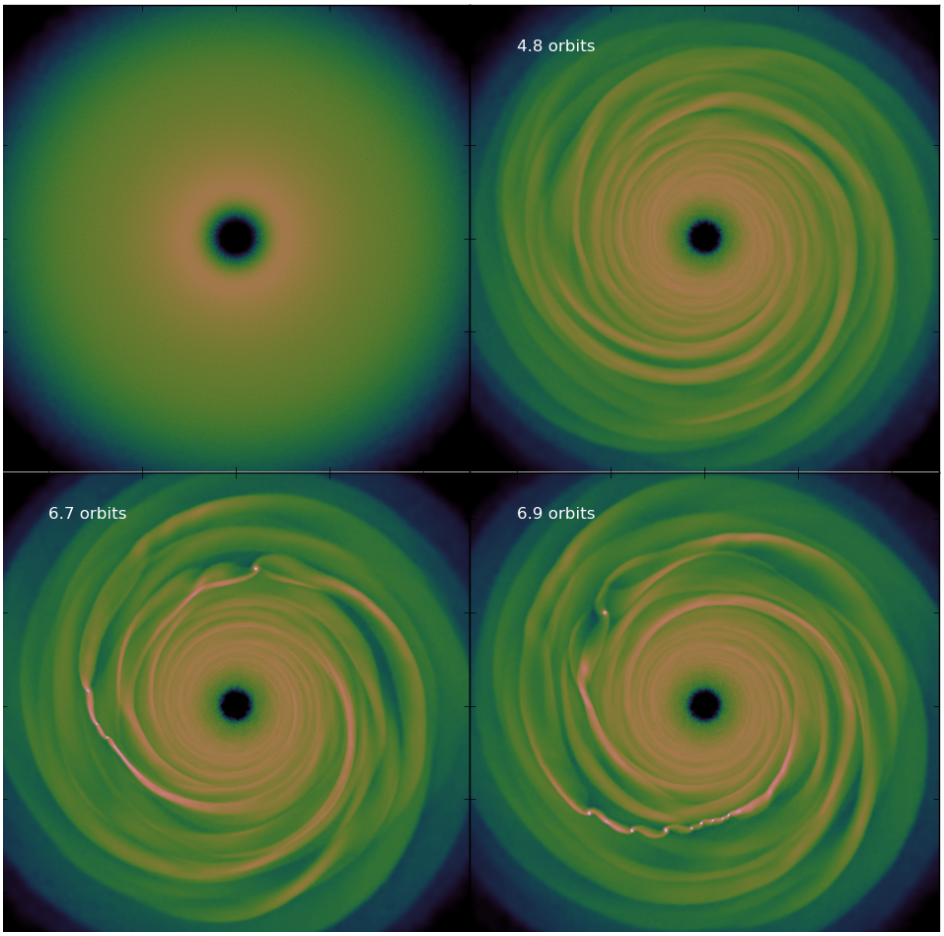



Figure 3.7: An example of the typical stages of clump formation. Clockwise, from top-left: (1) Initial conditions $Q_{\text{eff}} < 1$ (2) Strong spiral structure develops. (3) Spiral arms become overdense and break apart into clumps. (4) The disk begins fragment strongly.

The SKID algorithm is then applied, which uses a friends-of-friends clustering algorithm followed by a gravitational unbinding procedure to determine which clumps are gravitationally bound. By visual inspection, this provides robust results over a large range of disk and simulation parameters, and importantly it avoids marking high-density spiral arms as clumps. Figure 3.8 shows an example of this clump finding applied to a highly unstable PPD.

To track clumps over many time steps, clumps are first found in all simulation snapshots. They are then tracked over time by comparing clumps in adjacent time steps and seeing which have the most particles in common. Mergers (including multiple mergers), clump destruction, and clump formation are accounted for. Clump parameters such as mass, density, size, and location are all calculated and followed as a function of time.

We find that clumps form according to the same general picture, as shown in figure 3.7. Disks with a minimum $Q_{\rm eff} \lesssim 1.6$ (or about $Q_{min} \lesssim 1.3$) will form noticeable spiral structure after several orbital periods. Disks with an initial $Q_{eff} < 1$ will grow overdense spiral arms which will collapse and fragment into several dense clumps. Clumps initially form near the most unstable disk radius. For these simulations, that is approximately at $R_d$. They form after several to tens of orbital periods at the most unstable radius. Average clump masses are around $0.3 M_{Jupiter}$, with some rapidly growing to $1 M_{Jupiter}$. The timescales for disk fragmentation increase rapidly as $Q_{\rm eff} \to 1$ (fig. 3.5). After this stage, the isothermal approximation begins to break down. The disks then undergo a rapid, violent fragmentation.

## *3.5 Discussion*

### *3.5.1 Thermodynamics*

For these simulations we used a locally isothermal approximation for several reasons. We wished to perform a large scan of parameter space without compromising resolution too strongly. A computationally fast isothermal EOS is straightforward to implement. We also desired to build on previous work and to extend it to poorly studied M-dwarf systems. Our work here is directed at exploring the dependence of the fragmentation boundary on

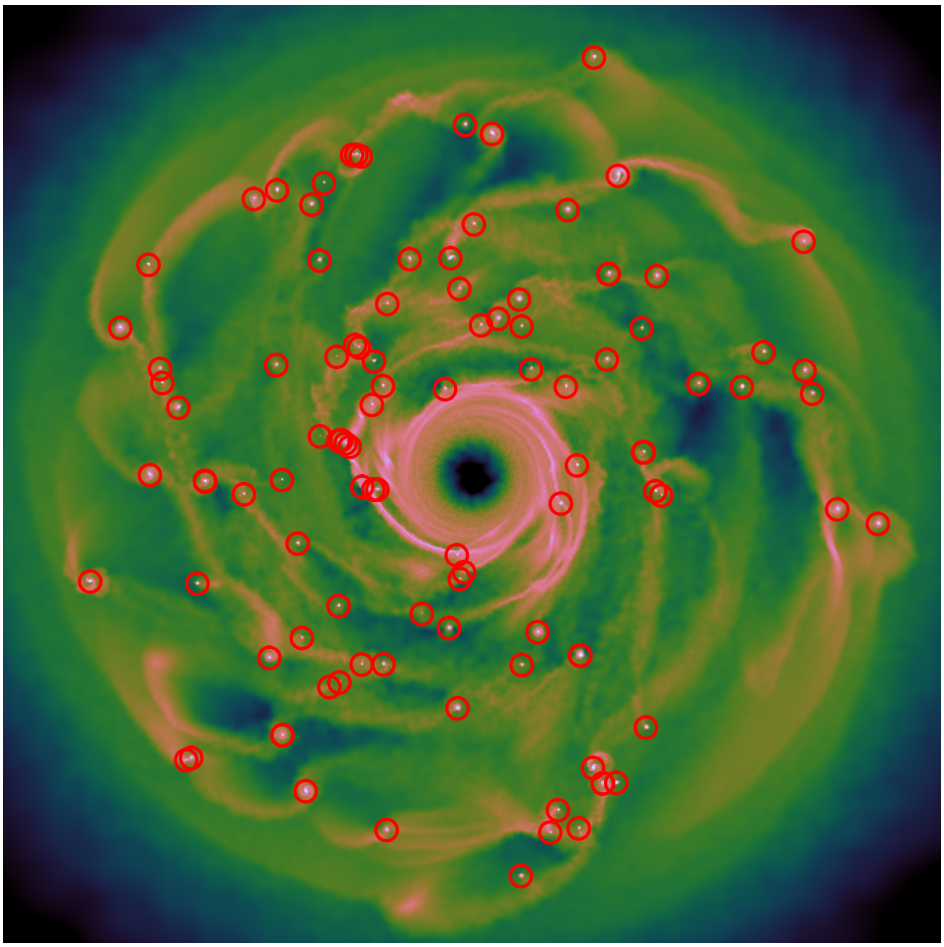


Figure 3.8: A demonstration of the clump finding algorithm used here. Integrated column density for the gas is pictured with a logarithmic color scale. Red circles mark the detected clumps. At the end of the simulation, this highly unstable disk formed 108 distinct, gravitationally bound clumps. The algorithm picks out clumps with a high success rate without reporting false positives from other high density structure such as spiral arms.

stellar/disk mass, disk height, and ICs. We leave the dependence on EOS for future work.

A non-isothermal EOS introduces non-trivial numerical issues, especially in the context of SPH simulations of protoplanetary disks. Unwarranted, poorly understood heating terms, especially from artificial viscosity (AV), are introduced into the energy equation. Previous results [**?????**] and our own initial tests indicate that in the context of PPDs, AV heating can dominate disk thermodynamics. Non-isothermal PPD simulations may not converge [**?**], whereas in §2.4 we demonstrate that our approach does converge. It is also unclear how rapidly disks will radiatively cool, an important parameter for the possibility of disk fragmentation [**???**]. We hope to investigate these effects in future work.

The isothermal approximation used for these simulations limits the scope of our results. An isothermal EOS would well approximate a disk where stellar radiation, viscous accretion heating, and radiative losses to infinity, are nearly balanced and control the temperature of the disk. Furthermore, temperature is independent of density, which is only appropriate for an optically thin disk. The isothermal approximation applies only to scenarios where dynamical timescales are much longer than the timescales for heating/cooling back to thermal equilibrium with the background.

These conditions may not hold in the disks under consideration. This experiment therefore does not realistically follow the thermal evolution of the disk. We are limited to a preliminary investigation into the large-scale dynamics of the disk before the putative equilibrium temperature profile would be expected to be strongly altered.

During the initial stages of disk evolution, dynamical and physical timescales are of order the orbital period and disk radius, respectively. During this stage, the isothermal EOS can still provide insight into the global dynamics of a GI disk at a certain stage. However, once clumps form, the isothermal approximation no longer provides much insight. Clumps should get hot as they collapse. The dynamic timescales of dense clumps will be short as they accrete matter, decouple from the disk, and scatter with other clumps. Pressure support of clumps, which is poorly captured by an isothermal EOS, should tend to increase their size and their coupling to the disk, meaning that the violent fragmentation of disks after initial

clump formation which we see in our simulations may not be the final state of a typical fragmenting disk. Clumps which are sufficiently dense may decouple from the disk enough to experience strong shocks and tidal interactions which will cause heating. These processes will strongly influence clump growth, evolution, and survival, all of which are still under investigation [**??**].

We therefore limit ourselves to discussing the early stages of clump formation. Our results indicate that the critical value of $Q_{min} \lesssim 0.9$ required for fragmentation is significantly lower than some previous results which found closer to $Q_{min} \lesssim 1.5$. Although this makes requirements for fragmentation more stringent, it does not rule out GI and disk fragmentation as important mechanisms during planet formation in PPDs around M-dwarfs.

### *3.5.2 Previous results*

We find that for most disks, $Q \lesssim 0.9$ is required for disk fragmentation. Re-parameterizing $Q$ as $Q_{eff}$ to include the stabilizing effect of disk height provides a more precise way to predict fragmentation. The ratio of $Q_{eff}/Q$ can vary by 30% for reasonable disk parameters. This ratio will vary even more when considering solar type stars in addition to M-dwarfs. However, $Q$ still provides a reasonable metric for disk fragmentation.

Other isothermal studies have found different boundaries. In contrast to our results, **?** found the threshold to be $Q \leq 1.5$. **?** found $Q$ as high as 1.3 would fragment. **?** found, using an isothermal EOS or diffusive radiative transfer, that $Q = 1.3 - 1.5$ would fragment. **?** found $Q = 1.4$ isothermal disks would fragment. **?** found that cooling a disk from $Q = 1.8$ to $Q = 0.9$ caused it to fragment. Their clumps did not survive, although as they note that may be due to numerical issues. **?** found that disks could be pushed below $Q = 1$ by mass loading and still not fragment, by transporting matter away from the star and thereby decreasing $\Sigma$ and increasing $Q$. It should be noted that some of these simulations were run at much lower resolution than ours. Differences may also be due in part to simulation methods: using cylindrical grids, spherical grids, or SPH methods; applying perturbations; or even 2D [**?**] vs 3D simulations.

Since all the details of previous work are not available, the source of the discrepancy in critical $Q$ values is uncertain. One source of discrepancy is simply how $Q$ is calculated. As mentioned in §3.3, approximating $Q$ by ignoring disk self-gravity and pressure forces can overestimate $Q$ on the 10% level for heavy disks.

The discrepancy may also be due to the different methods of constructing equilibrium disks. As demonstrated in §2.3, overestimating velocities at less than the percent level can force a disk to fragment. Disks near the fragmentation boundary are very sensitive to ICs. Initial conditions are not in general available for previous work, however we can note that some studies appear to display a rapid evolution of $Q$ at the beginning of the simulation. For example, some runs of **?** evolve from $Q_{min} = 1.5$ to $Q_{min} = 1$ in fewer than 3 ORPs (see their Figure 14). Figure 1 of **?** shows two isothermal simulations which evolve from a $Q_{min}$ of 1.38 and 1.65 to $Q_{min} = 1$ in fewer than 2 ORPs. This is indicative of ICs which are out of equilibrium.

In contrast, even our very unstable disks display a remarkably smooth and gradual initial evolution. Figure 3.9 shows the behavior of the minimum $Q_{\rm eff}$ (normalized by its initial value) for our fragmenting runs as a function of time until fragmentation. For all the runs, $Q_{\rm eff}$ decreases gradually for most of the simulation until dropping rapidly shortly before the disk fragments. For our runs near the fragmentation boundary, this is much more gradual and much less pronounced than for the runs of **?** or **?** mentioned above. For us, a much smaller change in $Q$ takes around 10 ORPs. $Q_{\rm eff}$ evolves even more slowly for non-fragmenting runs. $Q_{min}$ follows a similar behavior, although with more scatter (in large part because $Q$ does not determine the timescale until fragmentation as well as $Q_{\rm eff}$ does).

However it is not certain how close to equilibrium ICs should be to capture the relevant physics of PPDs. Actual disks are constantly evolving from the early stages of star formation until the end of the disk lifetime. We chose to use disks as close to equilibrium as possible, seeded only with SPH poisson noise, to avoid introducing numerical artifacts. Some authors introduce density perturbations which are controllable. If sufficiently large, they may serve to ameliorate the problems mentioned above by explicitly having fragmentation be driven

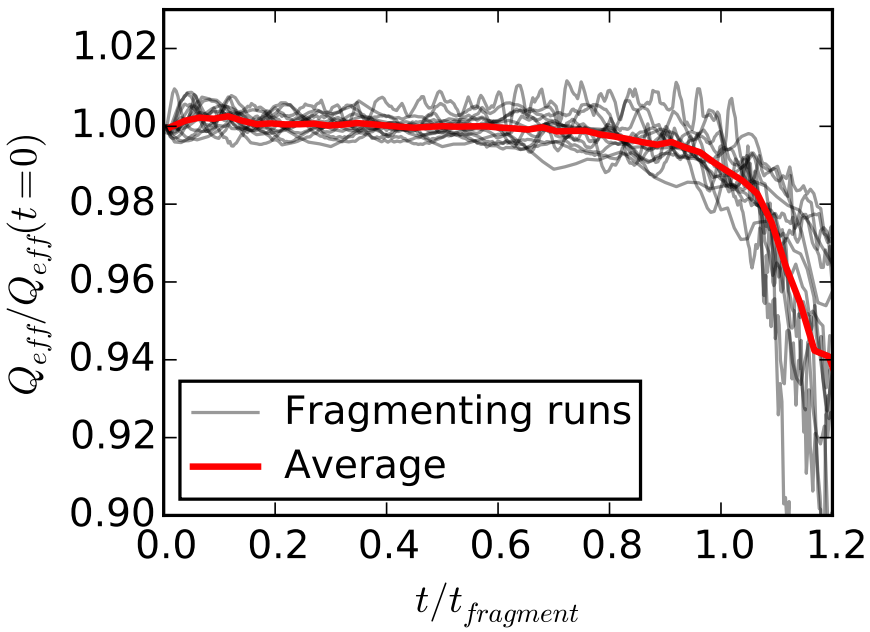



Figure 3.9: Minimum $Q_{\text{eff}}$ normalized by the initial minimum $Q_{\text{eff}}$ vs fraction of time until fragmentation for all the fragmenting runs. The average of these runs is plotted in red. Disk fragmentation occurs at $t/t_{fragment} = 1$. All runs follow similar trajectories in this plot, even though a significant range of initial $Q_{\text{eff}}$ and $t_{fragment}$ values are represented here (all the fragmenting runs in Fig. 3.5 are presented here). The simulations undergo an initially gradual decrease in $Q_{\text{eff}}$ which steepens sharply shortly before fragmentation.

by physically reasonable spiral modes (e.g. **?**) or intentionally large random perturbations (e.g. **?**). Others have considered mass loading as a means to grow to low $Q$ [**?**].

### *3.5.3 GI in PPDs*

Observed disks around M-dwarfs do not appear to have low enough inferred $Q_{min}$ values to be sufficiently unstable for fragmentation under GI, however this is what would be expected given the short timescales for the fragmentation of an unstable disk. Reported disk ages are of order $10^6$ years [**?**], orders of magnitude longer than fragmentation timescales, making the observation of a highly gravitationally unstable disk unlikely. This is a strong selection effect for observed disk parameters.

Although fragmentation timescales are very rapid, disks may persist much longer in moderately unstable configurations where GI drives large scale structure but which have not grown sufficiently unstable as to be prone to fragmentation. Early work is being done on trying to observe GI driven structures, but with current instrumentation such structures will be difficult to resolve [**??**].

As shown in figure 3.6, the fiducial values for disk parameters adopted here place observed disks near the boundary for fragmentation. The fact that observations indicate normal disk parameters which are close to the boundary, rather than orders of magnitude off, suggests it is plausible that a significant portion of PPDs around M-dwarfs will undergo fragmentation. This would predict a sharp transition in the distribution of inferred $Q_{\rm eff}$ values around $Q_{\rm eff} = 1$. Older disks tend to expand radially [**?**], thereby decreasing $\Sigma$, increasing $Q$, and pushing them away from the $Q_{\rm eff} = 1$ boundary. We note that while $Q$ is a reasonably strong predictor of fragmentation, disk height is an additional parameter worth measuring to predict fragmentation.

Given the results of these isothermal simulations, we expect GI to play a large role in the early stages of planet formation around M-dwarfs. The exact role of star mass/type on fragmentation remains unclear. The parameter space we scanned is sufficiently large that adding an extra dimension was prohibitive, we therefore only studied one star mass. Future

work should be able to determine what stars are the most suitable for fragmentation. Once large-scale density perturbations are formed via GI, the fate of the disk remains unclear. Future work should include more sophisticated thermodynamics to follow the evolution of the gaseous component of the disk to better determine under what conditions clumps will form, and what is required for them to survive. Additionally, decreasing the resolution in our isothermal SPH simulations appears to drive fragmentation (see §2.4). The importance of resolution in simulations is subject to much investigation and should be further pursued (e.g. **???**).

Planet formation will of course require the concentration of solids as well. Including dust in simulations of young PPDs will be required. In a fully 3D, highly non-axisymmetric environment, we may investigate how GI affects solids. Dust enhancement through pressure gradients, dust evolution through collisions and coagulation, and dust coupling to disk opacity and cooling, will all strongly affect prospects for planet formation.

| Sim. | Profile | $R_d$ (AU) | $T_0$ (K) | $M_d$ ($M_\odot$) | $Q$ | $Q_{eff}$ | fragment? |
|---|---|---|---|---|---|---|---|
| 0 | powerlaw | 1/3 | 64 | 0.02 | 0.78 | 0.76 | yes |
| 1 | powerlaw | 3 | 64 | 0.03 | 0.82 | 0.86 | yes |
| 2 | powerlaw | 30 | 64 | 0.04 | 0.88 | 1.01 | — |
| 3 | powerlaw | 1/3 | 129 | 0.03 | 0.81 | 0.84 | yes |
| 4 | powerlaw | 3 | 129 | 0.04 | 0.86 | 0.96 | yes |
| 5 | powerlaw | 30 | 129 | 0.06 | 0.92 | 1.12 | — |
| 6 | powerlaw | 1/3 | 258 | 0.04 | 0.85 | 0.93 | yes |
| 7 | powerlaw | 3 | 258 | 0.06 | 0.91 | 1.07 | — |
| 8 | powerlaw | 30 | 258 | 0.09 | 0.98 | 1.26 | — |
| 9 | powerlaw | 1/3 | 64 | 0.01 | 0.95 | 0.94 | yes |
| 10 | powerlaw | 3 | 64 | 0.02 | 0.99 | 1.05 | — |
| 11 | powerlaw | 30 | 64 | 0.04 | 1.04 | 1.23 | — |
| 12 | powerlaw | 1/3 | 129 | 0.02 | 0.98 | 1.03 | — |
| 13 | powerlaw | 3 | 129 | 0.03 | 1.02 | 1.16 | — |
| 14 | powerlaw | 30 | 129 | 0.05 | 1.09 | 1.35 | — |
| 15 | powerlaw | 1/3 | 258 | 0.03 | 1.01 | 1.13 | — |
| 16 | powerlaw | 3 | 258 | 0.04 | 1.07 | 1.29 | — |
| 17 | powerlaw | 30 | 258 | 0.07 | 1.12 | 1.48 | — |
| 18 | powerlaw | 1/3 | 64 | 0.01 | 1.12 | 1.12 | — |
| 19 | powerlaw | 3 | 64 | 0.02 | 1.15 | 1.25 | — |
| 20 | powerlaw | 30 | 64 | 0.03 | 1.21 | 1.45 | — |
| 21 | powerlaw | 1/3 | 129 | 0.02 | 1.14 | 1.22 | — |
| 22 | powerlaw | 3 | 129 | 0.03 | 1.19 | 1.37 | — |
| 23 | powerlaw | 30 | 129 | 0.04 | 1.24 | 1.56 | — |
| 24 | powerlaw | 1/3 | 258 | 0.02 | 1.18 | 1.33 | — |
| 25 | powerlaw | 3 | 258 | 0.04 | 1.23 | 1.51 | — |
| 26 | powerlaw | 30 | 258 | 0.06 | 1.27 | 1.7 | — |
| 27 | powerlaw | 1/3 | 64 | 0.01 | 1.28 | 1.3 | — |
| 28 | powerlaw | 3 | 64 | 0.02 | 1.32 | 1.45 | — |
| 29 | powerlaw | 30 | 64 | 0.03 | 1.37 | 1.67 | — |
| 30 | powerlaw | 1/3 | 129 | 0.01 | 1.31 | 1.41 | — |
| 31 | powerlaw | 3 | 129 | 0.02 | 1.35 | 1.59 | — |
| 32 | powerlaw | 30 | 129 | 0.04 | 1.4 | 1.78 | — |
| 33 | powerlaw | 1/3 | 258 | 0.02 | 1.34 | 1.54 | — |
| 34 | powerlaw | 3 | 258 | 0.03 | 1.39 | 1.73 | — |
| 35 | powerlaw | 30 | 258 | 0.05 | 1.42 | 1.93 | — |
| 36 | powerlaw | 1 | 64 | 0.02 | 0.8 | 0.81 | yes |
| 37 | powerlaw | 9 | 64 | 0.03 | 0.84 | 0.92 | yes |
| 38 | powerlaw | 1 | 129 | 0.03 | 0.83 | 0.89 | yes |
| 39 | powerlaw | 9 | 129 | 0.05 | 0.89 | 1.03 | — |
| 40 | powerlaw | 1 | 258 | 0.04 | 0.87 | 0.99 | — |
| 41 | powerlaw | 9 | 258 | 0.07 | 0.94 | 1.15 | — |
| 42 | powerlaw | 1/3 | 64 | 0.02 | 0.87 | 0.85 | yes |
| 43 | powerlaw | 1 | 64 | 0.02 | 0.88 | 0.9 | yes |

| Sim. | Profile | $R_d$ (AU) | $T_0$ (K) | $M_d$ ($M_\odot$) | $Q$ | $Q_{eff}$ | fragment? |
|---|---|---|---|---|---|---|---|
| 44 | powerlaw | 3 | 64 | 0.02 | 0.9 | 0.96 | yes |
| 45 | powerlaw | 9 | 64 | 0.03 | 0.93 | 1.02 | — |
| 46 | powerlaw | 30 | 64 | 0.04 | 0.96 | 1.12 | — |
| 47 | powerlaw | 1/3 | 129 | 0.02 | 0.89 | 0.93 | yes |
| 48 | powerlaw | 1 | 129 | 0.03 | 0.92 | 0.99 | — |
| 49 | powerlaw | 3 | 129 | 0.04 | 0.94 | 1.06 | — |
| 50 | powerlaw | 9 | 129 | 0.04 | 0.97 | 1.14 | — |
| 51 | powerlaw | 30 | 129 | 0.06 | 1.0 | 1.23 | — |
| 52 | powerlaw | 1/3 | 258 | 0.03 | 0.93 | 1.03 | — |
| 53 | powerlaw | 1 | 258 | 0.04 | 0.96 | 1.1 | — |
| 54 | powerlaw | 3 | 258 | 0.05 | 0.99 | 1.18 | — |
| 55 | powerlaw | 9 | 258 | 0.06 | 1.02 | 1.27 | — |
| 56 | powerlaw | 30 | 258 | 0.08 | 1.06 | 1.38 | — |
| 57 | viscous | 5 | 129 | 0.04 | 0.91 | 1.01 | yes |
| 58 | viscous | 9 | 129 | 0.04 | 1.0 | 1.13 | — |
| 59 | viscous | 11 | 129 | 0.04 | 1.04 | 1.19 | — |
| 60 | viscous | 13 | 129 | 0.04 | 1.09 | 1.26 | — |
| 61 | viscous | 16 | 129 | 0.04 | 1.13 | 1.32 | — |
| 62 | viscous | 19 | 129 | 0.04 | 1.17 | 1.38 | — |
| 63 | viscous | 27 | 129 | 0.04 | 1.25 | 1.51 | — |

Table 3.1: The suite of runs presented here.

# Chapter 4

# DUST PROCESSING IN GRAVITATIONALLY ACTIVE PROTOPLANETARY DISKS

## *4.1 Introduction*

Growth of solids from micron to meter scales and above—a requirement for the formation of terrestrial planets and gas giant rocky cores—is not trivial in protoplanetary disk (PPD) environments. Overcoming the barriers to grain growth will likely require significant concentration of millimeter and smaller grains above stellar nebulae abundances. Growth to centimeter sized grains (pebbles) requires that small grains collide and stick without destruction. This requires surpassing the "bouncing barrier" [**???**] and the "destruction barrier" [**?**]. Sufficiently small grains are strongly coupled to the gas and therefore have low relative velocities in most regions of the disk. Laboratory experiments have demonstrated that under these conditions, grains tend to bounce during collisions, thus preventing growth. Larger grains which decouple more from the gas and have larger velocity dispersions must avoid destruction during collisions in order to grow.

Growth beyond meter sized objects requires overcoming the "meter barrier" [**?**]. The gas component of a PPD experiences a small amount of radial pressure support which the dust does not. The dust therefore orbits more quickly and experiences a headwind. This generates a torque which decreases the angular momentum of the dust. Very small dust grains couple strongly to the gas and tend to orbit with the gas, while the drag force is insignificant for very large solids which decouple from the gas. Objects between these scales—from mm to km sizes depending on local conditions, though around a meter for typical PPDs—will experience very rapid radial migration from 1 AU on time-scales of hundreds of years, much faster than planet formation time-scales, and will fall rapidly into the central star.

Overcoming these growth barriers may require significant concentration of solid material. Current instruments, in particular ALMA, are sensitive to observing mm sized dust grains in PPDs and are beginning to provide a sensitive laboratory for small solids in PPDs [**??**]). Our work, which takes a theoretical approach, is aimed at addressing solid-gas interactions in PPDs and how they dictate the conditions of solid growth and planet formation.

One mechanism for dust concentration is vertical dust settling. Gas and dust in the disk will feel a vertical component of gravity toward the midplane. For a light, thin disk with a uniform vertical temperature profile, the gas has a nearly gaussian vertical density profile with a scale-height of $H_g = c_s/\Omega$ due to pressure support, where $c_s$ is the isothermal sound speed and $\Omega$ is the orbital angular velocity. The dust component is not pressure supported and will therefore tend to settle toward the midplane. However, dust settling may be limited by disk turbulence or large-scale vertical motion of the gas in a non-axisymmetric disk [**??**].

Dust may also be enhanced due to local pressure maxima in the disk. It has been shown numerically that dust can accumulate at pressure maxima in PPDs [**?**]. Spiral arms due to gravitational instability (GI) naturally generate well defined pressure maxima and may generate significant dust enhancement. Sufficiently cool or massive disks will naturally generate spiral structure for Toomre $Q$ parameter [**?**] $Q = c_s\kappa/\pi G\Sigma$ less than about 2 [**??**, Chapter 3], where $\kappa$ is the epicyclic frequency (equal to $\Omega$ for a Keplerian disk) and $\Sigma$ is the disk surface density. Indeed spiral structure has been observed in PPDs, in particular around Elias-27, which has been imaged using ALMA in dust SEDs [**?**].

Until recently, fully realistic 3D hydrodynamic simulations of dust in PPDs have not been possible due to computational and numerical constraints. Naive approaches are not computationally feasible given the harsh time-stepping constraints imposed by modeling the diffusion-like dynamics of dust grains. **?** (hereafter **?**) proposed a smoothed particle hydrodynamics (SPH) implementation of a terminal velocity approximation for small dust grains, which avoids these time-stepping constraints. This method is self-consistent, including the dynamical effect of dust on the gas, and potentially fast. Recent studies have used this method to model dust migration in a circumbinary PPD [**?**] and to model dust enhancement

in molecular clouds [**?**].

Here we study the effect of GI on dust concentration and migration in PPDs, paying particular attention to the role of spiral-arms. We use the one fluid terminal velocity dust algorithm of **?**, which we implement in ChaNGa. We test and analyze the **?** dusty-gas algorithm in the context of realistic PPDs. To that end we present a suite of SPH PPD simulations, including the highest resolution SPH PPD simulation to-date with $10^8$ particles.

In §4.2 we discuss the dusty-gas algorithm of **?** as implemented here, along with test results, which motivate our resolution choices for our PPD simulations. Section 4.3 describes the disk model presented here, the initial conditions (ICs), and the run parameters chosen. We present our results in §4.4, including the gas and dust dynamics. The problem of dust mass conservation inherent to this algorithm is presented and discussed in §4.5. Our discussion is presented in §4.6.

### *4.2 Dusty-gas algorithm*

To model dust dynamics, we use the one-fluid SPH implementation of **?** (hereafter **?**) with some minor modifications required to adapt it for our simulation code ChaNGa. **?** use a one fluid SPH implementation of dust-gas coupling via a drag force. Dust grains of one size are modeled. Compared to two-fluid approaches [**??**], the one-fluid approach is much less noisy and is faster due to needing only gas particles rather than gas and dust particles. For small dust grains at low Stokes number $\mathrm{St} \ll 1$ (typically cm or smaller, depending on local conditions), the punishing time-stepping constraint of the diffusion equation is avoided by assuming a terminal velocity approximation:

$$\Delta\mathbf{v} \equiv \mathbf{v}_d - \mathbf{v}_g = t_s \Delta\mathbf{f} \tag{4.1}$$

Where $\mathbf{v}_d$, $\mathbf{v}_g$ are the dust and gas velocities, $t_s$ is the dust stopping time and $\Delta\mathbf{f} \equiv \mathbf{f}_d - \mathbf{f}_g$ is the difference in acceleration due to external forces acting on the dust and gas components. Note that the opposite sign is often used in drag studies. This approximation is strictly valid for only low Stokes numbers. For the disks studied here, this approximation holds well

everywhere in the disk except at extremely large $R$. In most cases, for high Stokes numbers this will under-diffuse the dust. Under this approximation, the time-stepping constraint for the dust-gas interaction is satisfied by the Courant conditions, which is already required for hydrodynamics.

To implement the one-fluid approach, an extra SPH quantity is integrated. The dust fraction is defined as $\epsilon \equiv \rho_d/\rho$, where $\rho$ is the volume density of the dust and $\rho$ is the total (gas + dust) volume density. The gas density is then given by $\rho_g = (1-\epsilon)\rho$.

Here we reproduce the hydrodynamic equations of **?** for this approximation, neglecting external heating/cooling terms, which are not used in this study.

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho(\nabla \cdot \mathbf{v}) \tag{4.2}$$

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = (1-\epsilon)\mathbf{f}_g + \epsilon\mathbf{f}_d + \mathbf{f_{ext}} \tag{4.3}$$

$$\frac{\mathrm{d}\epsilon}{\mathrm{d}t} = -\frac{1}{\rho}\nabla \cdot [\epsilon(1-\epsilon)\rho t_s \Delta\mathbf{f}] \tag{4.4}$$

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -\frac{P}{\rho_g}(\nabla \cdot \mathbf{v}) + \epsilon t_s(\Delta\mathbf{f} \cdot \nabla u) \tag{4.5}$$

where $\mathbf{f_{ext}}$ is external specific force acting on both the dust and gas components (e.g., gravity) and $\mathbf{f}_d$ and $\mathbf{f}_g$ are the external specific forces (or accelerations) acting on only the dust, gas components, respectively. Pressure gradients factor only into the $\mathbf{f_g}$ term. The operator $\frac{\mathrm{d}}{\mathrm{d}t} \equiv \frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla$ is the standard convective derivative used in Lagrangian hydrodynamics. The standard specific internal energy of the gas is given by $u$, and $\mathbf{v}$ is the barycentric velocity of the dusty-gas mixture. In the case of hydrodynamics and gravity only (as used in this study):

$$\mathbf{f}_d = 0; \tag{4.6}$$

$$\mathbf{f}_g = -\frac{\nabla P}{\rho_g} + \mathbf{f}_{AV}; \tag{4.7}$$

$$\mathbf{f}_{ext} = \mathbf{f}_{grav}; \tag{4.8}$$

where $\mathbf{f}_{grav}$ is the specific force due to gravity and $\mathbf{f}_{AV}$ is due to artificial viscosity. As with **?**, $\mathbf{f}_{AV}$ is only applied to the gas component.

### *4.2.1 Implementation*

Equation (4.2) is unchanged from the standard SPH continuity equation. In the case of hydrodynamics + gravity (which we use here), eqs.(4.6)-(4.8) imply that the momentum equation becomes (4.3):

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\frac{\nabla P}{\rho} + \mathbf{f}_{grav} + (1-\epsilon)\mathbf{f}_{AV} \tag{4.9}$$

This is identical to standard SPH, but with the factor of $(1-\epsilon)$ multiplying $\mathbf{f}_{AV}$. Therefore, our implementation of dusty-gas is primarily restricted to the dust fraction and energy equations.

To apply SPH discretization to these equations, **?** propose two approaches. Eqs.(4.6) & (4.7) imply that with hydrodynamics eqs.(4.4) & (4.5) involve second derivatives. The "two-loop" method requires an extra loop over neighbors to calculate the second derivatives in eq.(4.4), whereas the "one-loop" method estimates the second derivatives directly.

The "one-loop" method avoids an extra loop over neighbors, making it the computationally faster approach. However, it has two main flaws. First, the one-loop method makes assumptions about the AV scheme used. In particular, the derivation of the method assumes negligible artificial viscosity such that $\mathbf{f}_{AV} \approx 0$ in eq.(4.7). The derivation also implicitly assumes a certain form of AV and the momentum equation [**?**]. Our tests demonstrate very bad behavior when a different AV scheme is used. Additionally, this method may be inconsistent with some weighting schemes used in calculating the momentum equation (eq. 4.3).

Second, the one-loop approach suffers more strongly from noise in the second derivative estimates. Noisy second derivatives are a well known problem in SPH [**?**]. As demonstrated in the PPD simulation results presented in §4.4, this dusty-gas method suffers generally from strong SPH noise. It is therefore important to avoid noisy second derivatives.

While the two-loop method is slower, in the production version of the code the cost is fairly insignificant. In ChaNGa, the majority of time is spent building the Barnes-Hut [**?**] tree, calculating gravity, and initializing the SPH loops over neighbors. An extra loop over neighbors is rapid. For a PPD in ChaNGa, the two-loop method typically requires an

increase of a few percent in wall clock time.

We perform the two-loop method by first looping over neighbors using the standard SPH approach, which allows us to estimate:

$$\Delta\mathbf{f} = -\frac{1}{1-\epsilon}\left(\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t}\right)_{SPH} \tag{4.10}$$

in which $\mathbf{f}_d = 0$ and $\left(\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t}\right)_{SPH}$ is the hydrodynamic acceleration on the particle, including AV terms and pressure gradients. A second loop over neighbors is then performed to calculate the dust fraction and energy equations. Let:

$$G_{ab} \equiv \frac{\epsilon_a(1-\epsilon_a)t_{s,a}}{\rho_a}\Delta\mathbf{f}_a \cdot \nabla W_{ab}(h_a) \tag{4.11}$$

where $W_{ab}(h_a)$ is the SPH kernel $W(|r_a - r_b|, h_a)$ and $h_a$ is the smoothing length of particle $a$. We can write the SPH discretization of eqs.(4.4)&(4.5) as:

$$\frac{\mathrm{d}\epsilon_a}{\mathrm{d}t} = -\sum_b m_b(G_{ab} - G_{ba}) \tag{4.12}$$

$$\frac{\mathrm{d}u_a}{\mathrm{d}t} = \frac{1}{1-\epsilon_a}\left(\frac{\mathrm{d}u_a}{\mathrm{d}t}\right)_{SPH} - \frac{1}{1-\epsilon_a}\sum_b m_b G_{ab}(u_a - u_b) \tag{4.13}$$

where $\left(\frac{\mathrm{d}u_a}{\mathrm{d}t}\right)_{SPH}$ is calculated from the normal SPH energy equation and represents heating by gas-gas interactions. The sum in eq.(4.13) is due to gas-dust interactions. These sums are essentially the same as **?** eqs.(26) & (27), but we have not specified the AV scheme or how exactly $\left(\frac{\mathrm{d}u_a}{\mathrm{d}t}\right)_{SPH}$ is calculated. Note that the term $\Omega_a \equiv 1 - \frac{\partial h_a}{\partial \rho_a}\sum_b m_b \frac{\partial W_{ab}(h_a)}{\partial h_a}$ in the SPH formulation used by **?** is replaced with $\Omega_a \to 1$ under our SPH implementation. This is explained in **?**.

As with **?**, we calculate the stopping time in the Epstein regime, valid for small dust grains at Stokes number $\ll 1$:

$$t_s = \sqrt{\frac{\pi\gamma}{8}}\frac{\rho_{\mathrm{grain}} s_{\mathrm{grain}}}{\rho c_s} \tag{4.14}$$

where $c_s = \sqrt{\gamma k_B T/m}$ is the adiabatic sound speed, $\rho_{\mathrm{grain}}$ is the intrinsic grain density (e.g. $\sim 3g/cm^3$ for silica) and $s_{\mathrm{grain}}$ is the size of the grains, e.g. the radius for spherical grains. Note that the $\sqrt{\gamma}$ factor in the expression for $t_s$ is canceled out by $\sqrt{\gamma}$ in $c_s$. The sound speed here is used as a proxy for the thermal velocity.

*Conservation*

Equation (4.13) is derived from demanding conservation of energy. As shown in **?**, eq.(4.12) conserves total dust mass exactly when (i) $m_a$ is constant with respect to $t$ for all particles and (ii) all particles are on the same time step.

Condition (i) is met since we do not change gas particle masses. Strictly speaking, condition (ii) is violated, since in production codes such as ChaNGa, multi-stepping is used to dramatically decrease computation cost. In ChaNGa, particles are dynamically assigned time-steps of the form $\Delta t = 2^{-N}\Delta t_0$ according to dynamical time-step constraints such as the Courant condition and acceleration constraints. $\dot{\epsilon}_a$ is only updated for particles on active time-steps [**??**]. In practice this produces minimal dust mass non-conservation. Occasionally, $\epsilon$ can go outside of the range [0,1]. As with **?**, when this happens we set $\epsilon = 0$ or 1. This can cause significant mass non-conservation, which we examine in detail in §4.5.

### *4.2.2 2D settling test*

In Appendix B we reproduce the results of the four tests presented in **?**. Our implementation is able to reproduce their results. Here we discuss some test results that are particularly relevant to PPDs.

The 2D dust settling test of **?** is a simple, idealized model of dust settling in a PPD disk. Turbulence, stirring due to spiral arms and disk structure, shearing, and 3D effects are all ignored. The results of this test motivate our choice of resolution for our 1e8 PPD run and indicate that we should be able to capture the basic vertical dust settling dynamics.

We follow their procedure for setting up the ICs. The set-up has an external vertical potential applied due to a star of $M_* = 1M_\odot$ at 50 AU along the $z$ direction. The test is run in a 2D box of width $\Delta x = 5$ AU with periodic boundary conditions along $x$ and open boundary conditions along $z$, without self-gravity and with an isothermal EOS. ICs are generated from gas-only particles on a hexagonal grid, stretched to give an approximately gaussian vertical density profile with an initial scale height $H/R = 0.05$, fixed by setting the

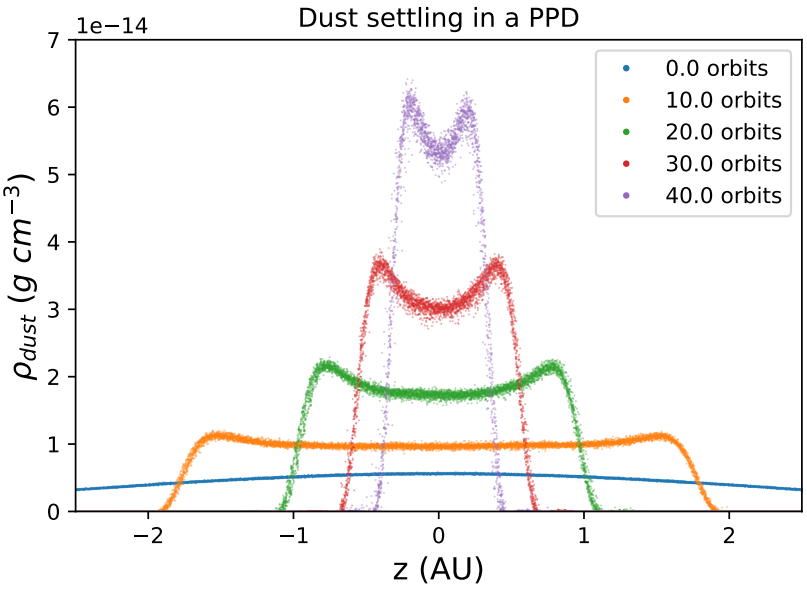



Figure 4.1: Results for the high resolution **?** dust settling test. This reproduces the results in the second panel of their Fig. 8. The dust density is plotted for all SPH particles at 5 epochs. This is a 2D test in x-z in a fixed external potential along z which represents a patch of a protoplanetary disk. The ICs are in equilibrium and are a nearly gaussian vertical density profile with a uniform dust fraction of $\epsilon \approx 0.01$.

sound speed to $c_s = \Omega H$ for $\Omega = \sqrt{GM_*/R^3}$. The ICs are then run with a velocity damping force to approach equilibrium before adding a uniform dust fraction of $\epsilon = 0.0099$. We use a grain size $s_{\rm grain} = 1$ mm and grain density $\rho_{\rm grain} = 3$ g cm$^{-3}$.

After the initial settling, the density profile is still very nearly gaussian, with a scale height of $H = 2.52$ AU. This yields a midplane density of $\rho_0 = 5.55 \times 10^{-13}$ g cm$^{-3}$ and a midplane stopping time of $t_s = 0.51$ yr[1]. As with **?**, the settling test was run at three resolutions: 896 (LoRes), 3552 (MedRes), and 14208 (HiRes) particles. The test was run for 50 ORPs, defined as $2\pi/\Omega = 10^{3/2}$ yr. For all runs we use $N_{smooth} = 28$ neighbors and the

[1] This value of $t_s$ does not agree with the value reported in **?**, due to a typo in their §4.4.1 for the $c_s$–$H$ relation which has an extra factor of $c_s$ which propagates into their reported $t_s$ value (eq. 58).

Wendland $C_4$ kernel. Figure 4.1 shows the results of the HiRes dust settling test (similar to **?**, Fig. 8). Dust density as a function of $z$ is plotted for all particles at several time steps. As expected, the dust settles to the midplane, with settling occuring more rapidly at higher elevations due to the increased $t_s$ and vertical component of the star's gravity, resulting in the double peaked profile apparent after 10 ORPs.

To check the validity of the approach, we have compared the results to a semi-analytic solution. Let us assume a negligible dust fraction $\epsilon \ll 1$ such that $\rho \approx \rho_g$ and such that the gas remains in equilibrium with a time-independent density profile of the form $\rho_g = \rho_0 \exp(-z^2/2H^2)$. Let us also assume very small dust particles, tightly coupled to the gas such that the terminal velocity approximation holds exactly, i.e. dust velocity is given by $v_d = t_s f_d$ where $f_d$ is the specific force acting on the dust, excluding gas-drag. For an isothermal EOS with uniform $c_s$, the dust velocity is then given as a function of position:

$$v_d(z) = \frac{\mathrm{d}z}{\mathrm{d}t} = -GM_* \sqrt{\frac{\pi\gamma}{8}} \frac{\rho_{\mathrm{grain}} s_{\mathrm{grain}}}{\rho_0 c_s} \frac{z e^{z^2/2H^2}}{(z^2+R^2)^{3/2}} \tag{4.15}$$

The lagrangian continuity equation can be evaluated as:

$$\frac{\mathrm{d}\rho_d}{\mathrm{d}t} = -\rho_d \frac{v_d}{z} \left[ 1 + z^2 \left( \frac{1}{H^2} - \frac{3}{z^2+R^2} \right) \right] \tag{4.16}$$

where the time derivative is a convective derivative. These equations can easily be numerically integrated.

Figure 4.2 compares $\rho_d$ for the semi-analytic solution with the test results at all 3 resolutions after 40 ORPs. For the semi-analytic solution, we integrate positions from $-3H$ to $3H$, equivalent to setting $\rho_d = 0$ for $|z| > 3H$. As resolution increases, the tests better approximate the semi-analytic curve, and the noise decreases. The MedRes and HiRes runs perform fairly well at approximating the width of the profile. Although still smoothed, the HiRes run is able to broadly capture the dual-humped nature of the profile. The semi-analytic profile is likely much more strongly peaked than a realistic solution would be, due to neglecting: the force of the dust on the gas as it gets highly concentrated; any 3D effects; disk shearing; and turbulence. Additionally, the semi-analytic profile neglects the SPH smoothing. When smoothed by the SPH kernel, the semi-analytic result more closely matches the simulation.

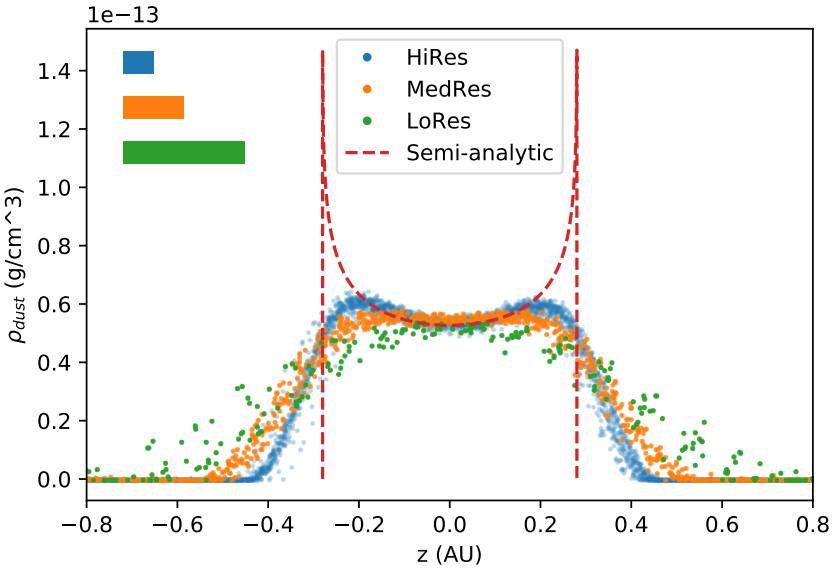



Figure 4.2: 2D dust settling test: resolution comparison. Plotted is the dust density for the 2D dust settling tests at 3 resolutions after 40 ORPs, for all particles. The semi-analytic solution is overlaid in red, assuming no dust beyond $3H$ for the ICs. The horizontal bars represent the midplane smoothing length for the different runs.

These tests indicate that we should be able to capture global dust settling rates in a full PPD simulation, and if our resolution is sufficient we should be able to capture the "outside-in" nature of the settling. We can naively scale the particle count of these tests from a 2D patch to a full 3D PPD. For a disk with radius $R$ and uniform $\Sigma$, the equivalent particle count is $N = \pi R^2 N_{2D}^{3/2}/(\Delta x)^2$. For these tests, that gives $N = 8.5 \times 10^6$, $6.7 \times 10^7$, and $5.3 \times 10^8$ particles for the LoRes, MedRes, and HiRes runs, respectively. Capturing the overall dust settling dynamics seems to require a resolution somewhere between the MedRes and the HiRes runs. Noise in $\rho_d$ is fairly well suppressed for the MedRes test. We therefore choose to use $10^8$ particles for our largest run as a decent compromise between computational cost and resolution. This represents the largest PPD run to-date that we are aware of.

Table 4.1: Simulations. $N_{SPH}$ is the number of SPH particles, $s$ is the grain size, $Q_{\rm eff}$ is the effective Toomre $Q$ (see §3.3) with a correction for scale height. Disks with $Q_{\rm eff} < 1$ are expected to fragment. The simulation runtime is given in ORPs, the rotational period at $R_d = 20$ AU. The 1e6 VeryHiQ run has the same $\tau_{\rm settle}$ as the 20 mm runs.

| Run | $N_{SPH}$ | $s$ (mm) | $Q_{\rm eff}$ | Frag? | $t_{run}$ (ORP) |
|---|---|---|---|---|---|
| 1e5 | $10^5$ | 1 | 1.03 | No | 132 |
| 1e6 | $10^6$ | 1 | 1.03 | No | 46 |
| 1e7 | $10^7$ | 1 | 1.03 | Yes | 14.8 |
| 1e8 | $10^8$ | 1 | 1.03 | No | 9.6 |
| 1e7 HiQ | $10^7$ | 1 | 1.10 | No | 68 |
| 1e7 20mm | $10^7$ | 20 | 1.10 | Yes | 5.4 |
| 1e7 20mm exp cut | $10^7$ | 20 | 1.10 | Yes | 4.7 |
| 1e7 20mm hard cut | $10^7$ | 20 | 1.10 | Yes | 2.7 |
| 1e7 5mm | $10^7$ | 5 | 1.10 | No | 34 |
| 1e6 20mm | $10^6$ | 20 | 1.03 | No | 50 |
| 1e6 VeryHiQ | $10^6$ | 0.27 | 63 | No | 50 |

### *4.3 Overview of runs*

For a full list of runs, see Table 4.1. The disk model here is similar to the fiducial model of Chapter 3 based on observations [**??**] of disks around M-Dwarfs, representing a typical young disk. The disk mass we use is slightly increased to place the disk near the $Q_{\rm eff} = 1$ fragmentation boundary, where $Q_{\rm eff}$ is the Toomre $Q$ parameter, modified to account for disk height, defined as $Q_{\rm eff} \equiv \beta Q (H/R)^{\alpha}$ (eq. 3.8) where we adopt $\alpha = 0.18$ and $\beta = 2.1$. This places our disks a little below $Q = 1$. Gaseous disks with $Q_{\rm eff} < 1$ are susceptible to

fragmentation. The star mass is $M_* = M_\odot/3$.

Equilibrium ICs were generated with the method of §2.2. The surface density profile used is a $\Sigma \propto R^{-1}$ powerlaw, with a smooth step-function interior cutoff applied at 2 AU and a gaussian cutoff applied at $R_d = 20$ AU. The central star is treated as a sink particle with a sink radius of about 1 AU. The surface density normalization is set by the required $Q_{\rm eff}$. The temperature follows a powerlaw of $T \propto R^{-q}$. This fiducial model uses a power law of $q = 0.59$, from the mean value for fits to dust SEDs for M-Stars in **?**, where we set the temperature at 1 AU to 130 K. Figure 4.3 shows the IC profiles for this disk model.

In order to investigate the resolution dependence of this dusty-gas algorithm in the context of a gravitationally active PPD, it was necessary to seed spiral arm growth with an artificial perturbation. Spiral arm growth in SPH simulations of axisymmetric PPDs is seeded by SPH noise, which is highly resolution dependent, making direct comparison of spiral arm dynamics at different resolutions very difficult. The spiral arm growth timescales vary with resolution and the dominant modes (especially early in the simulation) can vary stochastically.

An $m = 2$ density perturbation was applied to the disk after generating the equilibrium ICs. A surface density profile of:

$$\Sigma(r, \phi) = \Sigma(r)[1 + \delta\Sigma \sin(m\phi)] \tag{4.17}$$

was applied by shifting particles by a small amount along $\phi$:

$$\Delta\phi = \frac{\delta\Sigma}{m} \cos(m\phi) \tag{4.18}$$

For these runs we used a fractional surface density perturbation of $\delta\Sigma = 0.02$.

All the simulations performed here use a locally isothermal EOS with a mean molecular weight of 2, representing molecular hydrogen. As discussed in Chapter 3, while the isothermal approximation is an idealization, in the context of the early stages of gravitationally active PPDs it can reasonably capture the dynamics. Once a disk fragments, the approximation breaks down. However, current SPH approaches are not able to accurately

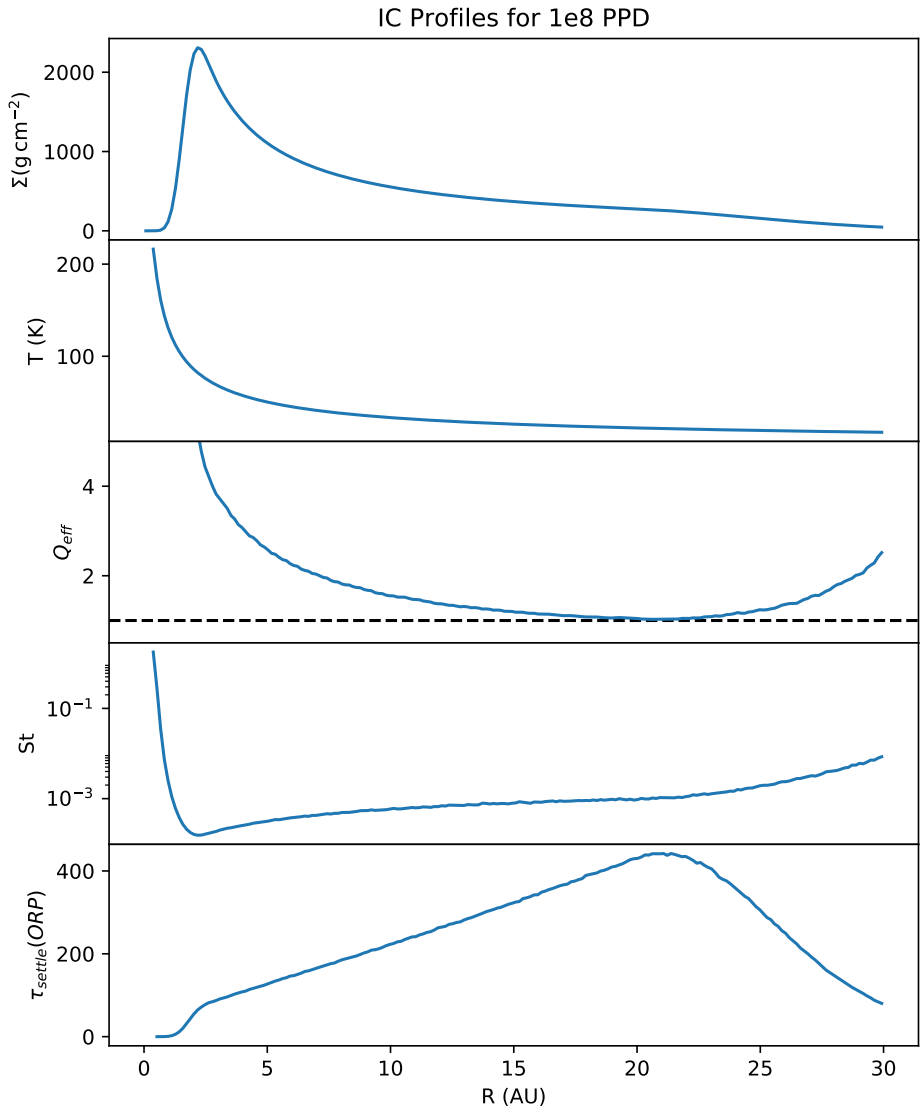



Figure 4.3: Radial profiles for the 1e8 PPD ICs. Plotted (from top to bottom) are the total surface density (gas + dust), temperature, the effective Toomre $Q$ ($Q_{\rm eff}$), the vertically mass averaged Stokes number $\mathrm{St} \equiv t_s\Omega$, and the dust settling timescale $\tau_{\rm settle}$. The surface density is a powerlaw of $-1$ with an interior cutoff at 4 AU and a smooth gaussian cutoff from 20 to 30 AU. Disks with a minimum $Q_{\rm eff} < 1$ are prone to fragmentation under GI [**?**].

run non-isothermal PPDs **???**. Our own test results indicate very strong artificial viscosity driven heating, with temperatures doubling on outer orbital time scales.

For our primary, fiducial runs we chose a dust grain size of $s = 1$ mm, to represent physically reasonable grain sizes in early PPDs. Current instruments such as ALMA are sensitive to observing dust in this regime. For our disks, this grain size also obeys the short stopping time $t_s \ll \Delta t$ and low stokes St $\ll 1$ assumptions of this algorithm (see figure 4.3). To probe more dust settling times ($\tau_{\rm settle}$) we also ran several 5 mm and 20 mm runs. To test the impact of disk stability, we ran a disk with decreased gas mass and with dust size scaled to keep $t_s$ the same as for the 20 mm runs (run 1e6 VeryHiQ in table 4.1).

Since the approach used to model dusty-gas in this chapter is fairly new we have studied the resolution requirements of this algorithm in a PPD environment. The test results of §4.2.2 indicate that unprecedented resolution may be required to properly capture dust settling in a PPD. An additional concern is the noise present in this algorithm, a problem common to SPH methods employing second derivatives. Given the test results, we have chosen to run our fiducial 1 mm disk at a highest resolution of $10^8$ particles (run 1e8 in table 4.1). To understand the effects of resolution we additionally ran the 1 mm fiducial disk at $10^5$, $10^6$, and $10^7$ particles (1e5, 1e6, and 1e7 in table 4.1). Due to early fragmentation of the $10^7$ particle disk we also re-ran it at a slightly increased $Q_{\rm eff}$ (run 1e7 HiQ).

The dust settling test of **?** adopts $N_{\rm smooth} \approx 28$ for the number of nearest neighbors used for smoothing[2]. Scaling this naively from 2D to 3D gives $N_{\rm smooth} = 28^{3/2} \approx 148$. For our runs we adopted $N_{\rm smooth} = 128$. A neighbor count significantly larger than the often used $N_{\rm smooth} = 32$ seems necessary to reduce noise in the dust algorithm. We used the Wendland $C_4$ smoothing kernel to prevent particle pairing instabilities at this high neighbor count [**?**].

A gravitational force accuracy (node opening) criterion of $\theta_{BH} = 0.7$ was used. Timesteps are set by a Courant condition of $\eta_C = 0.3$ and an acceleration criterion of $\Delta t_i = \eta\sqrt{\frac{\epsilon_i}{a_i}}$ where $\epsilon_i$ and $a_i$ are respectively the softening and acceleration of a particle, and $\eta = 0.2$. As with **?**

[2] While **?** use a variable number of nearest neighbors, we use a fixed number of neighbors for smoothing operations.

and **?**, we set the gravitational softening length to $0.5\,\langle h \rangle$, where $\langle h \rangle$ is the SPH smoothing length, calculated over the $N_{\rm smooth}$ nearest neighbors, and averaged over all particles in the simulation.

## *4.4 Run results*

### *4.4.1 Gas dynamics*

As discussed in §4.3 an initial $m = 2$ perturbation was applied to the disks presented here to ease comparison across resolutions spanning 3 orders of magnitude. These initial small perturbations rapidly grow (within fewer than 1 ORP) to become well defined $m = 2$ spiral arms extending to the edge of the disk. The spiral pattern speed is uniform across most of the disk and is not strongly resolution dependent for the $m = 2$ mode. For the 1e8 disk, $\Omega_s/\Omega_{\rm ORP} = 0.57$ where $\Omega_{\rm ORP}$ is the keplerian angular velocity at $R_d = 20$ AU and $\Omega_s$ is the spiral angular velocity. This places co-rotation at $R = 29$ AU, which is also the outer edge of the disk where $\Sigma$ has significantly decreased due to the external cutoff (see figure 4.3). Once the spiral arms develop, they extend out to corotation.

As has been reported before [**???**, §2.4] we find that gas dynamics in SPH simulations of PPDs can display strong resolution dependence, even up to typical resolutions of $10^6$ particles. Figure 4.4 compares the midplane total density $\rho$ at the 4 resolutions of the 1 mm dust simulations presented here at the end of the 1e8 run, after 9.6 ORPs. At $10^7$ and $10^8$ particles the disk is still composed of two well formed spiral arms, but at $10^6$ and $10^5$ particles the well defined arms vanish. In general, increasing resolution produces more coherent, well defined, and long-lived spirals. At lower resolutions, SPH noise can dominate the initial $m = 2$ perturbation which introduces significant higher mode power. However it is not clear if the different results are due simply to increased noise at lower resolutions. The effects of resolution dependent SPH features on PPD evolution, such as artificial viscosity, smoothing lengths, or gravitational softening lengths, remain poorly characterized. Additionally, the depth of perturbations in $\rho$ due to SPH poisson noise in the ICs is approximately res-

olution independent for fixed $N_{\rm smooth}$: what changes is primarily the length scales of the perturbations.

For the 1e8 disk, the $m = 2$ mode dominates throughout the simulation and higher order modes (except harmonics) do not grow relative to the $m = 2$ mode. The spiral arms remain coherent and stable, with a nearly uniform pattern speed which extends from the inner regions of the disk to the outer regions. For the 1e7 disk, an $m = 3$ mode grows and the pattern speed begins to vary radially. This causes spiral arms to cross, resulting in small, very high density regions at the outer disk which eventually fragment, forming small gravitationally bound clumps. For the lower resolution disks, higher order modes grow strongly and more chaotic spiral arms develop, but no regions become locally dense enough to drive fragmentation.

It is perhaps surprising that the 1e7 run fragments after about 14 ORPs while the other disks do not fragment. The disk fragments in a manner consistent with that presented in Chapter 3. Generally, decreasing resolution tends to make an isothermal SPH disk more likely to fragment, however in our case the 1e5 and 1e6 disks do not fragment. These runs are all very close to the $Q_{\rm eff} = 1$ fragmentation boundary, and the 1e7 run is just unstable/noisy enough to enter a chaotic phase and has sufficiently high resolution that small regions may have enough density (due to decreased smoothing length) to fragment.

To study the 1e7 disk at later times, we ran it again with a slightly higher $Q_{\rm eff} = 1.1$ (1e7 HiQ in table 4.1) by decreasing the gas mass slightly. This run did not fragment. Similar to the 1e6 and 1e7 runs it enters a chaotic phase, but rather than fragment it recovers and returns to an $m = 2$ dominated state again. During the chaotic phase, radial mass transport increases, spreading disk mass and decreasing $\Sigma$, placing the disk into a higher $Q$ state. To roughly quantify the mass transfer rate, we estimate a disk spreading time-scale by defining the disk width $w$ as the mass-weighted standard deviation of $R$ and defining the spreading timescale as $\tau_{spread} = w/\frac{{\rm d}w}{{\rm d}t}$. We estimate $\frac{{\rm d}w}{{\rm d}t} \approx \Delta w/\Delta t$ over a time scale of a few ORPs. At the beginning, chaotic, and final stages this gives a $\tau_{spread}$ of roughly 300, 60, and 1100 ORPs, respectively (equal to 0.3, 0.06, and 1.1 Myr). At the end of the chaotic phase, the

spiral power extends over a smaller region, mainly out to about $R = 15$ AU, and persists until the end of the run (68 ORPs).

We also detect long lived low Mach number shocks at the edges of spiral arms. The shocks grow in strength as resolution increases and the spiral arms become more well defined. For an isothermal shock, the Mach number is given by $\mathcal{M} = \sqrt{\rho_{g,2}/\rho_{g,1}}$. In the 1e8 disk this yields typical Mach numbers of $\mathcal{M} \approx 1.2$. Nowhere do we see the strong shocks required for chondrule formation, even in our strong and well defined spiral arms at very high resolution. For the inner regions of our disk, this gives shock speeds of $v \approx 1\ km/s$. For comparison, **?** find shocks of $v = 5km/s$ may form chondrules and **?** argue for shocks of order $v = 8km/s$.

We also observe that spiral arms have significantly decreased scale heights. Figure 4.5 shows the scale heights of our 1 mm disks. Here, we define the scale height as the vertically mass weighted mean of $|z|/R$. Comparison with figure 4.4 demonstrates that the high density regions of spiral arms have a decreased scale height, an effect which is present at all resolutions for the various spiral arm configurations. The post-shock region in a self-gravitating isothermal disk is expected to compress vertically (see **?** discussion around eq. 4.56), which is what we find here. However we are in the weak shock regime which is poorly studied analytically. Typical pre/post-shock scale height ratios are of order 1.2, similar to our Mach number. We also observe small, sharp jumps in the disk height of around 10% at the shock, which is only expected for adiabatic shocks.

### *4.4.2 Spiral arm enhancement*

We find that spiral arms can drive dust enhancement at rates comparable to or greater than dust settling or radial migration. Figure 4.6 shows the vertically mass-averaged dust fraction for the four 1 mm dust runs after 9.6 ORPs. This is equivalent to $\Sigma_d/\Sigma$. We find that the structure of $\epsilon$ in $R$ and $\theta$ does not follow the gas spiral arm structure and is resolution dependent, requiring very high resolution to be captured.

At $10^5$ and $10^6$ particles almost no structure is visibile along $(R, \theta)$ but rather is dominated by noise. In general, $\epsilon$ displays significantly more noise than the gas does (compare figures 4.4

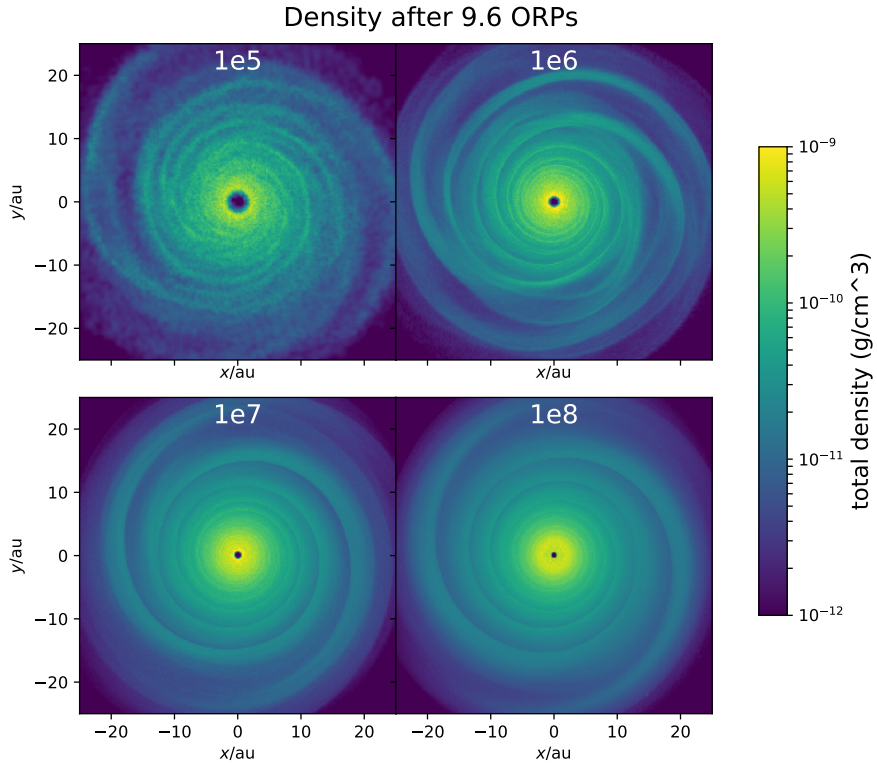



Figure 4.4: Midplane total density $\rho$ after 9.6 ORPs (the end of the 1e8 run) for the 1 mm simulations at 4 resolutions, listed in table 4.1. As resolution increases, the $m = 2$ spiral mode becomes increasing well defined, coherent, and long-lived. At $10^6$ and fewer particles, SPH noise and possibly other resolution effects drive the disk into a more chaotic state with higher order modes contributing significantly. The $10^7$ disk is beginning to enter a chaotic state at this point, and after several more ORPs it fragments. The $10^8$ disk shows no signs of entering such a state. The $m = 2$ pattern speed and winding angle are not strongly resolution dependent.

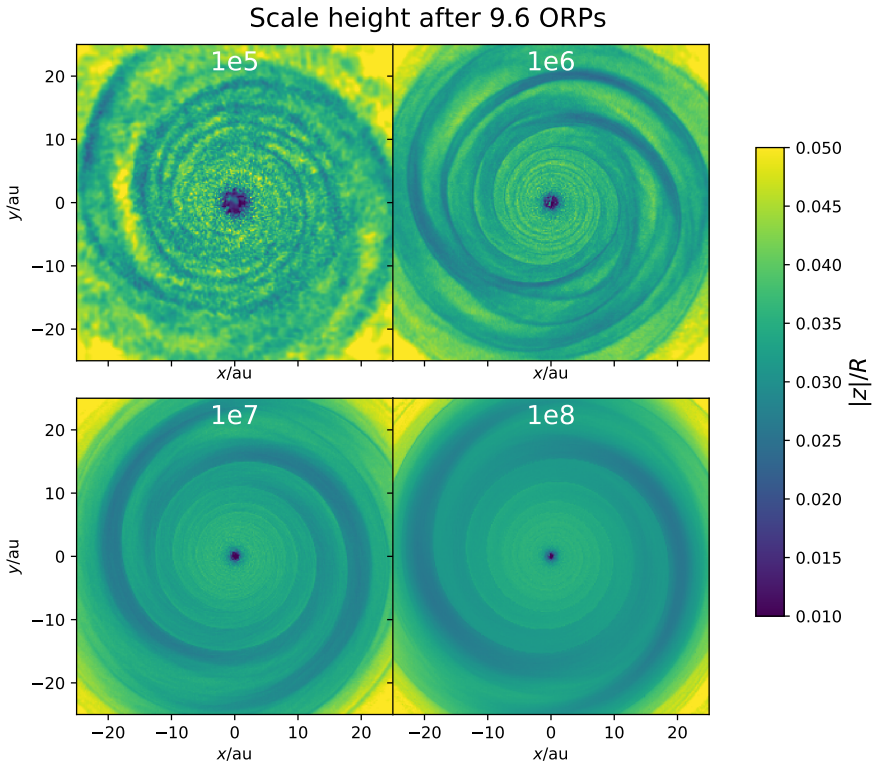



Figure 4.5: Scale height for the total mass (gas + dust) after 9.6 ORPs (the end of the 1e8 run) for the 1 mm simulations at 4 resolutions, listed in table 4.1. The scale height is defined as $|z|/R$, mass weighted and vertically averaged. The scale height is highly correlated with spiral arms, decreasing in the spiral arms due to self-gravity. Short, rapid spikes in the scale height are visible at the edges of spiral arms, particularly at higher resolution in the 1e8 run.

and 4.6). By $10^7$ particles, ring-like structures are visible in $\epsilon$ which become more prominent at $10^8$ particles.

The growth of these ring-like structures is generated by small, localized regions of dust enhancement around the shocks along the spiral arm edges. Figure 4.7 shows the dust density enhancement rate, defined as $|\rho\frac{\mathrm{d}\epsilon}{\mathrm{d}t}|$, where the time derivative is the standard material derivative used in SPH. Enhancement due to spiral arms increases as resolution increases and noise in the dust fraction decreases. Capturing structure along $(R, \theta)$ in $\epsilon$ requires extremely high resolution.

As the spiral arms pass, a given location will experience a brief period of dust enhancement. The sweeping motion of the spiral arms, coupled with the shearing of the disk which tends to un-wind spiral perturbations, results in the tightly packed ring-like structures. These rings are not stationary and grow in depth as the simulation progresses (i.e. the contrast in $\epsilon$ increases). We are unable to determine the ultimate outcome of these rings since they only form at very high resolution and we are unable to run such large simulations for sufficiently many dynamical times. Whether they continue to develop and create regions of highly concentrated dust or simply vanish due to shearing or some other dynamical effect remains to be seen.

These ring-like structures are also imprinted on the dust scale height. Figure 4.8 shows the ratio of the dust to gas scale heights, $h_d/h_g$. At high resolution, the ring structures are clearly visible. The dust scale height increases in regions of increased $\epsilon$. Dust gets concentrated and lofted upward by spiral arms. At $10^5$ and $10^6$ particles, $(R, \theta)$ structure in $h_d/h_g$ largely vanishes as the dust scale height tracks the gas scale height.

### *4.4.3 Dust settling and migration*

The disk scale height in the gaseous component is primarily due to pressure gradients, which the dust will not feel. Therefore, the dust is expected to settle toward the midplane. Dust settling should be a primary mechanism for concentrating dust in PPDs. We can estimate a dust settling time scale from the expression in eq. 4.15 for $v_d(z)$, the vertical velocity for

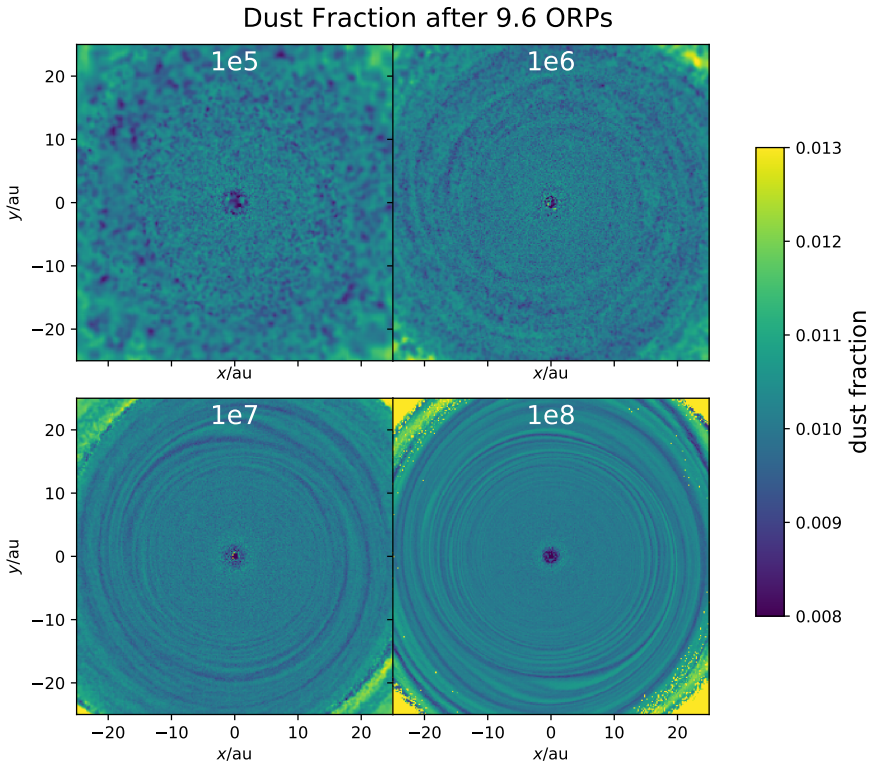



Figure 4.6: Vertically averaged dust fraction $\epsilon$ for the 1 mm simulations at 4 resolutions, listed in table 4.1. The dust enhancement is not co-located with the spiral arms. Structure in the $x - y$ plane increases with resolution. At $10^8$ particles, rings of dust-enhanced regions develop due to spiral arm driven dust concentration. This behavior is less prominent but still visible with $10^7$ particles, but vanishes at lower resolutions. By $10^5$ particles, noise completely dominates and no structure is visible in the $x - y$ plane.

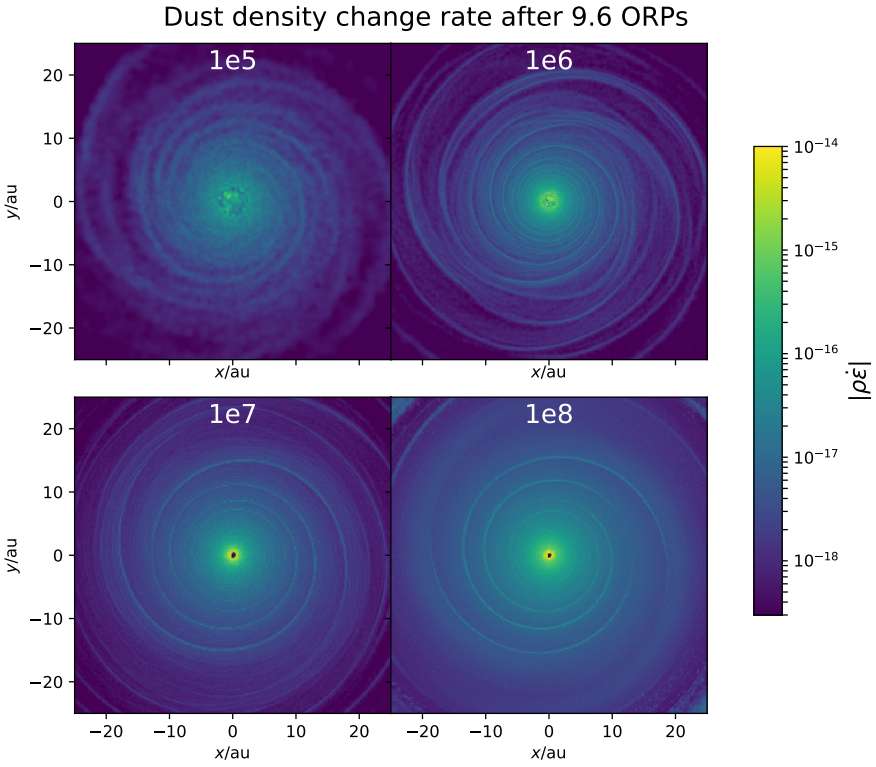



Figure 4.7: Absolute value of the dust density enhancement rate $|\rho\dot{\epsilon}|$, vertically mass averaged. The time derivative is a material derivative, i.e. on a particle basis. Spiral arms drive strong local dust concentration at the interface of two arms. Since spiral arms grow more coherent and well defined with resolution, this effect increases with resolution, however it remains visible even in the more chaotic spiral arms of the 1e6 run.

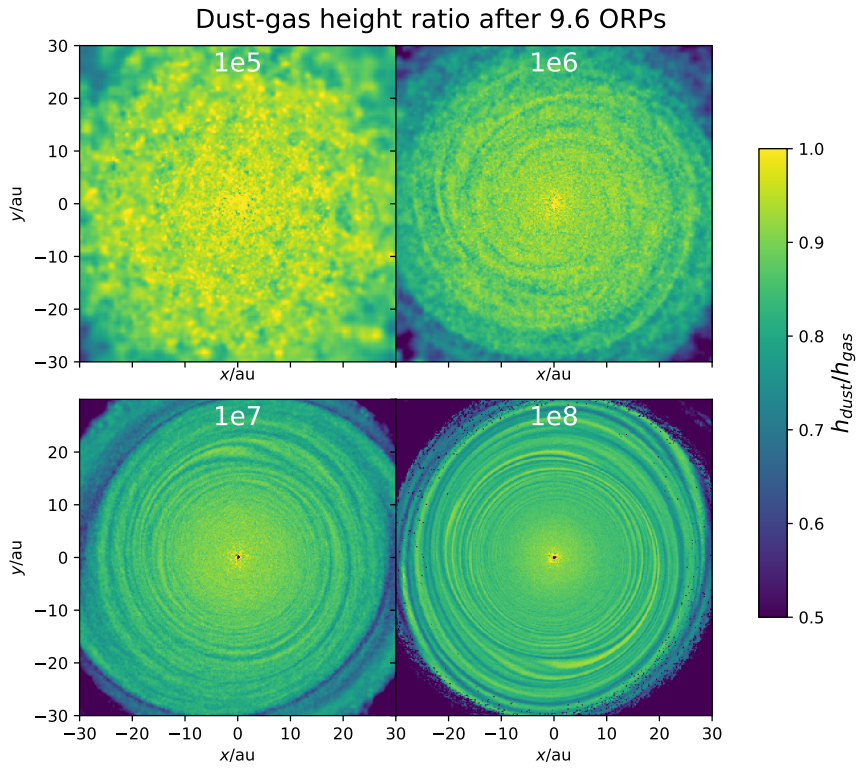



Figure 4.8: Dust height to gas height ratio $h_d/h_g$ for the 1 mm simulations at 4 resolutions, listed in table 4.1. Particularly at high resolution, structure in the $x-y$ plane largely follows the dust fraction structure (figure 4.6). With dust scale height increasing in regions where $\epsilon$ increases. As with $\epsilon$, at the highest resolution prominent ring-like structures form which are still visible but less prominent at $10^7$ particles, and which vanish at lower resolution. At $10^5$ and $10^6$ particles, the dust to gas height ratio is more uniform along $\theta$ as it tracks the gas height more strongly than at high resolution.

dust tightly coupled to the gas settling in a gaussian vertical gas density profile. Letting the dust settle at a characteristic velocity $v_d(H)$ from a characteristic height of $H = c_s/\Omega$ we obtain a dust settling timescale of

$$\tau_{\rm settle} = \frac{1}{e^{1/2}\Omega^2 t_{s,0}} \tag{4.19}$$

where $t_{s,0}$ is the midplane dust stopping time (see eq. 4.14). Figure 4.3 (bottom panel) shows $\tau_{\rm settle}$ vs $R$ for the 1-mm ICs. The mass weighted mean settling time is about $\tau_{\rm settle} \approx$ 220ORP.

Figure 4.9 shows the mass weighted mean dust-gas height ratio $h_d/h_g$, averaged out to $R = 30$ AU as a function of time for the 1 mm dust runs. The dust settling rate is resolution dependent and increases with resolution. For the 1e5 disk, settling has nearly ceased by the end of the run, saturating at a mean dust to gas height ratio of $h_d/h_g \approx 0.75$, comparable to the ratio reported in **?** for dust grains of a similar size in a circumbinary disk. This settling timescale for the 1e5 disk is much shorter than predicted by eq. 4.19 and the limiting is due entirely to numerical SPH effects, in particular increased SPH noise and increased smoothing lengths. The settling rate for the first 10 ORPs appears to converge by $10^7$ particles and agrees with the analytic approximate mean settling time. Settling continues until the end of all the 1 mm runs at resolutions above $10^5$ particles. During the chaotic spiral transition period present in the 1e6 and 1e7 HiQ disks around 25 ORPs, settling is strongly limited.

Probing dust settling requires high resolution disks. This introduces high computational cost and therefore reducing the $\tau_{\rm settle}$ is needed to explore settling at reasonable cost. To achieve that, we performed a series of runs with increased grain sizes of $s = 5$ mm and 20 mm, which reduce $\tau_{\rm settle}$ by factors of 5, and 20 respectively to $\tau_{\rm settle} = 44$ ORPs and 11 ORPs. Unfortunately, these decreased settling times introduce dust mass conservation errors, particularly at large radii (see §4.5). These errors restrict our analysis to the interior of the disk (out to about $R_d = 20$ AU) and to short time-scales.

With those caveats, we can still note some general dust settling behavior. Figure 4.10 shows the mass weighted mean dust-gas height ratio $h_d/h_g$, averaged out to $R = 20$ AU as

a function of time for several runs with decreased $\tau_{\rm settle}$: 1e7 5 mm, 1e7 20 mm, 1e6 20 mm, and the 1e6 VeryHiQ run. The 1e6 VeryHiQ run has the same $\tau_{\rm settle}$ as the the 20mm runs. The 1 mm 1e7 HiQ run is included for comparison. As expected, early in the simulation for a given resolution the settling rate is approximately proportional to the grain size, as can be seen by comparing the curves for the $10^7$ particle runs. The settling rate also increases with resolution. The 1e6 20mm run settles at a similar rate to the 1e7 5mm run, even though by physical reasoning the settling rate is expected to increase by a factor of 4 due to the increased grain size. Dust settling is also strongly limited at the end of the lower resolution $10^6$ particle runs, although it is not yet stalled as with the 1e5 1 mm run (figure 4.9).

One possible explanation for the reduced settling observed at lower resolutions is an increase in chaotic spiral power and turbulence-like SPH noise on length scales comparable to the disk scale height. To test this we ran a very stable $10^6$ particle disk (1e6 VeryHiQ in table 4.1), produced by increasing the $Q$ of the 1e6 20 mm run by decreasing the disk mass and proportionally decreasing the grain size to keep $t_s$ and $\tau_{\rm settle}$ constant. This disk shows almost no non-axisymmetric power and very little chaotic behavior, and yet as shown in figure 4.10 it settles significantly more slowly than the 1e6 20 mm run. This indicates that spiral arms are not limiting the settling and that chaotic spiral arms driven by low resolution do not explain the decrease in settling that we observe at lower resolutions. Rather, low resolution directly limits settling.

### *4.5 Mass conservation*

In the absence of multi-stepping or the dust floor/ceiling, the dust mass in this algorithm would be exactly conserved. The dust fraction time derivative can be written as a sum of interaction terms $\dot{\epsilon}_a = \sum_b \dot{\epsilon}_{ab}$, where the terms are of the form $\dot{\epsilon}_{ab} = -m_b(G_{ab} - G_{ba})$. Since particle masses are constant with respect to time, $\dot{m}_{dust} = m\dot{\epsilon}$. The change in dust mass in particle $a$ due to this interaction is:

$$\frac{{\rm d}m_{dust,ab}}{{\rm d}t} = m_a \frac{{\rm d}\epsilon_{ab}}{{\rm d}t} = -m_a m_b (G_{ab} - G_{ba}) \tag{4.20}$$

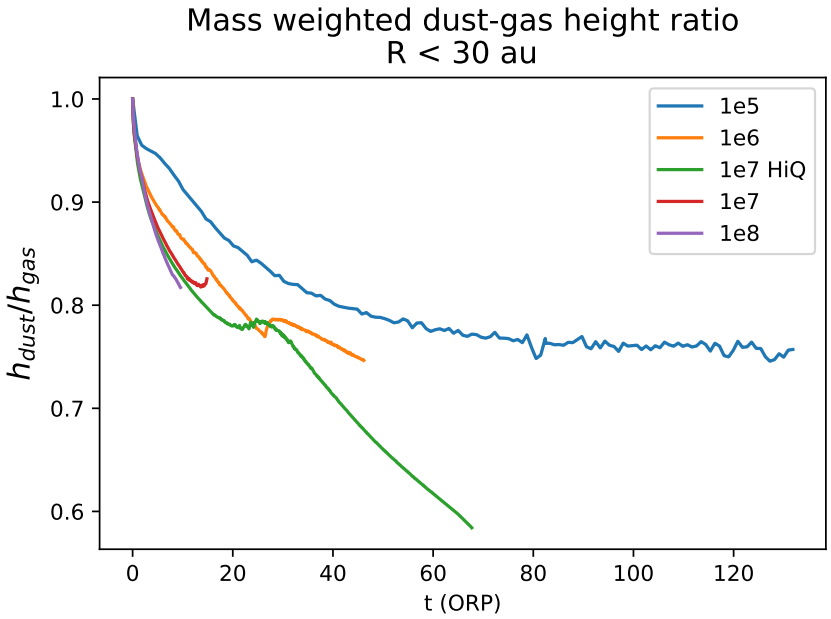



Figure 4.9: Mean dust to gas height ratio weighted by the dust mass as a function of simulation time for the simulations with 1 mm grains. Dust settling rates increase with resolution, appearing to converge by $10^7$ particles for the duration simulated (about 10 ORPs). By the end of the 1e5 run, settling has been largely stalled. The small up-tick at the end of the 1e7 run is due to fragmentation. The "bump" at around 25 ORPs for the 1e6 and 1e7 HiQ runs is due to the chaotic spiral transition period. The mass weighted settling timescale is $\tau_{\rm settle} = 220$ ORP for these runs.

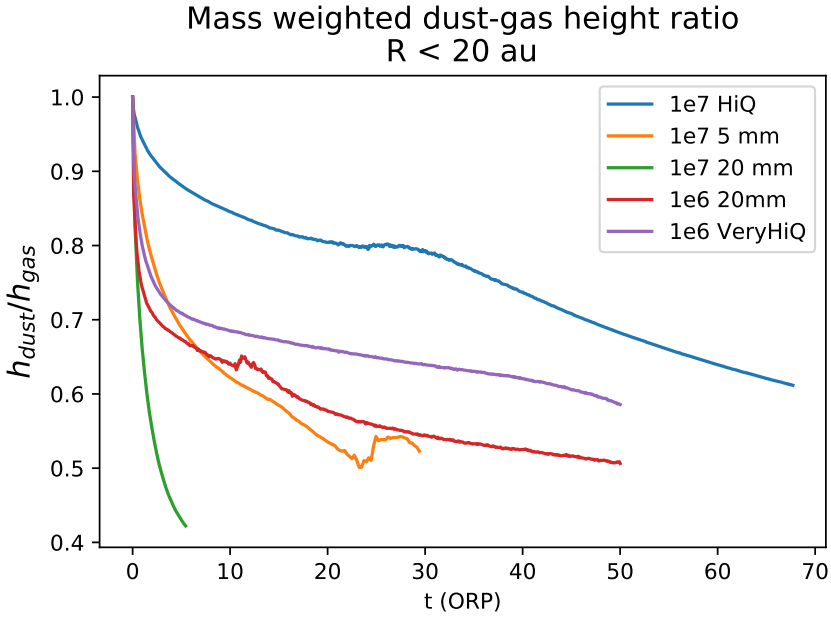



Figure 4.10: Mean dust to gas height ratio weighted by the dust mass for several simulations. The 1 mm 1e7 HiQ run is included for comparison. As expected, increased grain size increases the settling rate. However, as with figure 4.9, resolution increases the dust settling rate, which can dominate the effect of changing grain size. The 1e7 5 mm disk (expected $\tau_{\text{settle}} =$ 44ORP) measured settling rate is comparable to that of the 1e6 20 mm disk (expected $\tau_{\text{settle}} = 11$ ORP). The 1e6 VeryHiQ run has nearly zero non-axisymmetric power (no spiral arms). It has the same expected $\tau_{\text{settle}}$ as the 20 mm runs and does not settle faster than the gravitationally active 1e6 20 mm run.

Under particle exchange, clearly $m_a\dot{\epsilon}_{ab} = -m_b\dot{\epsilon}_{ba}$. Therefore this interaction adds nothing to the total mass. Any dust gained by $a$ is exactly canceled by dust lost by $b$.

Multi-stepping can break the dust mass conservation by changing when particle time derivatives are updated. In ChaNGa, multi-stepping is handled under the kick-drift-kick paradigm. A big time step $\Delta t_0$ is defined as a run-time parameter. Particles are dynamically assigned a time step $\Delta t_N = \Delta t_0 2^{-n}$ to obey time stepping constraints such as the Courant condition, gravitational accelerations, and Saitoh timestepping[**?**] to handle strong shocks. For a typical PPD multi-stepping reduces the total number of required time-steps by an order of magnitude or more.

Multi-stepping appears to be a minor contribution to dust-mass non-conservation, typically accounting for $10^{-2}$ or less of the total mass change. This is easily demonstrated by running a simulation with and without multi-stepping. In practice, the overwhelming majority of interacting neighbor pairs are on the same timestep. Additionally, particle time-stepping constraints mean that $\dot{\epsilon}_a$ should not vary much over the course of a time step, meaning that $\dot{\epsilon}_a$ and $\dot{\epsilon}_b$ will be roughly consistent in most cases. The majority of mass non-conservation is instead caused by requiring $\epsilon \in [0, 1]$.

### *4.5.1 Dust Fraction Limiting*

There is no guarantee that integrating the numerically estimated time derivative of $\epsilon$ will keep $\epsilon$ in the physically meaningful range of $[0, 1]$. To cope with this, we follow **?** and whenever $\epsilon$ leaves that range, we limit it by setting $\epsilon = 0$ or 1. In practice, we model environments with small dust fraction and therefore primarily encounter $\epsilon$ dipping below zero. This means that dust fraction limiting generates dust mass.

**?** also propose integrating a quantity $s \equiv \sqrt{(\rho\epsilon)}$ which guarantees $\epsilon = s^2/\rho \geq 0$. Initial tests of ours indicate that this can drive unphysical behavior as $s$ goes negative, driving oscillations in $\epsilon$. If we integrate $\epsilon$, the dust fraction going below zero is intrinsic to the algorithm, which we illustrate here in the case of 2D dust settling. These results can be extended by analogy to inward radial migration at the disk outer regions. The issue can be

| Simulation | $\Delta m_d/m_d$ |
|---|---|
| diffusion | 1.71e-13 |
| shock | 0.0 |
| wave | -1.22e-15 |
| settling HiRes | 0.11 |
| settling MedRes | 0.17 |
| settling LoRes | 0.24 |
| PPDs (9.6 ORPs) | |
| 1e5 | 7.3e-5 |
| 1e6 | -1.6e-2 |
| 1e7 | -3.0e-4 |
| 1e8 ($R < 30$ AU) | 0.13 (1.9e-3) |
| 1e7 HiQ | 1.2e-2 |
| 1e7 20mm | 3.3* |
| 1e7 20mm exp cut | 3.2* |
| 1e7 20mm hardcut | 2.6* |
| 1e7 5mm | 0.77 |
| 1e6 20mm | 0.17 |
| 1e6 VeryHiQ | 0.40 |

Table 4.2: Mass non-conservation for simulations and tests. The second column shows the fractional change in total dust mass. For the 2D dust settling tests, increasing the resolution decreases the mass non-conservation. For the PPD simulations, except for runs that finished earlier, the numbers are quoted after 9.6 ORPs, corresponding to the end of 1e8 run. Two numbers are quoted for the 1e8 PPD run–mass non-conservation for the entire disk and for $R < 30$ AU.

*These numbers are for the end of these runs, just before fragmentation. See table 4.1 for fragmentation times.

illustrated by considering particles at the vertical boundaries in the 2D dust settling test. For our algorithm, the contribution to $\dot{\epsilon}_a$ due to the interaction with $b$ is:

$$\dot{\epsilon}_{ab} = -m_b(G_{ab} - G_{ba}) \tag{4.21}$$

$$G_{ab} \equiv \frac{\epsilon_a(1-\epsilon_a)t_{s,a}}{\rho_a}\Delta\mathbf{f}_a \cdot \nabla W_{ab}(h_a) \tag{4.22}$$

Let particle $a$ be at the largest $z$ in its neighborhood such that $z_a > z_b$ for all neighbors $b$. After a dynamical time or so, particles at the boundaries lose dust and have a very low dust fraction, so it is safe to assume $\epsilon_a, \epsilon_b \ll 1$. Let us assume the gas is isothermal along $z$ such that $c_s =$ constant and let us assume the gas density follows a gaussian profile $\rho_g \approx \rho = \rho_0 \exp(-z^2/2H^2)$. Also, let the particle spacing $\Delta z_{ab} \equiv z_a - z_b \ll z, H$. In these limits, the stopping time can be written as $t_s = \tau/\rho c_s$ where $\tau$ is a constant. In general, the kernel gradient is inversely proportional to the smoothing length, and can be written as: $\nabla_a W_{ab}(h_a) = -(|F_{ab}|/h_a)\hat{\mathbf{r}}_{\mathbf{ab}}$, where $F_{ab}$ is some scalar, defined by that equation and is symmetric under particle exchange, $F_{ab} = F_{ba}$.

For SPH and gravity only (ignoring artificial viscosity), $\Delta\mathbf{f} = \nabla P/\rho_g = c_s^2 \nabla\rho_g/\rho_g$. For a gaussian profile, this reduces to $\Delta\mathbf{f} = -(c_s/H)^2\mathbf{z}$. Under all these assumptions, $G_{ab}$ simplifies to:

$$G_{ab} = \left(\frac{\tau c_s |F_{ab}|}{H^2}\right)\frac{\epsilon_a z_a}{\rho_a^2}\hat{\mathbf{r}}_{\mathbf{ab}} \cdot \hat{\mathbf{z}} \tag{4.23}$$

The smoothing length depends weakly on $\rho$ as $h \propto \rho^{-1/3}$ so we can use $h \equiv h_a \approx h_b$. Using $\hat{\mathbf{r}}_{\mathbf{ab}} = -\hat{\mathbf{r}}_{\mathbf{ba}}$ and $\hat{\mathbf{r}}_{\mathbf{ab}} = \hat{\mathbf{z}}$, we can write:

$$\dot{\epsilon}_{ab} = -m_b\frac{\tau c_s |F_{ab}|}{H^2}\left(\frac{\epsilon_a z_a}{\rho_a^2} + \frac{\epsilon_b z_b}{\rho_b^2}\right) \tag{4.24}$$

Since $a$ is at the boundary, $\dot{\epsilon}_{ab} < 0$ and $a$ will lose dust. Consider the terms $g \equiv \frac{\epsilon z}{\rho^2}$ in eq. 4.24. In the limit of infinite resolution, the particles would be very close and $g_a = g_b$. Putting this into the equation above tells us $\epsilon_a$ would exponentially damp since this entails $\dot{\epsilon}_{ab} \propto -\epsilon_a$. Similarly, $\epsilon_a$ will exponentially damp in the limit $g_a \gg g_b$. In the case of finite resolution $\epsilon_a$ may not asymptote to 0. Whenever $g_b$ is appreciably different from $g_a$, $\epsilon_a$ may

not damp to 0. Expanding $g$ gives $g_b \approx g_a - \Delta z_{ab} \frac{\mathrm{d}g}{\mathrm{d}z}$. For simplicity, let us assume $\epsilon$ also follows a gaussian with scale height $H_\epsilon < H$. The derivative of $g$ can be estimated as:

$$\frac{\mathrm{d}g}{\mathrm{d}z} = \frac{g}{z}\left[-\left(\frac{z}{H_\epsilon}\right)^2 + 2\left(\frac{z}{H}\right)^2 + 1\right] \tag{4.25}$$

The particles are at high elevations, so $z$ will be at several scale heights. Once the dust has settled to about $H_\epsilon < H/\sqrt{2}$, this derivative will be negative and $g_b > g_a$, meaning $\epsilon_a$ will decrease faster than exponentially and will be driven below zero, regardless of the time step. Therefore, dust fraction limiting will be required in any dust settling environment and this will cause mass non-conservation. For this idealized case, increasing resolution will mitigate the mass non-conservation. This analysis agrees with our 2D test results. Decreasing $\Delta t_0$ does not mitigate the mass non-conservation, while increasing the resolution can, as shown by the 2D settling test results in Table 4.2.

### *4.5.2 PPD mass conservation*

Table 4.2 shows the percent change in dust mass ($\Delta m_d/m_d$) over the course of the simulations for the tests and the PPD runs. For the PPD runs, dust mass accreted onto the star sink particle is accounted for. For the 2D dust settling tests, $\Delta m_d/m_d$ decreases with resolution, as predicted by the analysis in §4.5.1.

Mass non-conservation in the PPD runs is more complicated. In general, mass non-conservation for PPDs tends to worsen with increasing resolution, grain size, or $Q$. Comparing the 1 mm runs in table 4.2, all runs below $10^8$ particles demonstrate mass non-conservation on the percent level or below after 9.6 ORPs, but the 1e8 run shows a dust mass increase of 13%. Runs with larger grains show even worse mass non-conservation. Comparing the 1e7 5mm and 1e7 20mm runs shows that increasing grain size (which decreases $\tau_{\mathrm{settle}}$) drives more mass non-conservation. Comparing the 1e6 20mm and 1e7 20mm runs also shows that increasing resolution can greatly increase the mass non-conservation. We also ran a highly stable version of the 1e6 20mm run by decreasing the gas density and the

dust size by a factor of 75 (1e6 VeryHiQ), which keeps $t_s$ constant. The increased $Q$ run displays significantly worse mass conservation.

For PPDs, mass non-conservation is driven by dynamics at large radii. As with our test runs, multi-stepping contributes very little to mass non-conservation. In the turbulent, shearing, 3D, and glass-like environment of the PPDs there is very little mass non-conservation due to vertical dust settling at $R < R_d$. However, at large radii mass non-conservation can become severe. At the outer boundaries of the disk at large $R$, $\tau_{\rm settle}$ becomes short due to rapidly decreasing gas density. As a result, vertical dust settling occurs rapidly. This vacates dust from particles at higher elevations and drives mass non-conservation as described in §4.5.1 as particles with no dust try to lose dust. A similar effect happens along the radial direction: dust migrates rapidly radially inward at large radii. As with vertical settling, this results in boundary regions with no dust which try to lose dust, driving mass non-conservation. Modifying the outer boundary by applying a hard step function in $\Sigma$ at $R = R_d$ (1e7 20mm hardcut) or applying an exponential cutoff in $\epsilon$ at $R_d$ (1e7 20mm exp cut) does not improve mass conservation. Decreasing $Q$ or decreasing resolution makes these boundaries more noisy and less cleanly defined, which actually helps limit mass non-conservation.

Dust mass is then loaded at large radii and rapidly migrates radially inward. As gas density increases toward the center of the disk and stopping times decrease the dust migrates more slowly and piles up, forming dense regions at the outer edge of the disk. Figure 4.11 shows the dust mass per unit radius as a function of $R$ for several time steps for the 1e7 20mm run. The snapshots are plotted from start (dark) to end (light) of the run. Dust mass is loaded primarily from $R > 40$ AU. As it migrates inward (while being continuously loaded) the peak density increases, until at the final snapshot when the disk is driven unstable and fragments. Fragmentation is reflected in the jagged final profile.

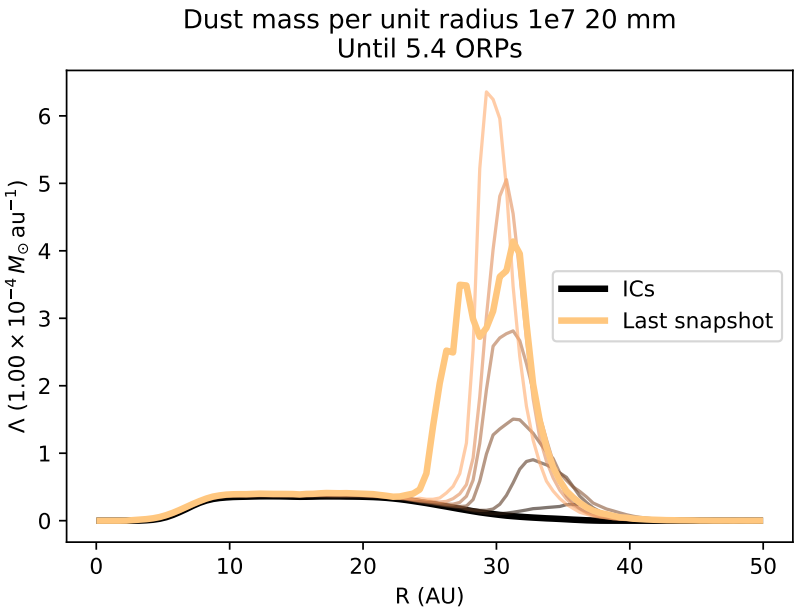



Figure 4.11: Dust mass per unit radius as a function of $R$ for the '1e7 20mm' run at several time steps. Dust mass is not conserved here, increasing by a factor of $\sim 4$. Dust mass is added at large $R$ (well beyond $R_d = 20$ AU) and migrates radially inward, piling up as it approaches regions of higher gas density where the dust radial migration velocity decreases. The peak of the mass profile grows until driving the disk to fragmentation. Fragmentation is evident in the jagged last snapshot profile.

## *4.6 Discussion*

In this chapter we presented SPH simulations of the gas and dust components of a young, gravitationally active PPD with spiral arms around an M-Dwarf. We used the one fluid dusty-gas SPH method of **?** to study the migration and concentration of millimeter-sized dust grains. This method, which uses a terminal velocity approximation, is attractive for being less noisy than many two-fluid methods and for potentially being fast by avoiding punishing time step criteria. However, its speed may be offset by strict resolution requirements.

The simple, 2D dust settling test of **?** hints at potentially strict resolution requirements for modelling dust settling in a realistic PPD. Resolution effects are often non-trivial and not intuitively obvious with hydrodynamic simulations, particularly with algorithms such as the dusty-gas method used here which are fairly new and poorly studied. Fortunately, a high resolution simulation can be run and analyzed at lower resolutions for almost no extra computer or human time: running a simulation at 0.1 and 0.01 of the maximum resolution introduces almost no overhead. In full, 3D PPDs, at resolutions of $10^5$ or $10^6$ particles, settling is strongly limited, with typical dust/gas scaleheight ratios of $h_d/h_g \approx 0.7$ or so. At higher resolutions, settling is not limited within the duration of simulations presented here, reaching values as low as 0.4 without indications of asymptoting. Globally, settling rates appear to converge by $10^7$ particles, at least for the stages presented here. At resolutions below $10^7$ particles, dust noise dominates any structure in $\epsilon$ in the $x-y$ plane. Limited settling, limited dust concentration, and a lack of $x-y$ structure mean that particularly at resolutions below $10^7$ particles this method is very similar to simply "painting" the dust on, i.e. running a gas-only simulation and assuming a constant dust fraction.

We also encounter significant dust mass non-conservation which depends on resolution and on the physical environment. This is due to poorly capturing the vacuum boundary conditions natural to a PPD simulation as explained in §4.5.1. The basic mechanism is that dust will tend to settle away from disk boundaries at large $z$ or $R$. Due to finite spatial resolution, particles at the boundaries will lose dust rapidly, driving their dust fractions

negative, regardless of time-step. Forcing $\epsilon \geq 0$ will load dust mass at the boundaries. This effect is worse with increasing resolution and decreasing spiral power. Smooth, non-chaotic boundaries suffer more strongly from this effect. This is an example of a more general difficulty: when a quantity is transported in a non-lagrangian way, boundary conditions must be handled. Properly handling boundary conditions (particularly vacuum boundaries) is generally difficult, a difficulty which is compounded in SPH where boundaries are not precisely defined. Such difficulties have been encountered in the case of photospheric cooling in SPH [**?**].

Changes to the algorithm as presented here may help address the mass non-conservation problem, however they will require more work. We have explored other approaches as well. It is possible to not limit the dust fraction and let it leave $[0, 1]$, but this can generate large waves in the dust fraction as it oscillates around 0. Any external term which drives $\epsilon$ to 0 is non-conservative or may drive oscillations. Other explicitly conservative approaches, such as exchanging dust mass with neighbors, are sensitive only to the gradient of $\epsilon$ (not the value of $\epsilon$) and do a poor job of limiting the dust fraction. **?** suggest integrating the quantity $s \equiv \sqrt{\rho\epsilon}$. This forces the dust fraction, calculated as $\epsilon = s^2/\rho$, to always be positive. However, our experiments indicate that this can lead to bad wave-like behavior. In regions where $\epsilon$ should be driven to 0 but integration over-shoots and $s$ is driven negative, we are now integrating a quantity allowed to go negative but which experiences a restoring "force" towards 0, an effect which may drive oscillations.

A different approach may be to "upwind" the dust fraction by calculating the direction of the pair-wise dust flow by the sign of terms in the dust fraction equation. SPH noise-driven diffusion is already a serious problem with this algorithm. Initial experiments with dust upwinding bear this out and indicate that mass non-conservation remains a problem in realistic PPDs.

For the gas dynamics of our runs, spiral arms are increasingly coherent and long lived as resolution increases. By $10^7$ particles, $m = 2$ arms appear long lived, which is even more pronounced at $10^8$ particles. Short lived, chaotic spirals due to e.g. variable pattern angular

velocity, often seen in PPD simulations, may often be due to low resolution.

In order to test the dusty-gas method used here with full PPDs, in a resolution independent way, we applied a 2%, $m = 2$ spiral perturbation, which mitigates perturbations driven by SPH noise. A low order, $m = 2$ perturbation provides the best resolution to spiral perturbation scale. Additionally, there are physical reasons to model $m = 2$ disks. Indeed, clear $m = 2$ spiral arms are visible in recent observations of Elias-27 using ALMA [**?**]. **?** argue that disk arms may be due an external perturber and/or gravitational instability.

An $m = 2$ perturbation can be formed via tidal interactions with an external perturber or distant binary companion, both of which may be commonplace for young PPDs. The tidal interaction for a distant binary companion can be estimated at $R = R_d$ as $\delta a/a = 2(R_d/R_p)^3 M_p/M_*$ where $R_p$ is the distance to the companion of mass $M_p$ and $\delta a/a$ is the fractional perturbation to the centripetal force for a particle orbiting the central star of mass $M_*$. For an equal mass companion, a 2% perturbation would place the companion at $R_p \approx 5R_d$, larger companions would push that radius out.

For 1 mm sized dust grains, we find that dust enhancement does not follow the spiral arms. Dust does not pile up at these pressure maxima as it has been shown to do in solar nebulae [**?**]. At the edges of spiral arms where shocking and convergent flow occur, dust gets enhanced. As the spiral arms pass by a given region, and as the dusty-gas mixture is sheared by the disk, dust-enhanced ring-like structures form in the disk. This effect is almost not present at typical PPD simulation resolutions of $10^6$ particles and becomes well defined at $10^8$ particles. Recent observations have revealed ring-like structures in dust SEDs of disks such as HL-Tau. Their origins are still unknown and debated and may be due to entirely different mechanisms such as clearing by protoplanets, but the general mechanism of local, non-axisymmetric dust enhancement by a rapidly moving perturburber (in our case spiral arms) may produce ring-like structures in many shearing disk environments.

Our very high resolution simulations allow us to probe dynamics of spiral arms in an idealized isothermal environment. Spiral arms do not strongly limit dust settling, although we do not probe down to low dust/gas scale-height ratios. Well defined weak shocks of

$\mathcal{M} \approx 1.2$ are present. There is no evidence for the strong shocks required for astrophysical processes such as chondrule formation. While necessary to capture dust and gas dynamics, high resolution does have the undesirable consequence of driving dust mass non-conservation.

A main issue which caused mass conservation issues here is the presence of the large range of dust settling times present in a PPD. Our method has trouble simultaneously modeling the inner and outer disk. The mass non-conservation encountered here is roughly equivalent to dust-mass loading at large $R$ and leads us to propose a mechanism for PPD fragmentation: Toomre instability driven by dust accretion at large $R$.

For outer regions of the disk, the dust radial migration velocity generally increases with $R$. From the continuity equation, this means that dust mass will "pile up" as convergent dust flow generates regions of increased density. This behavior is seen in Figure 4.11: as dust migrates inward rapidly from regions of low gas density it hits the disk and piles up. In the limit of small dust fraction in an isolated disk, **?** demonstrated that out regions of the disk may show enhanced dust due to such outwardly increasing dust migration velocity. We suggest that if allowed to accrete from the surrounding nebula, the dust may reach a sufficient density to drive the disk Toomre unstable at fairly large $R$ which can lead to fragmentation.

For simplicity, consider a keplerian disk where at some radius we have powerlaw profiles in $T$ and $\Sigma$ of $q$, $p$ respectively. Let us assume a terminal velocity approximation such that the dust radial velocity is given by $v_d = t_s \frac{\nabla P}{\rho_g}$. From the continuity equation, the inward dust-mass flux rate is $\dot{M}_d = -2\pi R \Sigma_d v_d$. Assuming densities (for the gas or dust component) are given by $\rho = \Sigma/H$ with a gas scale height of $H_g = c_s/\Omega$ and a dust to gas height ratio $h \equiv H_d/H_g$. Ignoring the effect of dust on the gas component, to achieve a Toomre $Q = 1$ for $\Sigma_d = \Sigma_g$ gives a mass loading rate of:

$$\dot{M}_d = (p + q/2 + 3/2)\frac{2\pi\sigma c_s^2}{\Omega}\left(\frac{h}{h+1}\right) \tag{4.26}$$

where we define $\sigma \equiv t_s \rho c_s$, which for our stopping time proscription and for an isothermal EOS is $\sigma = \sqrt{\pi/8}\rho_{\rm grain} s_{\rm grain}$. Note that this is not extremely sensitive to the steepness of the disk or the exact nature of the outer cutoff region. In general, dust radial migration rates

will increase outward, driving this behavior.

For the disk studied in this chapter, at $R = 30$ AU and assuming a dust/gas height ratio of $h = 0.1$ and powerlaws of $p = 2$ (appropriate for a disk cut-off region) and $q = 0.59$, this gives $\dot{M}_d = 8 \times 10^{-13} M_\odot$/yr or $2 \times 10^{-10} M_\odot$/orbit at 30 AU. This is fairly modest and represents the minimum rate at which the external medium would have to supply dust. Supplying dust faster would enhance the pile-up and drive instability more rapidly.

The mass required for fragmentation may roughly be estimated by assuming we need enough dust to achieve $\Sigma_d = \Sigma_g$ and $Q = 1$ in a region of radial extent $\Delta R = 1/k_T$ where $k_T = \kappa^2/2\pi G\Sigma$ is the most unstable Toomre wave number. Under these assumptions we get

$$M_d = \frac{2c_s^2 R}{G} \tag{4.27}$$

where $R$ is the radius at which the pile-up occurs. From this a time scale for the pile-up to grow to instability can be estimated as:

$$\tau_{\rm growth} = \frac{M_d}{\dot{M}_d} = \frac{\Omega R}{\pi G\sigma}\left(\frac{h+1}{h}\right)(p + q/2 + 3/2)^{-1} \tag{4.28}$$

A growth time scale may also be calculated by estimating $\dot{\Sigma}_d$ from $v_d$ and the continuity equation which yields:

$$\frac{1}{\Sigma_d}\dot{\Sigma}_d = \left(\frac{3}{2} + \frac{p\Sigma_g}{\Sigma_g + \Sigma_d/h}\right)\frac{v}{R} \tag{4.29}$$

The growth timescale to $\Sigma_d = \Sigma_g$ and $Q = 1$ can be estimated in the limits $\Sigma_d \ll h$ and $\Sigma_d \gg h$. For simplicity, taking the harmonic mean of $\tau_{\rm growth}$ in both limits gives:

$$\tau_{\rm growth} = \frac{8\Omega^2 R^2}{\pi G\sigma c_s}(2p + q + 3)^{-1}(2p + 3 + 3h)^{-1} \tag{4.30}$$

For typical parameters, eqs.(4.28) and (4.30) yield similar values of $\tau_{\rm growth}$.

To roughly estimate a reasonable range of values for the $M_d$ and $\tau_{\rm growth}$, we performed a simple monte carlo test. Grain density, the temperature at 1 AU, $p$, and $q$ were all drawn randomly from uniform distributions of $\rho_{\rm grain} \in [1, 3]\ g/cm^3$, $T_0 \in [100, 300]\ K$, $p \in [0, 3]$, and $q \in [3/7, 3/4]$. Stellar mass, grain size, $h$, and $R$ were all randomly drawn from distributions

uniform in log space (i.e. $\log M$ was uniformly distributed) in the ranges of $M_* \in [0.1, 2] M_\odot$, $s \in [1\mu m, 10mm]$, $h \in [0.01, 1]$ and $R \in [20, 300]$ AU.

Calculating $10^6$ trials gives a total dust mass required of $M_d = 2 \times 10^{-2}(8 \times 10^{-3} - 7 \times 10^{-2})\ M_*$, where the numbers in parentheses provide the 1 standard deviation range (in log-space) and $M_*$ is the mass of the star used in the trial. If dust mass can be accreted more rapidly, much less mass may be required. The growth time as calculated in eq.(4.28) is $\tau_{\rm growth} = 3 \times 10^6\ (2 \times 10^5 - 6 \times 10^7)\ yr$ and the time in eq.(4.30) is $\tau_{\rm growth} = 5 \times 10^6\ (2 \times 10^5 - 8 \times 10^7)\ yr$. These growth time-scales represent the amount of time a disk has for the instability to form. Up to a limit, longer $\tau_{\rm growth}$ may make the instability more likely to form, as short $\tau_{\rm growth}$ requires loading mass rapidly to drive the dust pile-up. If dust can be supplied faster (as may be the case since $\dot{M}_d$ increases radially outward) growth of the pile-up will be much shorter.

We can place some cuts on the accreted dust mass and growth time-scale to estimate disk parameters which may allow dust pile-ups. Disk life times are Myrs [**???**] so we select a rough cut of $10^4 yr < \tau_{\rm growth} < 10^7 yr$. Assuming $\epsilon = 0.01$ for the surrounding medium and that at most we can deplete dust from a gas mass of $M_*$, we select $M_d < 0.01 M_*$. Selecting trials according to these cuts does not strongly change the distribution for disk parameters except for grain size and stellar mass, indicating a weak dependence on those parameters. Under these cuts, the dust pile-up fragmentation may form for $M_* > 0.3 M_\odot$ and $s > 40\mu m$. This is a plausible mechanism for realistic PPDs, but it requires a much more in-depth investigation.

Dust pile-ups may also be able to promote the collapse of clumps once fragmentation by Toomre instability has occurred. At sufficient $\epsilon$ the dust component may act as a pressure-less heat sink which can reduce the effective $\gamma$, reducing pressure support of gravitationally bound clumps. If the effective $\gamma$ can be driven below unity, direct collapse by Jeans instability becomes a possibility. However, the algorithm presented here will not realistically be able to probe this regime, due to mass non-conservation and due to resolution constraints which render the long integration times required computationally infeasible.

We find that the one-fluid, terminal velocity, dusty-gas method used here has limited use in PPDs due to resolution requirements and mass conservation issues. Within those limits, we find that dust settling in PPDs is not limited by spiral arms. Spiral structure can drive rings of dust enhancement, however dust does not build up in spiral arms. This has an important observational consequence: when spiral density features are observed in the dust component, they will be at least as well defined in the gas. Axisymmetric features in the dust, however, may not be as strong in the gas.

# Chapter 5

# GRAIN GROWTH

We have so far restricted our dusty-gas models to track a single grain size. A realistic disk will have a distribution of grain sizes, and importantly grains will grow, through sticking during collisions, and shrink by fragmenting under high velocity collisions. The details of grain growth and destruction are a topic of on-going research [**????**]. Dust growth can have important dynamical effects, in particular because dust migration is controlled to a large degree by grain size.

A recent paper by **?** proposed a mechanism which may overcome the fragmentation and meter barriers which they call self-induced dust traps. Dust may grow rapidly and without destruction in outer regions of the disk where relative velocities of dust particles during collisions are low enough to avoid fragmentation, even as grains grow to cm size. As the grains grow larger they will migrate radially inward. If they are large enough they will be sufficiently decoupled from the gas that stirring due to gas will be minimal and their relative velocities will be low enough to avoid fragmentation. As they approach the inner regions of the disk, dust may pile-up (similar to dust pile-ups in the outer regions of the disk which I proposed in §4.6). Piling-up may be promoted by increased gas density and by dust fragmentation at inner regions (relative velocities will tend to increase as $R$ decreases) as both will decrease $t_s$ and therefore the radial migration velocity. If $\rho_d$ can grow sufficiently (of order $\rho_g$), drag on the gas due to the dust may clear some gas from a small radial extent. This may create a local pressure maximum. Roughly speaking, the dust radial migration velocity $v_{r,d} \propto \frac{\mathrm{d}P}{\mathrm{d}r}$, and therefore dust tends to migrate toward pressure maxima. This can further clear the nearby regions of gas. This can stall radial migration, even for larger grains, which potentially solves the meter barrier. Additionally, due to decreased $\rho_g$, the relative

velocities of collisions will decrease, potentially solving the fragmentation barrier.

They demonstrated the formation of self induced dust-traps using a simple two-fluid dust SPH implementation. While very promising, their implementation has several draw-backs. Two-fluid SPH dust implementations, including the one they use, suffer from strong SPH noise [**????**]. Additionally, two-fluid dust particles are known to clump very strongly due to numerical problems [**?**]. They also were only able to run at the quite low resolution of $2 \times 10^5$ gas particles, with an unreported number of nearest neighbors used. These numerical issues drive large noise in densities and may cause unrealistically high dust over-densities to form and are potentially the seeds of dust trap formation.

The dust traps form quite close to the artificial disk inner boundary. As orbital velocities and gas densities increase rapidly toward the inner disk, dynamical considerations force time-steps to decrease rapidly toward the origin. It is therefore necessary to not model the inner disk. This is typically handled by creating a hole in $\Sigma$ and treating the star as a sink particle. For many SPH PPD studies, e.g. when studying GI, the inner boundary is not very important since the inner region of the disk is very Toomre stable. However for this study the inner disk may be very important, and the effects of the inner cut-off are poorly studied.

Here, I present our initial tests and implementation of a one-fluid, terminal velocity, dust grain growth model. A one-fluid approach, especially in a fast code such as ChaNGa which can potentially run at very high resolution, has the potential to avoid these problems. While using our different approach may not provide refutation of the self-induced dust traps model, it does have the potential to provide more robust confirmation as it suffers from different problems and works under different assumptions.

While our tests indicate that a one-fluid SPH method may be poorly suited to studying self induced dust-traps, it does provide some further insight into the mechanism and the tests of **?**. In §5.1 I describe the grain growth model of **?** used here. I outline our implementation of one-fluid variable dust size, explain how we handled dust advection in §5.2, our upwinding scheme to improve mass conservation in §5.3, and an alternative to the parallel SPH "gather-scatter" approach **?** in §5.4. Test results for 2D settling tests are presented in §5.5 and

analysis of PPD simulations are presented in §5.6.

### *5.1 Grain growth model*

While the Gonzalez implementation uses a 2 fluid SPH implementation, ChaNGa's dusty-gas implementation is a single fluid. This will force us to use a slightly modified approach.

For the one fluid approach, the SPH particles have a single grain size ($s$) associated with them. This treats the grains as having a single size locally with spatial variance. SPH noise will allow a grain size distribution in small local regions (by looking at the distribution among nearby particles). This one fluid approach is simple because advection of solids of different sizes is automatically handled, and one only needs to time evolve $s$ locally. We, however, will also need to handle advection of dust between particles as dust is transported between particles.

Local grain growth is handled according to the simple model of **?**:

$$\left(\frac{\mathrm{d}s}{\mathrm{d}t}\right)_{\text{growth}} = \begin{cases} \frac{\epsilon\rho}{\rho_{\text{grain}}}\mathrm{V}_{\text{rel}} & (\mathrm{V}_{\text{rel}} < \mathrm{V}_{\text{frag}}) \\ -\frac{\epsilon\rho}{\rho_{\text{grain}}}\mathrm{V}_{\text{rel}} & (\mathrm{V}_{\text{rel}} > \mathrm{V}_{\text{frag}}) \end{cases} \tag{5.1}$$

Where $\mathrm{V}_{\text{frag}}$ is the fragmentation velocity and $\mathrm{V}_{\text{rel}}$ is a sub-grid average relative velocity between dust grains for a given SPH particle. This approximately represents the velocity dispersion of dust grains. There are many plausible candidates for reasonable sub-grid turbulent models to estimate $\mathrm{V}_{\text{rel}}$. For simplicity, we follow the model of **?**:

$$V_{\text{rel}} = \sqrt{2 * 2^{1/2}\mathrm{Ro}\ \alpha_{\mathrm{S}}}\frac{\sqrt{\mathrm{St}}}{1+\mathrm{St}}c_s \tag{5.2}$$

Where Ro is the Rossby number for turbulent flow, which they take to be 3 and $\alpha_{\mathrm{S}}$ is the dimensionless Shakura & Sunyaev viscosity parameter for a self-similar viscous PPD [**?**]. The typically used value for PPDs is $\alpha_{\mathrm{S}} = 0.01$, which we adopt but is certainly debatable. St is the Stokes number. As with **?** we take this to be approximately equal to the ratio of the dust stopping to orbital times $t_s/T$:

$$\mathrm{St} \equiv \frac{\Omega_K \rho_{\text{grain}} s}{\rho_g c_s} \tag{5.3}$$

where $\Omega_K \equiv \sqrt{GM_*/R^3}$ is the keplerian orbital angular velocity. Note that the stokes number increases with grain size. Due to the dependence on stokes number which in turn depends on $\Omega_K$, this turbulent model assumes that we are modeling a keplerian PPD. For disks which have undergone fragmentation, have strong spiral-arms, or which have highly enhanced dust regions this assumption may not hold.

For this simple model, the fragmentation velocity $\mathrm{V_{frag}}$ is a free parameter and is meant to represent the average velocity at which colliding grains fragment rather stick. In principle this should be a function of many things (grain size, grain composition, grain growth history, etc.) but we will follow Gonzalez and ignore that for simplicity and just make it a constant. They consider values of $\mathrm{V_{frag}} = (10, 15, 20, 25)$ m/s. In principle, reasonable values of $\mathrm{V_{frag}}$ should be measurable in laboratory experiments and may be informed by rubble pile simulations, but for now we should consider them quite unconstrained.

This model captures a few general features. The growth/destruction rates increase with increased grain volume density as more collisions occur. Collisions with very high relative velocities tend to break grains apart, resulting in decreased average grain size. Above the fragmentation threshold, increasing collision velocity tends to drive more fragmentation, thereby decreasing the average grain size more rapidly. The relative velocity between dust grains $\mathrm{V_{rel}}$ increases as gas turbulent velocities increase (these grow with $\alpha_\mathrm{S}$ and $c_s$). In the limit of very small grains (very small St) dust is tightly coupled to the gas and $\mathrm{V_{rel}}$ is very small and therefore the so is the growth rate. In the large grain limit the dust decouples from the gas and is not stirred by it. The Stokes number is small and again, $\mathrm{V_{rel}}$ and the growth rate are small.

Figure 5.1 shows a diagram of $\mathrm{V_{rel}}$ as a function of St. The dashed lines represent several different $\mathrm{V_{frag}}$ parameters. For simplicity, imagine a disk with a static profile. For a particle with a fixed location, St will grow as long as $\mathrm{V_{rel}} < \mathrm{V_{frag}}$. Let $V_{max}$ be the maximum $\mathrm{V_{rel}}$ (at St = 1). For $\mathrm{V_{frag}} > V_{max}$ (for example the top dashed line), particles will continue to grow un-impeded. For $\mathrm{V_{frag}} < V_{max}$ (the bottom dashed line), particles with very small stokes numbers will grow up to a maximum size when $\mathrm{V_{rel}} = \mathrm{V_{frag}}$. Particles with very large St will

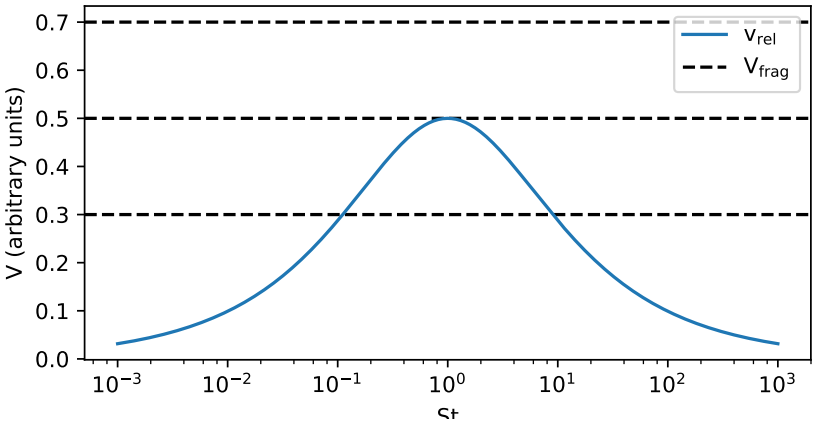



Figure 5.1: $V_{rel}$ as a function of stokes number (St) in arbitrary units. The dashed lines represent several different $V_{frag}$ parameters. The stokes number of particles with $V_{rel} < V_{frag}$ will grow and will decrease for $V_{rel} > V_{frag}$.

grow indefinitely. Particles with intermediate St (around St = 1) will have $V_{rel} > V_{frag}$ and will shrink to a minimum size where $V_{rel} = V_{frag}$.

## *5.2 Dust advection*

Since we have a one fluid approach and we transport dust between SPH particles, we need to also keep track of how much dust of a given size moves from one particle to the next. To achieve this, we update the particle grain size $s$ after advecting dust from neighboring particles as a mass weighted average of grain sizes. We can say a particle $a$ receives dust from a particle $b$ whenever $\dot{\epsilon}_{ab} > 0$. If a particle $a$ gets dust from some neighbors 1, 2, 3... we should update particle $a$'s new dust size to be:

$$s'_a = \frac{(m_{d,a} - |\Delta M_{d,a}|)s_a + \Delta m_{d,a1}s_1 + \Delta m_{d,a2}s_2 + ...}{(m_{d,a} - |\Delta M_{d,a}|) + \Delta m_{d,a1} + \Delta m_{d,a2} + ...} \tag{5.4}$$

Where $m_{d,a}$ is the mass of dust in $a$ at the beginning of the time step; $\Delta m_{d,a1}$ is the amount of dust given to $a$ by particle 1 (zero when $a$ gives dust to $b$); and $\Delta M_{d,a}$ is the mass of dust that $a$ loses to other particles during that time step, such that $\tilde{m}_{d,a} \equiv m_{d,a} - |\Delta M_{d,a}|$.

is the amount of its original dust that $a$ has after advecting dust. The term $\Delta M_{d,a}$ can be evaluated by adding up all the terms where $a$ gives dust to $b$. The denominator is just the total dust mass in $a$ after updating, $m'_{d,a}$.

It can be shown that this is equivalent to:

$$s'_a = \frac{s_a m_{d,a} + \sum_b \Delta m_{d,ab} s_{<a,b>}}{m'_{d,a}} \tag{5.5}$$

where we now take $\Delta m_{d,ab}$ to be the mass give to $a$ by $b$, positive when $a$ receives dust and negative otherwise. The index of $s$ is chosen according to:

$$< a, b > \equiv \begin{cases} a & \text{for } \Delta m_{d,ab} < 0 \\ b & \text{for } \Delta m_{d,ab} > 0 \end{cases} \tag{5.6}$$

This is a weighted mean, however the weights can be negative which will result in a poorly defined mean. Let $\mathbf{A}$ be the set of neighbors where $\Delta m_{d,ab} < 0$ (i.e. those that particle $a$ gives dust to) and let $\mathbf{B}$ be the set of neighbors where $\Delta m_{d,ab} > 0$ (i.e. those that particles $a$ gets dust from). Eq. (5.5) can be re-written:

$$s'_a = \frac{s_a(m_{d,a} - \sum_{b\in\mathbf{A}} |\Delta m_{d,ab}|) + \sum_{b\in\mathbf{B}} \Delta m_{d,ab} s_b}{(m_{d,a} - \sum_{b\in\mathbf{A}} |\Delta m_{d,ab}|) + \sum_{b\in\mathbf{B}} \Delta m_{d,ab}} \tag{5.7}$$

This is a weighted mean of $s_a$ and all $s_b$ terms. By definition, $\Delta m_{d,ab} > 0$ in the sum over $b \in \mathbf{B}$, so those weights are positive. Our weighted mean is guaranteed to be well behaved if the weight: $(m_{d,a} - \sum_{b\in\mathbf{A}} |\Delta m_{d,ab}|) > 0$. This is equivalent to saying $a$ must not give up all its dust. We therefore separately sum the positive and negative $\dot{\epsilon}_{ab}$ terms such that

$$\dot{\epsilon}_a = \sum_{b\in\mathbf{A}} \dot{\epsilon}_{ab} + \sum_{b\in\mathbf{B}} \dot{\epsilon}_{ab} = \dot{\epsilon}_{a,loss} + \dot{\epsilon}_{a,gain} \tag{5.8}$$

We enforce $(\epsilon_a + \Delta t \dot{\epsilon}_{a,loss}) \geq \delta$ where $\delta$ is some very small number. Using $\Delta m_{d,ab} \approx \dot{\epsilon}_{ab} \Delta t$ we can update $s'_a$:

$$s'_a = \frac{s_a(\epsilon_a + \Delta t \dot{\epsilon}_{a,loss}) + \Delta t \sum_{b\in\mathbf{B}} \dot{\epsilon}_{ab} s_b}{(\epsilon_a + \Delta t \dot{\epsilon}_{a,loss}) + \Delta t \dot{\epsilon}_{a,gain}} + \Delta t \left(\frac{\mathrm{d}s_a}{\mathrm{d}t}\right)_{\text{growth}} \tag{5.9}$$

where the final term, given by eq. (5.1), is due to local grain growth. The mass weighted dust grain size $m\epsilon s$ is conserved by the advection term except when we enforce $(\epsilon_a + \Delta t \dot{\epsilon}_{a,loss}) \geq \delta$, although for small $\delta$ and with proper time-stepping this should be limited. Conservation can also be broken when forcing $\epsilon \in [0, 1]$ or through multi-stepping. In practice, we find this non-conservation to be limited.

### *5.3 Dust upwinding*

An unfortunate effect of updating the grain size according the eq. (5.9) is that dust mass non-conservation can become particularly severe (see also §4.5). Initial tests with the standard 2D dust settling test indicate mass non-conservation, even with nearly uniform $s$ and local grain growth turned off, can be much more severe than for the uniform grain size method of **?**. To limit mass non-conservation we have implemented an upwinding scheme for the dust fraction.

Upwinding schemes use gradient estimators biased behind the direction of the flow or information propagation. In the simple case of a wave packet propagating along the $+x$ direction, an upwinding scheme would primarily use samples at $x < x_0$ to estimate a gradient at $x_0$. Such schemes are often used in advection-like scenarios, however they often suffer from strong numerical diffusion.

The interaction term between particles $a$ and $b$ in the dust fraction equation of the **?** one-fluid dust method are given by:

$$\dot{\epsilon}_{ab} = -m_b \left[ \epsilon_a \frac{(1-\epsilon_a) t_{s,a}}{\rho_a} \Delta \mathbf{f}_a \cdot \nabla W_{ab}(h_a) + \epsilon_b \frac{(1-\epsilon_b) t_{s,b}}{\rho_b} \Delta \mathbf{f}_b \cdot \nabla W_{ab}(h_b) \right] \tag{5.10}$$

which allows us to calculate $\dot{\epsilon}_a = \sum_b \dot{\epsilon}_{ab}$. This involves two terms, weighted by $\epsilon_a$ and $\epsilon_b$.

We introduce here an upwind alternative to this, using only the dust fraction of the upwind particle:

$$\dot{\epsilon}_{ab} = -\epsilon_{<a,b>} m_b \left[ \frac{(1-\epsilon_a) t_{s,a}}{\rho_a} \Delta \mathbf{f}_a \cdot \nabla W_{ab}(h_a) + \frac{(1-\epsilon_b) t_{s,b}}{\rho_b} \Delta \mathbf{f}_b \cdot \nabla W_{ab}(h_b) \right] \tag{5.11}$$

where:

$$< a, b >= \begin{cases} a & \text{for} \quad \dot{\epsilon}_{ab} < 0 \\ b & \text{for} \quad \dot{\epsilon}_{ab} > 0 \end{cases} \tag{5.12}$$

The term $\epsilon_{<a,b>}$ is the dust fraction in the particle from which the dust is coming. When dust is moving from $a$ to $b$ we use $\epsilon_a$. In this case, $a$ is on the "upwind" side (behind the direction of the flow), hence this is an upwinding scheme.

This scheme can be written formulaically as:

$$\begin{aligned} \dot{\epsilon}_{ab} &= -\epsilon_{<a,b>} m_b (G'_{ab} - G'_{ba}) \\ G'_{ab} &\equiv \frac{(1-\epsilon_a) t_{s,a}}{\rho_a} \Delta \mathbf{f}_a \cdot \nabla W_{ab}(h_a) \\ \epsilon_{<a,b>} &= [\epsilon_a \Theta(G'_{ba} - G'_{ab}) + \epsilon_b \Theta(G'_{ab} - G'_{ba})] \end{aligned} \tag{5.13}$$

where $\Theta$ is the heaviside step function. Dust mass is conserved by this method. The total dust mass time derivative due to this interaction is:

$$\frac{\mathrm{d}m_{d,ab}}{\mathrm{d}t} = m_a \dot{\epsilon}_{ab} + m_b \dot{\epsilon}_{ba} = 0 \tag{5.14}$$

by eq. (5.13). The corresponding energy equation is derived from conservation of total energy (thermal + kinetic), which in the absence of external heating/cooling terms reads:

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \sum_a m_a \left[ \mathbf{v}_a \cdot \frac{\mathrm{d}\mathbf{v}_a}{\mathrm{d}t} + (1-\epsilon_a) \frac{\mathrm{d}u_a}{\mathrm{d}t} - u_a \frac{\mathrm{d}\epsilon_a}{\mathrm{d}t} \right] = 0 \tag{5.15}$$

The first term in the sum is the kinetic energy term which represents PdV work and is the normal SPH $\dot{u}$ term. The third term is due to dust advection. There are several ways to formulate the energy equation which all satisfy the above equation and conserve total energy. A few alternatives are:

$$\frac{\mathrm{d}u_a}{\mathrm{d}t} = m_a \left( \frac{\mathrm{d}u_a}{\mathrm{d}t} \right)_{\mathrm{SPH}} + \frac{m_b \epsilon_{<a,b>}}{1-\epsilon_a} \times \begin{cases} G'_{ab}(u_a - u_b) \\ (G'_{ab} - G'_{ba}) u_a \\ -(G'_{ab} - G'_{ba}) u_b \end{cases} \tag{5.16}$$

Similar to the **?** scheme, we select the first option which can be understood as something analogous to a central finite difference scheme in $u$:

$$\frac{\mathrm{d}u_a}{\mathrm{d}t} = m_a \left(\frac{\mathrm{d}u_a}{\mathrm{d}t}\right)_{\mathrm{SPH}} + \frac{m_b \epsilon_{<a,b>}}{1-\epsilon_a} G'_{ab}(u_a - u_b) \tag{5.17}$$

This method introduces an implementation challenge, with an interesting generic solution. Certain asynchronously parallel SPH implementations do not allow for quickly and efficiently calculating such terms as $\Theta(G'_{ab} - G'_{ba})$. We discuss this below.

### *5.4 Gather-scatter*

ChaNGa uses the "gather-scatter" approach, originally presented in the **?** TREESPH paper. This method is a clever and efficient way to calculate SPH terms for particle $a$ (such as forces) due to an interaction with particle $b$ when they are a non-mutual particle pair. By this I refer to the situation where $b$ is a neighbor of $a$, but $a$ is not a neighbor of $b$ (see Fig. 5.2). In the gather-scatter approach, $a$ calculates terms for itself and for $b$ that are non-zero when $b$ is $a$'s neighbor, and vice-versa. An in depth explanation follows.

First, some definitions. We define the neighborhood of particle $a$ ($N_a$) as the region defined by where $a$'s kernel is positive. In ChaNGa this is defined as the region with a distance $2h_a$ of $a$, where $h_a$ is the smoothing length of particle $a$. Typically, the smoothing length is chosen such that there are a fixed number of particles ($N_{\mathrm{smooth}}$) within the neighborhood of $a$. The SPH kernel for an interaction between $a$ and $b$ is a function of the smoothing length and the distance between particles, which we can write as written as $W_{ab}(h_i) = W(|\mathbf{r}_a - \mathbf{r}_b|, h_i)$ where $h_i$ is the smoothing length of $a$ or $b$. This entails that the kernel is symmetric under particle exchange, i.e.: $W_{ab}(h_i) = W_{ba}(h_i)$. In ChaNGa we follow the convention:

$$W_{ab}(h_i) = 0 \quad \text{for} \quad |\mathbf{r}_a - \mathbf{r}_b| \geq 2h_i \tag{5.18}$$

From which we can say the kernel gradient is also negative outside the neighborhood of $a$:

$$\nabla_a W_{ab}(h_i) = -\nabla_b W_{ab}(h_i) = 0 \quad \text{for} \quad |\mathbf{r}_a - \mathbf{r}_b| \geq 2h_i \tag{5.19}$$

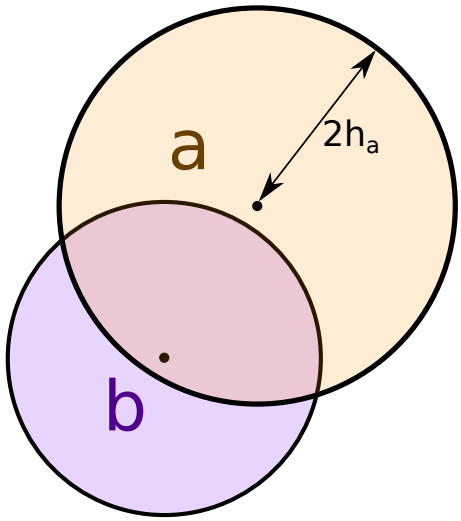



Figure 5.2: An example of a non-mutual particle pair. Particle $b$ is in the neighborhood of $a$, but $a$ is not in the neighborhood of $b$. The "gather-scatter" approach is a method to handle this situation in SPH when looping over neighbors.

Where $\nabla_a$ is the gradient taken with respect to particle $a$.

This can be summed by saying outside the neighborhood of a particle, the SPH kernel and the SPH kernel derivative go to zero, i.e.:

$$W_{ab}(h_a) = 0 \quad \text{for} \quad b \notin N_a \tag{5.20}$$

$$\nabla_a W_{ab}(h_a) = 0 \quad \text{for} \quad b \notin N_a \tag{5.21}$$

The kernel and its gradient typically show up in every term for the time derivative of an SPH quantity. For an arbitrary SPH quantity $x$ (such as velocity or internal energy), the time derivative is typically defined as a sum of the form:

$$\frac{\mathrm{d}x_a}{\mathrm{d}t} = \sum_b \dot{x}_{ab} = \sum_b F_{ab} + G_{ba} \tag{5.22}$$

where the sum is in general over all the particles in the simulation. $\dot{x}_{ab}$ is the interaction term between particle $a$ and $b$, and $F_{ab}$ and $G_{ba}$ are functions of the SPH kernel and/or its derivative and are zero when $b$ is outside the neighbor of $a$ such that:

$$|F_{ab}|, |G_{ab}| \begin{cases} \geq 0 & \text{for} \quad b \in N_a \\ = 0 & \text{for} \quad b \notin N_a \end{cases} \tag{5.23}$$

And of course the converse holds, i.e. $F_{ba} = 0$ for $a \notin N_b$.

The gather-scatter approach uses these properties to efficiently calculate the interaction terms $\dot{x}_{ab}$ and $\dot{x}_{ba}$, which are given by:

$$\dot{x}_{ab} = F_{ab} + G_{ba} \tag{5.24}$$

$$\dot{x}_{ba} = F_{ba} + G_{ab} \tag{5.25}$$

We start calculating the time derivatives by setting them to zero for all particles. We loop over all particles $a$ and their neighbors $b \in N_a$ and calculate the terms which are non-zero for $b \in N_a$, for both $a$ and its neighbor $b$, i.e.:

$$\dot{x}_a \mathrel{+}= F_{ab} \tag{5.26}$$

$$\dot{x}_b \mathrel{+}= G_{ab} \tag{5.27}$$

$$\tag{5.28}$$

When the outer loop over particles gets to particle $b$, if $a \in N_b$, this procedure will calculate the remaining terms:

$$\dot{x}_a \mathrel{+}= G_{ba} \tag{5.29}$$

$$\dot{x}_b \mathrel{+}= F_{ba} \tag{5.30}$$

If $a$ is not a neighbor of $b$ we won't add these terms, but they are zero anyway, so the result is the same. This ensures that the terms in eqs.(5.24) and (5.25) are all calculated and are calculated exactly once. Additionally, all these calculations can be done asynchronously in parallel in a lock-free manner since all the calculations above are addition and therefore associative. Gather-scatter pseudo-code is printed in algorithm 1.

**for** $a \in$ *all particles* **do**
  **for** $b \in N_a$ **do**
    Add interaction terms that are non-zero for $b \in N_a$.
    $\dot{x}_a \mathrel{+}= F_{ab}$
    $\dot{x}_b \mathrel{+}= G_{ab}$
  **end**
**end**

**Algorithm 1:** Gather-scatter method for calculating interactions of the form $\dot{x}_{ab} = F_{ab} + G_{ba}$. These calculations can be done in parallel asynchronously.

Calculating the time derivative as in eq.(5.22) assumes a loop over all particles. This method ensures that a particle only needs to loop over its neighbors, not over its neighbors and all particles for which it is a neighbor. This greatly simplifies the tree building/tree walking procedure and reduces the number of neighbors to loop over. As an added bonus, it usually also avoids calculating the kernel/kernel gradient twice, since typically $F_{ab}$ and $G_{ab}$ are functions of the same kernel.

### *5.4.1 Alternative gather scatter*

The above "gather-scatter" approach works for interaction terms of the form in eqs. (5.22) and (5.23), involving terms which are only a function of *one* particle's kernel/kernel gradient. If we want to extend this to terms which are a function of the kernels/kernel gradients of *both* particles, we need to change our approach. Consider terms that are only guaranteed to be zero when $a$ and $b$ are not neighbors:

$$H_{ab} = 0 \quad \text{for} \quad a \notin N_b \text{ and } b \notin N_a \tag{5.31}$$

For example, some advection schemes might follow a modified version of equation 5.24:

$$\dot{x}_{ab} = [F_{ab} + G_{ba}]\, \Theta(F_{ab} + G_{ba}) \tag{5.32}$$

where $\Theta$ is the heaviside step function. This is similar to the upwinding approach in eq. (5.13) The key difference with terms such as this is that we have to calculate the entire term all at once.

We cannot separately calculate the $F_{ab}$ and $G_{ba}$ terms asynchronously and combine them since they are not reduced under an associative operation such as addition. To efficiently calculate such terms, we can calculate the entire interaction at once with a condition which ensures it is only calculated once. This can be done as in the following pseudocode. See algorithm (2) below. A simplified equivalent implementation is included in algorithm (3).

```
for a ∈ all particles do
    for b ∈ N_a do
        if a ∉ N_b then
            // Non-mutual neighbors.  We can avoid double counting and
               safely calculate all interaction terms here since b's loop
               will not include a.  This is the situation in Fig. 5.2
            // We also know that h_a > h_b
            Calculate interaction terms for both a and b
            ẋ_a += H_ab
            ẋ_b += H_ba
        else
            // Mutual neigbors.  If we aren't careful, both a and b may
               calculate the interaction terms, resulting in a double
               counting.  This requires some check to ensure only one
               particle will calculate the terms.
            if h_a > h_b then
                Calculate interaction terms for both a and b
                ẋ_a += H_ab
                ẋ_b += H_ba
            else
                Do Nothing // Since these are both neighbors, a will be in
                   b's loop.  Except in extreme edge cases, we will have
                   h_b > h_a and this term will be calculated.  In the edge
                   case h_a = h_b, this term will be neglected.
            end
        end
    end
end
```

**Algorithm 2:** Alternative approach to "gather-scatter". This method has the advantage of allowing the entire interaction to be calculated at once, rather than splitting the calculation up between particles, but at the cost of introducing more if statements.

```
for a ∈ all particles do
    for b ∈ N_a do
        if h_a > h_b then
            // We can avoid double counting and safely calculate all
               interaction terms here since b's loop will not include a
            Calculate interaction terms for both a and b
            ẋ_a += H_ab
            ẋ_b += H_ba
        else
            // Do Nothing -- a and b are each others' neighbors and b will
               calculate the interaction.
        end
    end
end
```

**Algorithm 3:** Simple gather scatter alternative, a simpler equivalent to algorithm (2). Note that in the edge case $h_a = h_b$ we will miss a term. This edge case will only arise in SPH grid ICs with evenly spaced particles, before they have moved, and can be avoided by checking for equality and updating the particle with the lower particle ID.

### *5.5 2D dust settling tests*

Here we present the results of 2D dust settling tests using the terminal velocity one-fluid upwinding dust scheme presented here. We present results for the fixed grain size **?** 2D dust settling test and for a variant with variable grain size to test the dust grain advection which we call the "top-hat" test.

### *5.5.1 PL15 settling test*

The **?** settling test is detailed in §4.2.2. Briefly, it is a 2D test in $x - z$ in a fixed external potential along $z$ which represents a patch of a PPD. The ICs are in equilibrium and are a nearly gaussian vertical density profile with a uniform dust fraction of $\epsilon \approx 0.01$ and uniform grain size $s = 1$mm and a scale-height of $H = 0.252$. The test is run at 3 resolutions: 896 (LoRes), 3552 (MedRes), and 14208 (HiRes) particles. For all the tests presented here, $N_{\rm smooth} = 28$ and the Wendland $C_4$ kernel were used. Since the grain-size is uniform in this test, it demonstrates the effect that upwinding has on dust advection.

The primary motivation for introducing the upwinding scheme was to promote mass conservation. It should be noted that while the upwinding method in eq. (5.13) is conservative, the full approach is not since the dust fraction is forced to stay within the region $\epsilon \in [0, 1]$. This is discussed in detail in §4.5.1. With variable grain-sizes (before implementing upwinding), our approach showed significantly worse mass non-conservation at the $z$ boundaries of the 2D settling test. Additionally, up-winding is more conceptually coherent, since our dust size advection scheme transports grain sizes from one particle to another, in a fixed direction. The dust size is therefore upwinded and upwinding the dust fraction as well makes intuitive sense. Figure 5.3 plots the fractional change in dust mass for this test at all 3 resolutions for the upwinding and non-upwinding implementations. Upwinding implementations are marked with a "-gg" suffix. Without upwinding, after 50 ORPs the dust mass increase by 7-25%, increasing with decreasing resolution. With upwinding, the fractional change is limited to the percent level.

While the upwinding scheme shows marked improvement in mass conservation, as an upwinding scheme it does suffer from numerical diffusion. Figure 5.4 shows the vertical dust density profiles at several epochs for the **?** dust settling test using the one-fluid terminal velocity, fixed grain size method (fig. 5.4a) and for the variable grain-size, upwinding method presented in this chapter (fig. 5.4b). The upwinding scheme is certainly more diffusive–the double humped shape is much more prominent without upwinding, although the overall set-

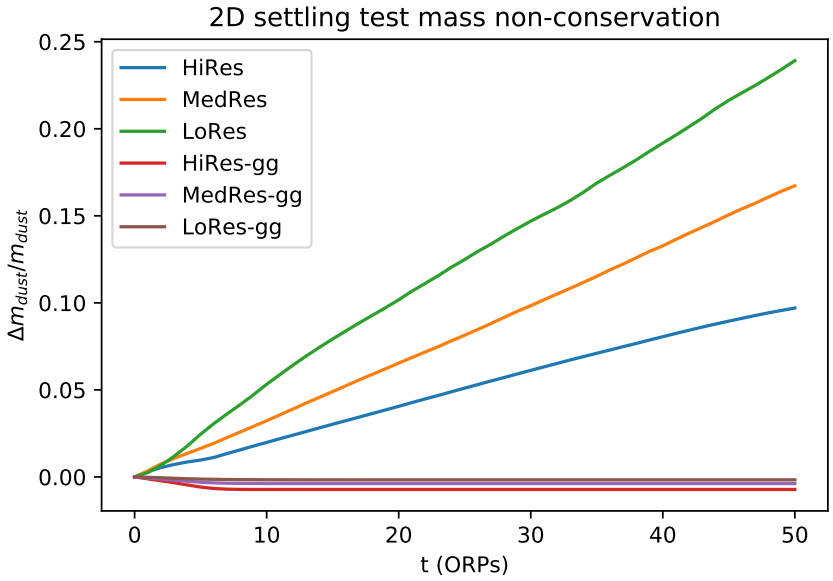



Figure 5.3: Fractional dust mass change for 2D dust settling tests at all 3 resolutions for the standard scheme and the upwinding scheme, marked by names ending in 'gg' (for grain-growth). The upwinding scheme displays much better mass conservation as it does not suffer as strongly from boundary error (see §4.5).

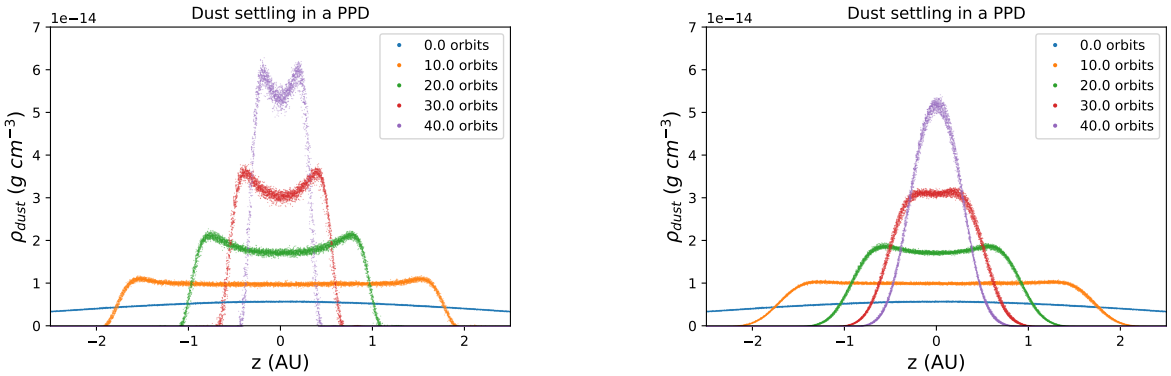
tling rate is still captured. Upwinding here has an effect similar to decreasing the resolution.

For reference, figure 5.5 shows a similar effect. The dust density after 50 ORPs is plotted for both methods, along with the semi-analytic solution (see §4.2.2) assuming no dust at $|z| > 3H$ for the ICs. The midplane smoothing length (a measure of spatial resolution at the midplane) is represented as bars in the upper left corner. The salient features of the semi-analytic solution are the double humped feature–a product of more rapid settling at larger $z$–and the overall width and height of the profile. The extremely sharp edges are due to neglecting turbulence or the dynamical effect of dust on the gas. The upwinding scheme underestimates but broadly captures the overall settling rate, but it fails to capture the outward-in nature of the settling. In this idealized 2D test, the upwinding affects the profile in a manner similar to decreasing the resolution.

### *5.5.2 Top-hat settling test*

The top-hat test is identical to the **?** settling test but with a double "top-hat" in the grain size. We set the grain size as a function of $z$: $s = \sqrt{3}$ mm for $0.75H \leq |z| \leq 2H$ and $s = 1/\sqrt{3}$ mm otherwise. This produces a double top-hat in the ICs of height $s_{max}/s_{min} = 3$ with a geometric mean size of 1 mm. Grain growth is turned off. The larger grains should settle more rapidly toward the midplane, producing an average size profile with peaks which move toward the midplane.

Figure 5.6 shows grain size $s$ profiles for the LoRes (fig. 5.6a) and the HiRes ((fig. 5.6b)) tests after 0, 10, 20, and 30 ORPs. The grain size is plotted vs $z$ in blue for the simulation and in green for the semi-analytic result. Note that the $z$ scale decreases in adjacent panels as the dust profile narrows and settles to the midplane. The semi-analytic solution is binned and averaged and overlaid as a dashed red line. The semi-analytic solution is calculated assuming the terminal velocity approximation and negligible dust mass such that the gas density profile (assumed to be gaussian) is static. This allows a vertical velocity profile $v_d(z)$ to be calculated (eq. 4.15). We then randomly seeded $4 \times 10^5$ points according to the gaussian gas density profile, each representing lagrangian dust particles. It is straightforward




(a) Reproduction of **?** 2D dust settling test using the one-fluid, terminal velocity dust scheme presented there with a fixed grain size. Dust upwinding is not used here.

(b) Dust settling test with dust upwinding (§5.3). Grain-growth was turned off by setting $\alpha_S = 0$. Dust-size advection is off by using uniform $s$.

Figure 5.4: Dust upwinding comparison in the HiRes 2D dust settling test presented in §4.2.2. The dust density is plotted for all 14208 SPH particles at 5 epochs. This is a 2D test in $x - z$ in a fixed external potential along $z$ which represents a patch of a PPD. The ICs are in equilibrium and are a nearly gaussian vertical density profile with a uniform dust fraction of $\epsilon \approx 0.01$ and uniform grain size $s = 1$mm. Dust upwinding was used in (b) and not (a). The diffusivity of the upwinding scheme is apparent in the smoothed profiles in panel (b). The double humped nature of the settling gets washed out. The upwinding results are similar to lower resolution for the non-upwinding scheme.

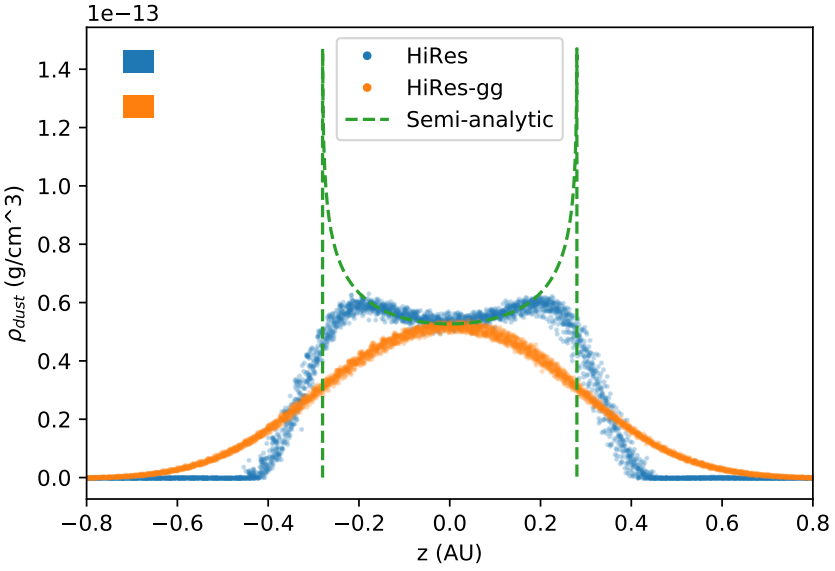


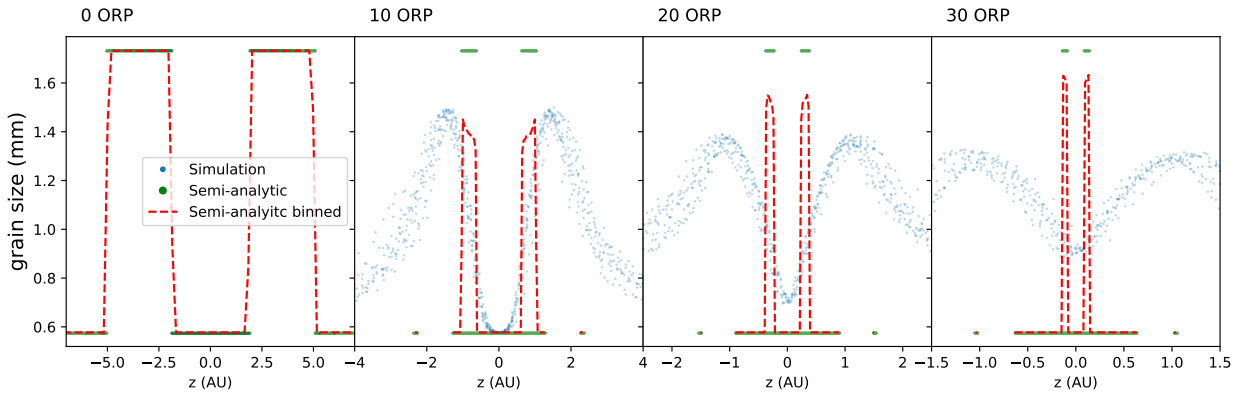

Figure 5.5: 2D HiRes dust settling test: upwinding comparison. The dust density is plotted vs $z$ after 50 ORPs for without upwinding (HiRes) and with upwinding (HiRes-gg). The semi-analytic solution assuming no dust beyond $3H$ is overlayed. The bars in the upper left indicate the midplane smoothing length. The upwinding scheme is significantly more diffusive and approximates the semi-analytic result worse.

to numerically integrate $v_d(z)$ to calculate the positions as a function of time.

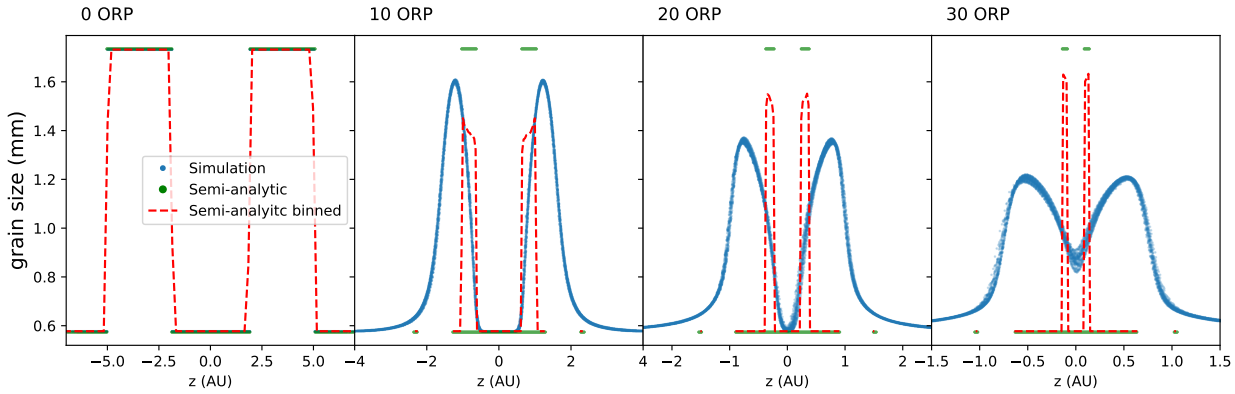
As our method is one-fluid, it only allows a single $s$ at a given location. We track the mean grain size. The simulation does reproduce two peaks in grain size that move toward the midplane. The binned and averaged semi-analytic result produces significantly sharper peaks, however if they were smoothed over several particles they would be similarly broad. Importantly, the settling occurs more slowly for the one-fluid approach, and the settling rate increases with resolution. However, the overall picture of a top-hat grain size distribution narrowing and settling to the midplane holds here, and the settling rate broadly agrees within a factor of order unity. Note also that the semi-analytic solution ignores any disk turbulence.

Figure 5.7 shows the dust density profiles for the same tests at the same times, with the semi-analytic density overlaid as a dashed red line. The semi-analytic density was estimated by histogramming the semi-analytically integrated lagrangian particles. Note the decreasing $z$ scale as time increases. As with the $s$ profiles, the semi-analytic result settles a bit more rapidly. The overall shape of 4 humps shows up in the simulation. The two, broad central peaks are from the rapidly settling larger grains. The two small peaks at the wings are not extremely evident for the simulation in this figure, but they are present at the correct locations. They are the peaks expected from particles rapidly settling at high altitudes, similar to the double humped profile in the 2D uniform $s$ settling test (fig. 5.4). The peaks are significantly more narrow in the semi-analytic solution, however this is not necessarily more realistic due to ignoring turbulence and the effect of dust on gas which will both tend to smooth out any such strong discontinuities in $\epsilon$.

As a general picture, these results indicate that our method can broadly capture dust settling and dust grains of variable sizes advected across many SPH particles. Upwinding essentially solves the problem of mass conservation in the idealized 2D settling case, although it introduces diffusion in $\epsilon$. The top-hat test represents a worst case scenario for advecting grain sizes in our method, as our method is poorly suited to handle steep gradients in the dust size since it assumes $s$ is locally uniform and continuous. We now proceed to examine results for a PPD test.




(a) LoRes top-hat test




(b) HiRes top-hat test

Figure 5.6: Dust size for the 2D top-hat settling test at several times and at two resolutions, LoRes (top) and HiRes (bottom). The ICs are identical to the **?** 2D settling test but with a double "top-hat" in the grain size, set such that $s = \sqrt{3}$ mm for $0.75H \leq |z| \leq 2H$ and $s = 1/\sqrt{3}$ mm otherwise. The semi-analytically integrated Lagrangian particles are displayed in green. The binned mean of $s$ for these are overlaid (dashed red line). The basic feature of two peaks settling radially inward is captured by the simulation, however the settling rate is decreased and the grain size is smoothed out, representing an average grain size.

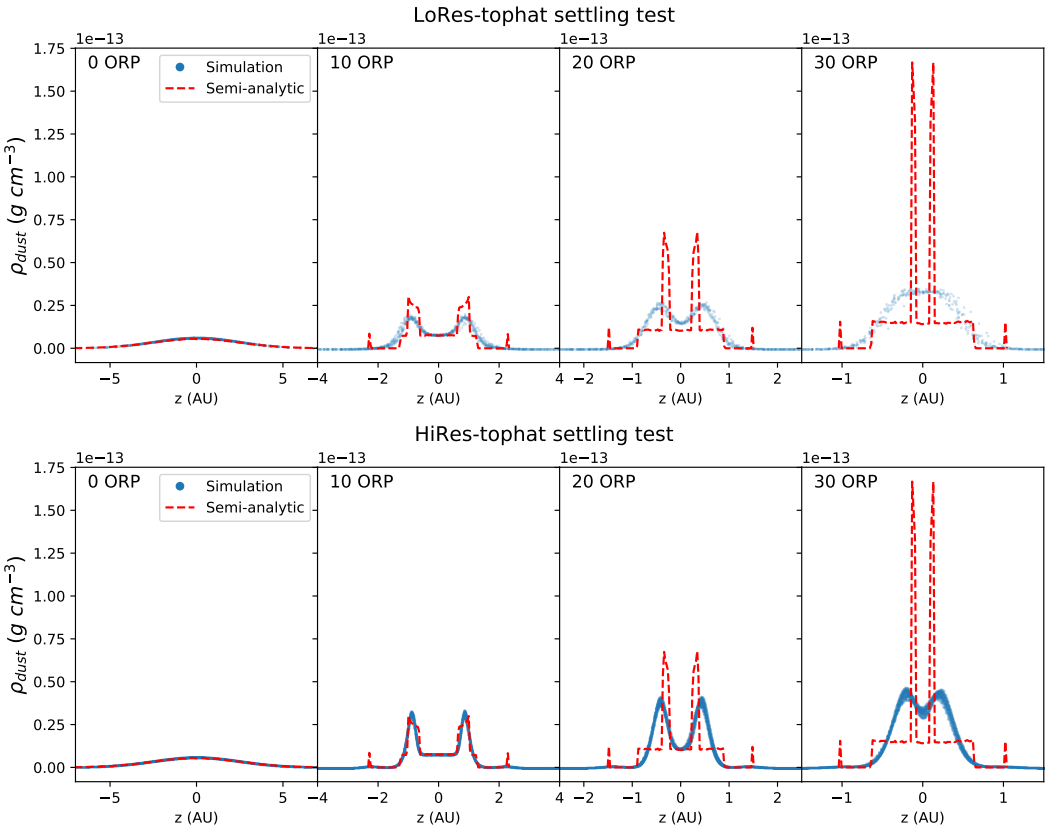



Figure 5.7: Dust density for the 2D top-hat settling test at several times and at two resolutions, LoRes (top) and HiRes (bottom). The ICs are identical to the **?** 2D settling test but with a double "top-hat" in the grain size, set such that $s = \sqrt{3}$ mm for $0.75H \leq |z| \leq 2H$ and $s = 1/\sqrt{3}$ mm otherwise (see fig. 5.6). The semi-analytic solution is overlaid assuming no dust beyond $3H$. The semi-analytic result displays much steeper peaks in $\rho_d$ which settle inward slightly more rapidly, although these may be unrealistically steep. All 4 peaks are present in the simulation and are captured better in the HiRes test.

### *5.6 PPD test*

Here we present the results of a PPD test using our one-fluid grain-growth and dust advection code. **?** tested two disk models to demonstrate their self induced dust-traps mechanism: a "flat" disk model with uniform $\Sigma$ and a "steep" disk model with $\Sigma \propto R^{-1}$. While the steep model is more realistic, we have chosen to test the flat disk since it demonstrates the dust-traps more strongly.

The uniform surface density is $\Sigma = 487.74$ kg/m$^2$. Interior and exterior cut-offs in $\Sigma$ are applied at $R = 4$ AU and $R_d = 120$ AU, respectively, as described in §3.2.1. The temperature follows a profile $T(R) = T_0(R/R_0)^{-1}$ with $T_0 = 623$ K and $R_0 = 1$ AU. A uniform dust fraction $\epsilon = 0.01$ and uniform grain size $s = 1$µm were used in the ICs. A grain density of $\rho_{\rm grain} = 1$ g/cm$^3$ was used, reasonable for ice or fluffy aggregate. For the grain growth model parameters we used the maximum fragmentation that Gonzalez considered, $\mathrm{V_{frag}} = 25$ m/s as a larger $\mathrm{V_{frag}}$ value promotes grain growth and formation of the self-induced dust traps.

The star is treated as a sink particle with radius $r_{sink} = 4$ AU. Gas particles which approach within $r_{sink}$ are accreted onto the star particle, conserving momentum and dust mass. For SPH parameters we used $N_{\rm smooth} = 128$, artificial viscosity parameters of $\alpha = 0.1$ and $\beta = 0.5$, and a Courant condition of $\eta_C = 0.3$. As with **?**, the gas gravitational softening length was chosen as $\epsilon_s = 0.5\,\langle h \rangle$ where $\langle h \rangle$ is the mean smoothing length, as discussed in §3.2.2.

**?** run their simulation without disk self gravity, i.e. the only gravity calculated is due to the central star. To easily achieve this with ChaNGa, which always calculates gravity, we simply scaled $\Sigma$ and $\rho_{\rm grain}$ by a factor of 1/50. This keeps the dust stopping times and dynamics unaffected and keeps the gas accelerations, from e.g. pressure gradients, constant. The gas+dust mass simply provides a negligible contribution to the gravitational forces.

Figure 5.8 shows the basic features of the disk model which may promote self-induced dust traps. Plotted are important dust grain sizes vs $R$ for the ICs. The initial dust size vs $R$ is plotted as the horizontal black line. The St = 1 line shows the vertically averaged $s$ which

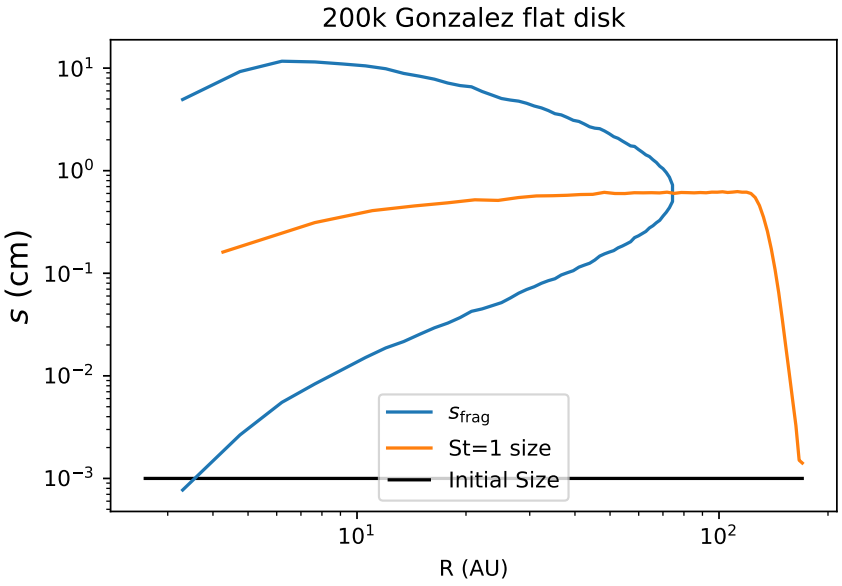



Figure 5.8: Important grain sizes for the Gonzalez flat disk ICs, vertically averaged. $s_{\text{frag}}$ (blue) is defined as the dust size(s) required for the transition between grain growth and grain destruction (eq. 5.1) $\text{V}_{\text{rel}} = \text{V}_{\text{frag}}$ (see eq. 5.2). The St = 1 size is the grain size that gives St = 1. The initial grain size is plotted in black. Grains with $s$ below the bottom $s_{\text{frag}}$ curve will grow. Grains with $s$ between the top and bottom $s_{\text{frag}}$ curves will decrease in size. All others will grow, including grains in regions where the $s_{\text{frag}}$ curve is not defined (e.g. at large $R$), which will grow indefinitely as $\text{V}_{\text{rel}} < \text{V}_{\text{frag}}$ for all grain sizes there.

would give St = 1. This grain size gives the maximum $\text{V}_{\text{rel}}$ for given disk conditions. In regions where this is greater than $\text{V}_{\text{frag}}$, this grain size provides the fastest grain destruction, while in regions where this is below $\text{V}_{\text{frag}}$ in generates the fastest grain growth, as per the model of eq. 5.1. In the uniform $\Sigma$ region of the disk, the St = 1 size gradually increase, dropping rapidly beyond $R_d = 120$ AU where the cutoff in $\Sigma$ is applied.

The blue $s_{\text{frag}}$ curve shows grain sizes at the transition between grain growth and destruction (vertically integrated). This transition is defined by $\text{V}_{\text{rel}} = \text{V}_{\text{frag}}$. For the $\text{V}_{\text{rel}}$ model

used here (eq. 5.2), solving for $s$ gives two roots to this equation:

$$s_{\rm frag}^{\pm} = s_{\rm drift}\left[\tilde{c_s}^2 - 1 \pm \tilde{c_s}\sqrt{\tilde{c_s}^2 - 2}\right] \tag{5.33}$$

with the definitions $\tilde{c_s}^2 \equiv \sqrt{2}{\rm Ro}\alpha_{\rm S}(c_s/{\rm V_{frag}})^2$ and $s_{\rm drift} \equiv \frac{\rho_g c_s}{\rho_{\rm grain}\Omega_K}$, which define the upper and lower portions of the blue curve in fig. 5.8. Grains with size $s < s_{\rm frag}^-$ will tend to grow until $s = s_{\rm frag}^-$, at which point growth stalls due to fragmentation. Larger grains, with $s > s_{\rm frag}^+$ are sufficiently decoupled from the gas that their relative velocities are not too large and will grow indefinitely. Grains with $s_{\rm frag}^- < s < s_{\rm frag}^+$ will shrink until $s = s_{\rm frag}^-$. At sufficiently large $R$, the temperature will decrease such that $\tilde{c_s}^2 < 2$ and there are no roots to eq. (5.33). In these regions, ${\rm V_{rel}} < {\rm V_{frag}}$ for all $s$ and the grains will tend to growth without destruction. For the disk presented here, this transition occurs around $R = 75$ AU. This model can therefore allow grains to grow at large $R$ where there is no $s_{\rm frag}$ boundary and migrate inward. Sufficiently large grains may migrate inward above the $s_{\rm frag}^+$ boundary and continue growing. It should be stressed that this is just an approximate, 1D analysis meant to illustrate the general behavior.

Figure 5.9 shows the grain size evolution. The grain size for all particles is plotted against $R$ at several epochs. The points are color-coded according the Stokes number. As with **?**, we get efficient grain growth at large $R$. Those grains migrate radially inward and the largest grains are able to form at small $R$. By 80 kyr, **?** form a separate population of large grains at large $R$ which migrate inward and get trapped at small $R$. Due to the averaging of our one-fluid $s$ advection method, we do not form a separate population of grains. Rather, our grain size shows a dependence on $R$ with little scatter. By the end of the simulation, at the inner disk our grains have grown to $s \gtrsim 10$ cm.

Although we do not form a separate population of large grains, we do form a sort of dust trap at the disk interior. Figure 5.10 shows the total surface density $\Sigma_{\rm tot} = \Sigma_d + \Sigma_g$ (top panel) and the dust surface density $\Sigma_d$ (bottom panel) at several times. Large spikes in the surface density due to dust pile-ups are evident by the end of the simulation. These regions of high dust density are formed as large grains migrate to the inner disk and clear gas away

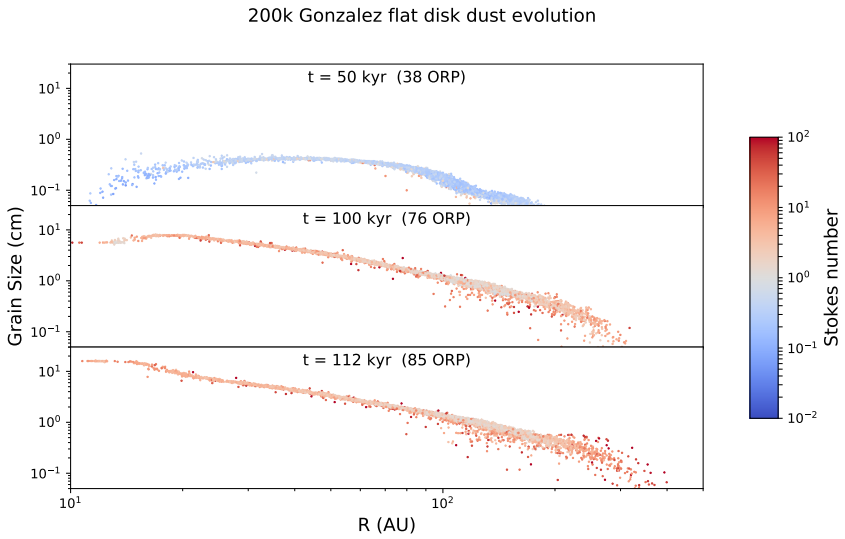



Figure 5.9: Grain-size evolution for the Gonzalez flat disk. The grain size is plotted as a function of $R$ and color coded according to the Stokes number. The grain size distribution is nearly a function of $R$, with little scatter, unlike with the two-fluid approach of Gonzalez which produces a separate population of large $s$, high St particles at inner regions of the disk.

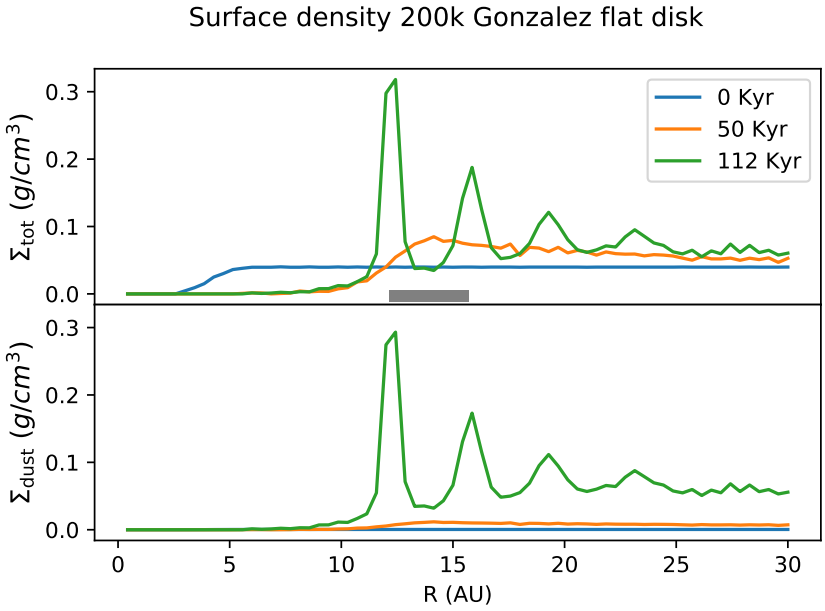



Figure 5.10: Total surface density (top panel) and dust surface density (bottom panel) vs $R$ at several times for the Gonzalez flat disk model. The gray bar represents twice the mean smoothing length $2\langle H \rangle$ at the location of the local trough in $\sigma$. This is the distance between a particle and its furthest interacting neighbor. This is approximately the distance between the peaks in $\Sigma$ at the end of the simulation.

from small annuli, creating pressure maxima which trap the dust. However, it should be noted that this is very likely a boundary effect. Due to the star sink particle, the very inner edge of the disk gets almost cleared of SPH particles out to $R \approx 10$ AU. The first spike in $\Sigma_d$ forms about one smoothing length exterior to this point, followed by several more spikes, each spaced by about twice the mean smoothing length, $2\langle H \rangle$ (pictured as a gray bar). This is the distance between a particle and its furthest SPH interacting neighbor. Given how unrealistic it is to have a sink particle vacuum boundary at small $R$ we cannot treat this effect as physical. It is very likely numerical.

It should be noted that the nature of our dust traps is visually different from those of **?**. Our traps are significantly more narrow, $\sim 5$ AU compared to $\sim 25$ AU. We appear to form

more traps, and at generally a smaller $R$, however at the end of their flat disk run (see their fig. 2) they appear to form a second dust trap beginning at the interior cut-off. Additionally, that run used a smaller $V_{frag} = 15$ m/sec, which may tend to form traps at larger $R$. Our traps may not be formed by exactly the same mechanism. It is uncertain from their plots whether boundary issues drive their dust traps as well, although their dust traps do form close to the inner boundary.

Unfortunately, even with upwinding, our approach still suffers from mass non-conservation, similar to our uniform grain-size simulations without upwinding. The basic mechanism is outlined (for the case without upwinding) in §4.5, but the basic effect is that vacuum boundaries can generate mass non-conservation as we must force $\epsilon \in [0, 1]$. This is exacerbated at large $R$ where $\Sigma_g$ is very low and settling times are very short and SPH particles without dust try to lose dust and therefore tend to force $\epsilon < 0$. Applying a floor to $\epsilon$ therefore tends to generate dust mass at large $R$ which then migrates rapidly inward. Figure 5.11 shows the dust mass per unit radius ($\Lambda$) at two times for the flat disk run with $10^6$ particles. Large amounts of dust mass are generated at $R > R_d$, migrate inward, and pile-up at the edge of the disk.

The inner disk is not as strongly affected, so it is unlikely that the dust traps which form are driven by this mass non-conservation. That being said, our simulation cannot be run for very long and all analysis should be restricted to the inner regions of the disk due to this numerical mass loading.

As found in §4.5.2, this error can be worsened with increasing resolution. Figure 5.12 shows the total dust mass normalized by the initial dust mass as a function of time for the flat disk at the 3 resolutions presented here. While the up-winding scheme is able limit mass non-conservation for a while, this gets broken at all resolutions at some point. As resolution increases, the mass non-conservation begins to occur earlier.

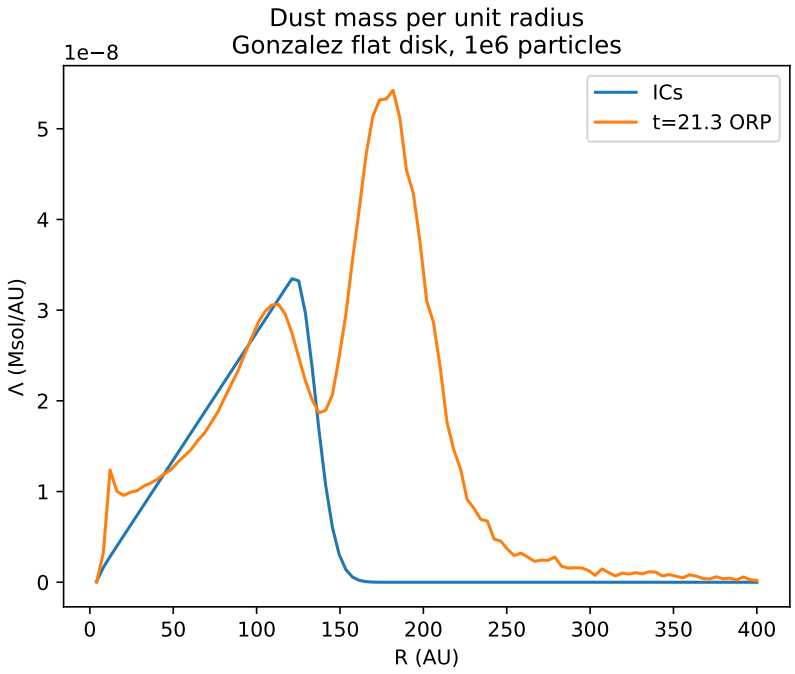



Figure 5.11: Dust mass per unit radius ($\Lambda$) for the Gonzalez flat disk with $10^6$ particles at two times. Dust mass is generated by the scheme at large $R$, well beyond the disk radius, and migrates radially inward, piling up as it hits the disk, as seen in §4.5.2 for uniform grain sizes. As seen by Gonzalez, the dust piles up at the inner disk. For us, this is likely an artificial boundary effect.

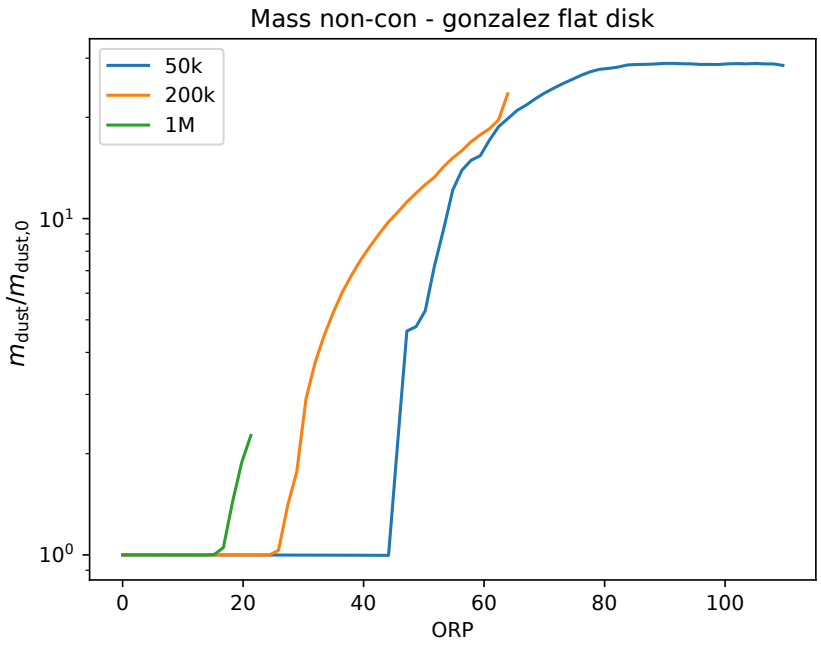



Figure 5.12: Dust-mass non-conservation for the Gonzalez flat disk, using the one-fluid, terminal velocity, upwinding dust scheme presented here. The total dust mass normalized by the initial dust mass is plotted as a function of time for 3 resolutions. The upwinding scheme is able to completely stall mass non-conservation for a while, but as with the non-upwinding scheme these PPDs undergo severe dust mass loading at large $R$ (see fig. 5.11), beyond the disk radius. The mass loading occurs earlier as resolution increases.

### *5.7 Conclusions*

In this section I presented a novel approach to modeling variable sized dust grains in gas with a one-fluid SPH approach. Dust grains are free to grow and fragment under collisions with the sub-grid turbulent model of **?**. Dust is allowed to be transported between SPH particles, as with $\epsilon$ in the one-fluid approach. Particles keep track of the average grain size $s$ which is free to be advected between particles, along with the dust. Our method will tend to smooth out grain size distributions due to this averaging, and the advection rate for $s$ will tend to be underestimated, but we broadly capture the expected behavior.

Altering the formulation of the dust fraction equation to use an upwinding scheme to integrate $\epsilon$ (eq. 5.13) and $u$ (eq. 5.17) essentially solves the mass non-conservation vacuum boundary problem in the 2D dust settling test, however in realistic disks it does not.

We use our implementation to test the self induced dust-traps model of **?**, and find that dust traps do indeed form, but they are likely caused by issues with the inner disk boundary. Since mass conservation is a significant problem in our simulations at large $R$ (beyond $R_d$), our analysis must be restricted to the inner disk. Because of this, our simulations are not entirely realistic and must be understood as a disk model which loads dust mass at large $R$. Over a short enough period (around the duration of these simulations) this does not have an enormous impact on the inner disk.

Since our one-fluid approach integrates the average $s$ in a region, we do not form the second population of large dust grains which putatively drive the dust-traps in simulations of **?**. However, the fact that we can generate dust traps numerically driven by boundary conditions problems may call into question their demonstration of dust-traps. While the self-induced dust traps model is very promising as a solution to the meter and fragmentation barriers, more work needs to be done to produce a fully realistic simulation of self-induced dust traps.

# Chapter 6

# CONCLUSIONS

## *6.1 Results*

The process of planet formation is long—beginning with a stellar nebula, through the stages of collapse, star and protoplanetary disk formation, solid growth, and eventual disk dissipation and evaporation leaving a planetary system—and it involves a broad swath of science, including chemistry, geology, fluid dynamics, magnetohydrodynamics, nuclear physics, and thermodynamics. Over the last several years, my work has focused on one aspect of that long story: the highly non-linear fluid and gravitational dynamics of young protoplanetary disks. I have investigated the role of global disk structure driven by gravitational instability, with an eye on explaining the formation of gas giants, on solid-gas interactions, and on illuminating the growth of solids from micron to meter sizes and beyond.

To model these very complex systems with dynamical, length, density, and force scales that span orders of magnitude, I have employed smoothed particle hydrodynamic methods in a highly parallel, supercomputing context. Certain features of SPH make it very well suited to studying PPDs, in particular: explicit simultaneous conservation of momentum, angular momentum, energy, and mass; natural handling of vacuum boundary conditions; and resolution naturally tracking the mass flow, allowing us to easily capture the 10 orders of magnitude or more of density variations which may arise in PPDs.

I have developed a fast, robust, and effective method to generate equilibrium initial conditions for PPDs. The importance of ICs in studying such non-linear, often unstable systems cannot be overemphasized. Perturbing ICs out of equilibrium can drive massive, disk wide instabilities and even cause artificially driven disk fragmentation. I have also studied the role of resolution and have demonstrated that decreasing resolution can drive

disks to fragment and may control the formation of spiral structure in gravitationally active disks. The effects of resolution are surprisingly poorly studied in the literature, and these effects are often non-trivial and unintuitive. For all my large simulations, I have performed resolution studies, and I urge the community to join me in doing so. Much published research does not share resolution tests or demonstrate convergence. Since the methods used to publish new science in computational astronomy are typically very new and poorly studied, publishing resolution tests should be standard practice. This can be done for almost no extra computational or human cost. Adding a few runs at say 0.1, 0.01, and 0.001 of the maximum resolution adds negligible overhead to the simulations, and since all analysis is done on computers with scripts, it adds almost no extra effort to the analysis chain.

Armed with a fast, simple method to generate ICs, and armed with our fast simulation code ChaNGa, I was able to produce the largest suite to-date of high resolution PPD simulations (64 simulations of $10^6$ particles) to study gravitational instability around M-Dwarfs. M-Dwarfs comprise around 70% of stars in our galaxy and may be the first stars around which we can find and characterize earth-like planets. We find that requirements for direct disk fragmentation via GI are more strict than previously thought, with Toomre $Q \lesssim 0.9$ being required for disk instability.

While a lower $Q$ may be required, it still appears that direct fragmentation is a plausible outcome of many disks and may explain gas giants at large radii. We also demonstrate that increased disk height tends to stabilize disks and introduce the effective Toomre $Q$ parameter $Q_{\rm eff} \equiv \beta Q (H/R)^{\alpha}$ which is a much better predictor of disk dynamics and fragmentation than $Q$. From our simulations, we derived parameters of $(\alpha, \beta) = (0.18, 2.1)$ that predict fragmentation for $Q_{\rm eff} < 1$. The cause of discrepancies between our results and previously published work is unknown because, in general, previous work is not reproducible.

While the gas component of a disk has been well studied with realistic, 3D full hydrodynamic and gravitational simulations, the solid components which eventually comprise terrestrial planets and gas giant cores, have received much less attention. This is primarily due to numerical and computational challenges. Individual dust grains are far too small

and numerous to model directly. In the ensemble, dust can behave poorly. It can have diffusion-like dynamics, meaning it will suffer from the harsh time-stepping constraint of the diffusion equation, requiring a timestep $\Delta t \propto (\Delta x)^2$. Additionally, dust can behave as having a negative diffusion coefficient, meaning it will tend to clump, which can drive large numerical instabilities and errors. Even in non-diffusive regimes, time-stepping constraints can remain harsh. Additionally, a distribution of dust sizes is expected at a given location, which may require integrating an entire population of dust grains, at enormous potential computational cost.

We have presented rigorous tests of an SPH method for simulating small dust grains in a dusty-gas mixture which is designed to be rapid by avoiding the diffusion time-stepping constraint and to suffer much less from the noise inherent to simulating dust. To test the method, we presented the highest resolution PPD simulation to-date ($10^8$ particles). We find that spiral arms drive dust enhancement rapidly, marked by rings of concentrated dust particularly at higher altitudes. Dust does not build up at the pressure maxima of spiral arms as previously speculated, and therefore spiral arms observed in the dust will be at least as well defined in the gas component. Coherent spiral arms do not limit settling, but chaotic spiral arms can.

The method of **?** used for simulating a dusty-gas mixture requires very high resolution to capture dust dynamics and therefore is not computationally fast. Dust settling is resolution dependent, increasing with increased resolution and converging at around $10^7$ or $10^8$ particles. Additionally, this method suffers from mass conservation and boundary problems that become worse with increasing resolution. This prevents the long-integration times required to achieve the high dust concentrations which would make dust dynamically important. We therefore find the method has limited use in the context of PPDs, although it may be well suited to more uniform environments without vacuum boundaries.

Inspired by the numerical dust mass-loading at large $R$ which is present in our simulations, we propose a mechanism for direct disk fragmentation at large $R$. Dust will tend to migrate radially inward in a PPD. At the outer regions of the disk, the migration velocity $v_r$ will

increase radially outward, causing convergent flow inwards which will cause dust to pile-up. If the disk can accrete sufficient mass from the surrounding nebula it may be driven Toomre unstable by the added dust mass and may fragment directly under GI. With so much dust mass present, direct collapse of clumps may be enhanced due to dust behaving like a pressure-less fluid and heat-sink that tends to migrate to pressure maxima. While plausible, this mechanism requires significantly more study to determine if it will operate in realistic disks.

Recently, **?** proposed the self induced dust traps model which may provide a natural means to overcome the meter and fragmentation barriers. Dust grains may grow efficiently in outer regions of the disk, migrate inward and pile-up, achieving sufficient densities to clear some gas from a region, creating a "trap" where dust will tend to collect. As gas is cleared away, the dust may grow rapidly and without migrating into the star. They provided a demonstration of this with a two-fluid dusty-gas code, but at low resolution using a technique which may suffer strongly from noise and which neglects disk self-gravity. We have developed an extension of the one-fluid approach to allow us to model variable grain size and to follow the local average of $s_{\rm grain}$, with the simple grain-growth and destruction model of **?**. We introduce an upwinding scheme to integrate the dust fraction equation which can mitigate dust mass non-conservation, although in a PPD simulation, the upwinded, one-fluid approach still suffers from mass non-conservation. While our simulations do not develop the separate population of grains which putatively drive the dust trap formation, we do form dust traps. These are due to boundary effects at the disk's artificial inner edge. This may explain the **?** results, although more work needs to be done to confirm or deny that.

### *6.2 Future work*

The underlying goal and over-arching trajectory of my research has been to work toward combining solid and gas dynamics to investigate the formation of gas giants via direct gravitational collapse and to explain terrestrial planet and gas giant core formation by studying dust migration and enhancement, and by understanding solid growth past the destruction

and meter barriers. There are several directions that this work should take, both as a direct continuation of my projects and as a task for the community at large.

Perhaps the biggest difficulty in SPH simulations of PPDs is to extend beyond an isothermal EOS. To date, no one has done so in a realistic manner. So-called "$\beta$ cooling" [**???**] simulations have been done, where disks are run with an adiabatic equation of state with external cooling, proportional to the temperature, with a time-scale proportional to the orbital period given by $t_{\rm cool} = \beta/\Omega$. However, the heating terms in these simulations may be primarily numerical, as evidenced by the reported non-convergence of the critical $\beta$ required to drive fragmentation [**??**].

The fundamental difficulty with a non-isothermal EOS in SPH simulations of PPDs is handling artificial viscosity. Artificial viscosity is required in SPH, particularly to handle the shocks which are naturally present in many astrophysical contexts. Even grid codes implicitly have some diffusive numerical terms. In the highly shearing environment of a PPD disk, AV is very active. With an isothermal simulation this is not a big problem since AV heating is suppressed, and it mainly leads to the disk behaving like a viscous disk with plausible levels of viscosity [**?**]. Once we relax the isothermal assumption, heating due to AV can dominate. For typical PPDs, our tests indicate that AV can triple the disk temperature within an ORP.

Modern SPH formulations all use the Balsara switch [**?**]. The switch is applied by multiplying artificial viscosity terms by the following factor, averaged over the interacting neighbor pair:

$$B \equiv \frac{|\nabla \cdot \mathbf{v}|}{|\nabla \times \mathbf{v}| + |\nabla \cdot \mathbf{v}|} \tag{6.1}$$

In the limit of infinite resolution in a Keplerian disk, $\nabla \cdot \mathbf{v} = 0$ and AV should be off, but due to discretization and SPH noise, estimates of $\nabla \cdot \mathbf{v}$ can be quite large and AV will be poorly damped. Some research has been done on alternative AV formulations designed specifically for accretion disks [**?**], but limited tests have been done on fully realistic, high resolution disks. A particular difficulty is designing an AV scheme which works well for keplerian flow and for flow in disks with spiral structure or which have undergone fragmentation.

If AV heating can be handled we will be able to model perhaps the most important dynamical effect that dust will have on disks: cooling. Dust is expected to strongly influence disk cooling, both as a heat sink and as an efficient black-body radiator. Dust will couple strongly to the disk opacity. As a first step, simple radiative cooling approximations might be used, appropriate for the somewhat optically thin regimes present in PPDs. One such approximation has been implemented in ChaNGa, similar to the work of **?** and **?** where the gradient of the radiative flux (which is proportional to the heating/cooling) is given locally by:

$$\nabla \cdot \mathbf{F} = -(36\pi)^{1/3} \frac{\sigma(T^4 - T_{irr}^4)}{r(\Delta\tau + 1/\Delta\tau)} \tag{6.2}$$

Where $T_{irr}$ is a background temperature profile, $r$ is the size of the SPH particle, and $\Delta\tau = r\kappa_d\rho$ is the optical depth of the particle with an opacity of $\kappa_d$.

Further work should also be done to explore and improve two-fluid dust methods. As we have demonstrated, particularly in the context of PPDs, one-fluid terminal velocity dust is of limited use in PPDs. One-fluid dust methods will always need to contend with the difficulties of conservation and of handling boundaries. Boundaries are non-trivial, even for grid codes where they are cleanly defined, and it is even more difficult for SPH [**?**]. Adding dust transport in a one-fluid method seems to unite all the difficulties unique to SPH with those unique to grid codes. While two-fluid dust may suffer from noise and clumping driven by SPH noise, there is some promising work currently underway. An improved, robust, two-fluid dust algorithm would allow us to probe important dust dynamics, including self-induced dust traps.

Investigating self induced dust traps in realistic simulations will also require further work on the disk inner-boundary: a source of much trouble for PPD simulations. The self induced dust traps model depends on both the inner and outer disk, and therefore we cannot restrict ourselves to simulating one region. Unfortunately, these regions have vastly different dynamical time-scales and simulating both is computationally very expensive. Cutting a hole at the inner disk is a simple way to handle this, but it has unknown consequences and may artificially drive dust trap formation. Vacuum boundary conditions are quite inappropriate

for the inner disk, and one might be able to help mitigate this by adding an external pressure term, centered on the star's location, to mimic gas orbiting close the the star.

We are unlikely to be able to solve the computational cost of the inner edge just by throwing computing power at it. Moving the inner edge in does not require adding many more particles to the simulation, but the time-steps rapidly decrease as $R \to 0$ and the gravitational and hydrodynamic forces diverge. Naive parallelization cannot help here: we simply have a handful of particles which need to be integrated over an enormous number of time-steps, while the rest of the simulation waits. For PPDs which have not fragmented, the simulation wall-time is dominated by the inner disk. Often, a handful of particles (fewer than 50 in a $10^6$ particle run) may be responsible for doubling or even tripling the wall-clock time.

Developing a load-balancing specific to the PPD geometry may help enormously. Handling accelerations due to the star's gravity using a method suitable for integrating Kepler's equations may relieve some of the time-stepping constraints. A method which I have played around with also shows some promise: increasing SPH particle mass within some radius in the ICs. This has the effect of decreasing the SPH resolution and increasing the time-steps in those regions, however it needs further testing, especially since variable mass SPH can raise difficulties.

In complement to the full, 3D simulation work, semi-analytic and simple 1D (along $R$) and 2D ($R + z$) hydrodynamical tests should be carried out to investigate both the self-induced dust traps and the outer-disk dust pile-ups. Both of these mechanisms may serve as a means to grow solids to the sizes required to form initial planetesimals, and both may serve to explain the formation of certain planets. Simplified tests and semi-analytic models allow us to explore a much larger range of parameter space, to test the plausibility of these models, and to select good parameters to be used for large-scale full simulations.

I hope to see these advances made in short order by the community. Our theoretical and observational understanding of protoplanetary disks and the processes of planet formation are advancing rapidly. This is the golden era for exoplanet science.

### *6.3 Code and data*

A major hindrance to progress in this branch of computational astronomy is a stunning lack of reproducibility. At a basic level, simulations and numerical methods are just math, and therefore results should be exactly reproducible. In my experience, this is almost never the case. Code and data are rarely published or made publicly available. Instead, cursory descriptions of methods and techniques are provided in papers. These descriptions are usually incomplete and very often contain inconsistencies, typos, or other such errors that make reproducing results impossible. This is not a case of a handful of offending papers, it is the norm. Not only is this a crisis for the believability of results, it introduces much re-inventing of the wheel. Since published techniques and methods—i.e. the code—cannot be directly used, they must often be re-developed from scratch with the literature as an incomplete guide.

In the digital era, your code and data are your math and your model. As such, they should be published. In that interest, I have made it a priority to make my code and data publicly available whenever possible. Unfortunately, journals typically have very low data caps for supplementary on-line material, so this requires finding alternative hosting resources.

Most of my tools for analyzing and handling PPDs are available online in my DISKPY package at `https://github.com/ibackus/diskpy`. IC generation is contained within the subpackage *diskpy.ICgen* and clump tracking is contained within the subpackage *diskpy.clumps*. TESTDUST, a package for generating and analyzing the dust tests of **?** can be found at `https://github.com/ibackus/testdust`. A simple package for generating SPH glasses (SPHGLASS) in periodic boxes is available at `https://github.com/ibackus/sphglass`. A public version of ChaNGa is available at `https://github.com/N-BodyShop/changa` and is required for those packages, including for IC generation. The group finding software SKID (required for clump finding) is available at `https://github.com/N-BodyShop/skid` and depends also on tipsy tools (`https://github.com/N-BodyShop/tipsy_tools`).

We have also made much of our data available online at the University of Washington's

ResearchWorks archive at `http://hdl.handle.net/1773/34933`. Initial conditions and final simulation snapshots are available for all the M-Dwarf simulations simulations presented in Chapter 3. Wengen test results (see appendix A) are also available there.

# Appendix A

# WENGEN TESTS

The Wengen tests are a series of code tests designed to compare different astrophysical hydrodynamic and gravity simulation codes and are available online at `http://www.astrosim.net/code/`. We ran Wengen test 4, an unstable isothermal PPD, with ChaNGa. We present the results of our simulation here. We find that our results are in good agreement with other simulation codes (previous test results can be found at `http://users.camk.edu.pl/gawrysz/test4/`). We have reproduced all the plots on the Wengen test website for the ChaNGa results, which are available at `http://hdl.handle.net/1773/34933`. Our ICs are the 200k-particle run. Here we present a few figures demonstrating the ChaNGa results.

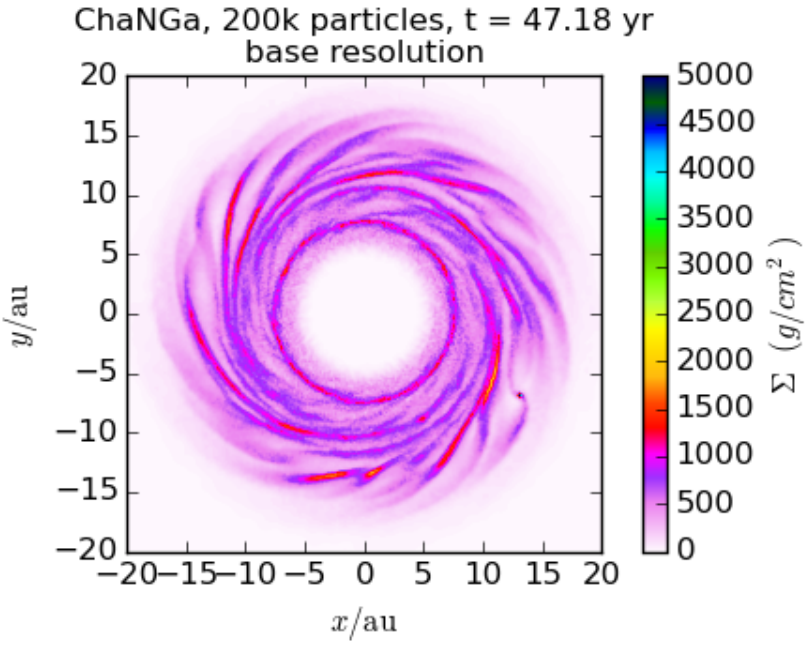



Figure A.1: Surface density map of the Wengen test simulation at $t = 47.18$ years (6 in code units). This reproduces the *surface density map* at $t = 6$ in the images table at http://users.camk.edu.pl/gawrysz/test4/#images. As with the other SPH codes and the higher resolution runs, a clump has formed.

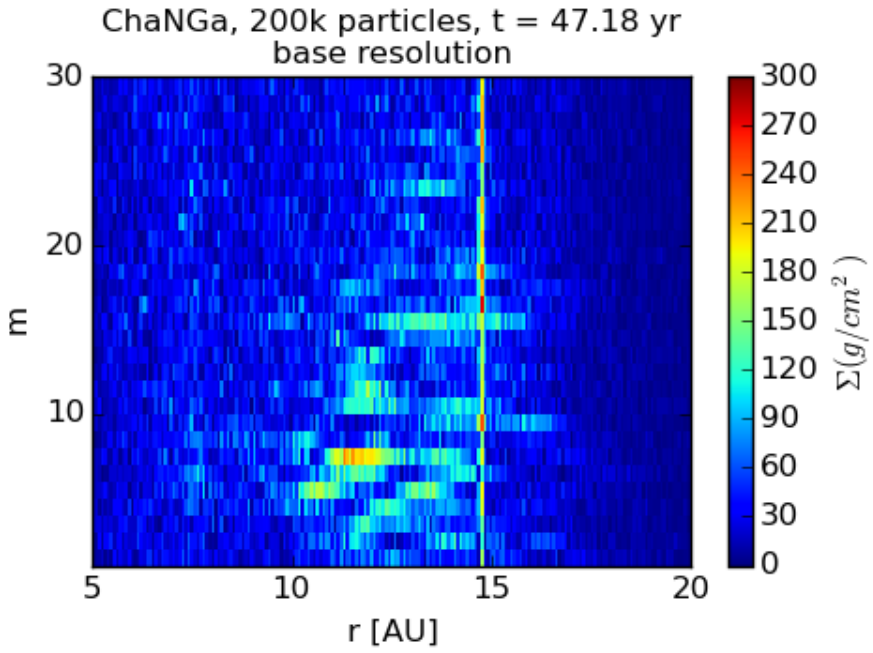



Figure A.2: Amplitude of the fourier transform of the surface density along the angular direction as a function of radius, for the same snapshot in figure A.1. This reproduces the *FFT of surface density* plots at $t = 6$ in the images table at http://users.camk.edu.pl/gawrysz/test4/#images. The clump which has formed shows up as a bright vertical stripe. 200 radial bins were used. The features in the plot agree well with those for other SPH codes and the high resolution codes.

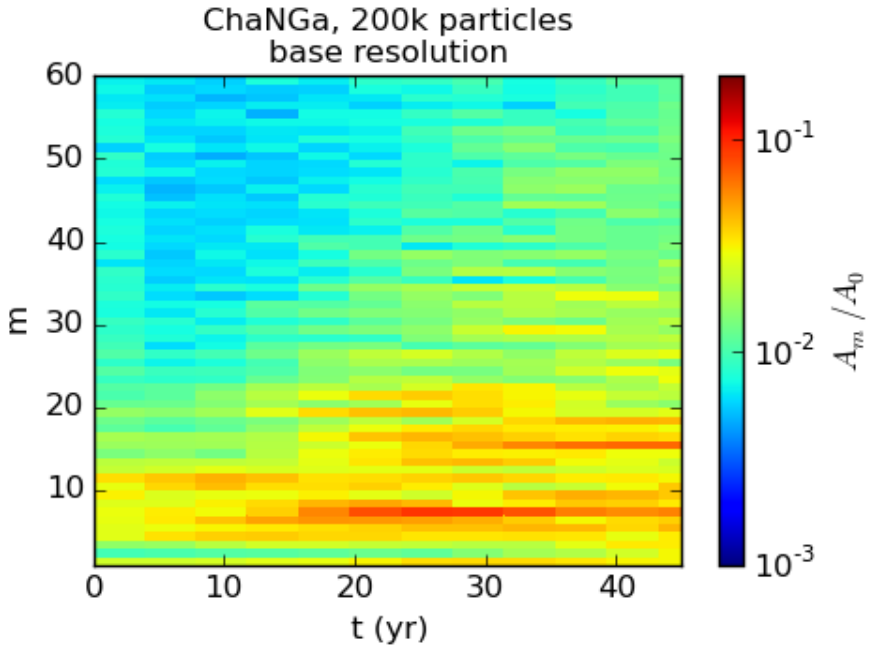



Figure A.3: Radially integrated fourier transform of surface density as a function of time. The amplitudes are normalized by the DC amplitude (not shown). This reproduces the *FFT integrated over r*, $A_m(t)$ plots in the images tables available at http://users.camk.edu.pl/gawrysz/test4/#images. The development of strong power around $m = 8$ and later around $m = 16$ agrees with other codes.

# Appendix B

# PL15 ONE-FLUID DUST TESTS

Here we present the results of running ChaNGa on the four dusty-gas tests of **?**: "dustydiffusion", 2D dust settling in a PPD, "dustywave", and "dustyshock". We refer the reader to **?** for a further description of the tests.

The "dustydiffusion" test is a dusty-gas only test in a periodic box with a uniform total density $\rho$ and with $\epsilon$ decreasing radially outward. An isothermal EOS is used. This means the gas density and the pressure increase radially outward. This generates a pressure gradient which drives dust diffusion radially outward. Particles are placed on a 3D $50 \times 58 \times 60$ and are fixed in place, which allows an analytic solution to be calculated. This tests the integration of the dust fraction equation only: the momentum and energy equations are not integrated. Figure B.1 plots $\epsilon$ vs spherical $r$ at several times for the dustydiffusion test with the analytic solution overlaid in red, similar to **?** figure 4.

Figure B.1 reproduces results for the 2D dust settling test, similar to **?** figure 8. This test is described in detail §4.2.2. This is an isothermal test with particles free to move (although their motion is generally small) and serves to test the momentum and dust fraction equations.

The dustywave test is a 1D SPH test of wave propagation in an adiabatic dusty-gas mixture. The particles are placed in a periodic box along $z$ with a small sinusoidal perturbation to the density, with corresponding velocity and pressure perturbations:

$$\rho(z) = \rho_0(1 + \delta\rho \sin kz) \tag{B.1}$$

$$v(z) = v_0 \sin kz \tag{B.2}$$

$$P(z) = P_0 + c_s v_0 \sin kz(1 + \frac{\delta\rho}{2} \sin kz) \tag{B.3}$$

where $k = 2\pi/L$, $\rho_0 = 1$, $\delta\rho = v_0 = 10^{-4}$ in code units. $P_0$ is the un-perturbed pressure, cal-

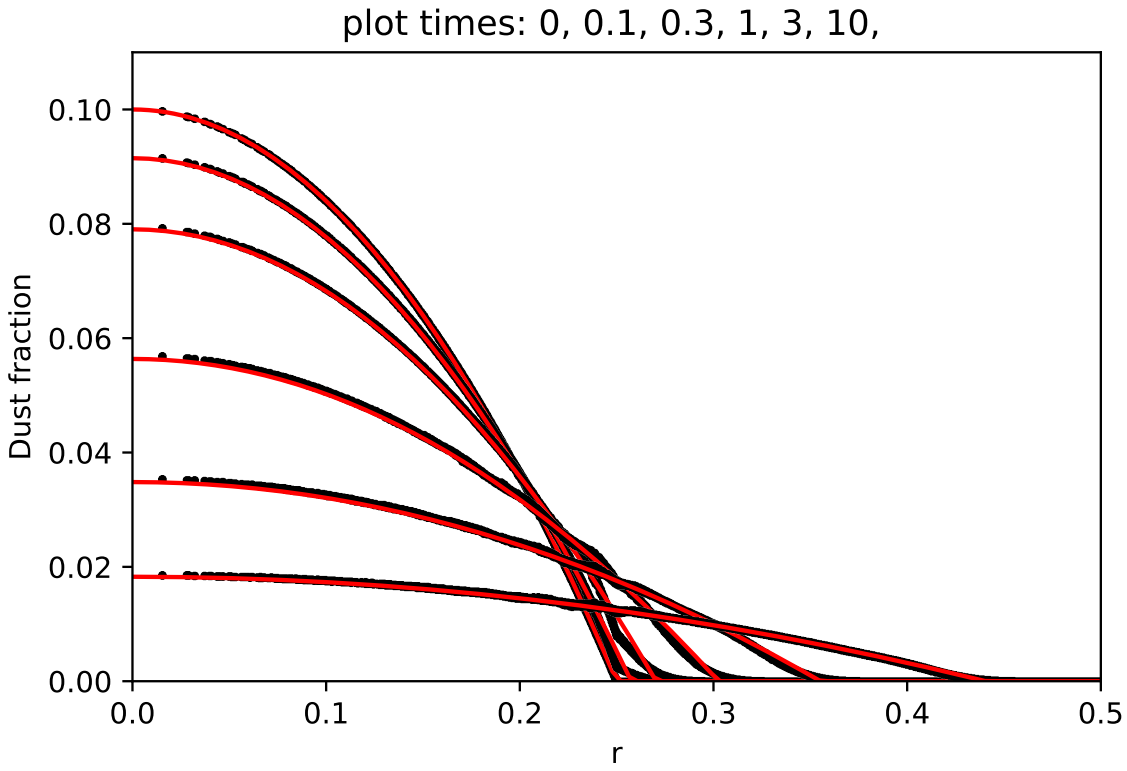



Figure B.1: Dust diffusion test of **?**, similar to their Fig. 4. The dust fraction for all particles at various times is plotted in black (dots) with the analytic solution overlaid in red. The peak dust fraction monotonically decreases with time and the profile broadens. The ICs is a periodic box with a uniform total density and $\epsilon$ decreasing radially. The test is run with an isothermal EOS and uniform $c_s$. Particles are placed on a 3D $50 \times 58 \times 60$ grid and fixed in place.

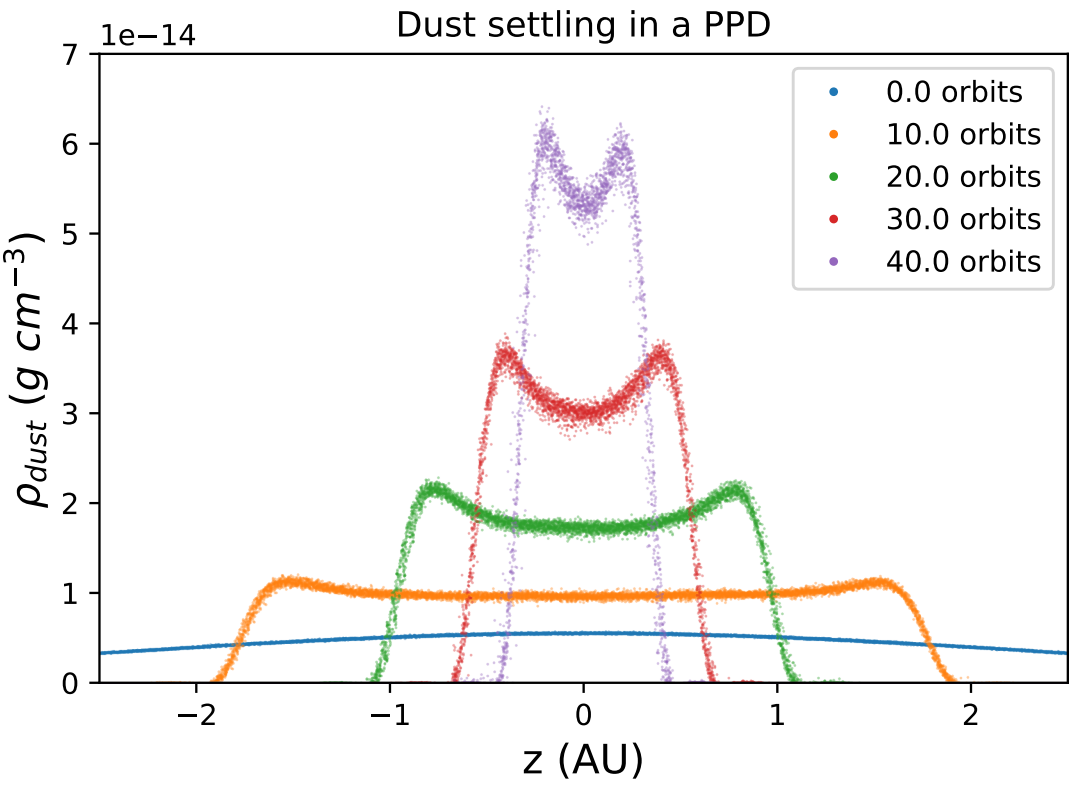



Figure B.2: Reproduction of Fig. 4.1 for completeness in this appendix. Results for the high resolution **?** dust settling test, similar to their figure 8.

culated from $c_s$. This tests the momentum, dust fraction, and energy equations. Figure B.3 shows the density after 4.5 periods with the analytic solution overlaid in red, similar to **?** figure 2. It should be noted that this is not a realistic test of wave propagation in SPH, since it is 1D and not a glass. It is also very sensitive to kernel and $N_{\rm smooth}$. Our tests show good behavior for the $M_4$, cubic spline kernel with $N_{\rm smooth}$ of 4 or 5 only, other neighbor counts ruin the wave propagation. The $M_6$ kernel performs poorly at all neighbor counts.

Figure B.4 shows the results for the dustyshock test, similar to **?** figure 3. The dustyshock test is an adiabatic shock tube test with $\epsilon = 0.5$ for small grains and is primarily a test of the momentum and energy equations. The analytic approximate solution is identical to an ideal gas with increased molecular weight and is overlaid in red.

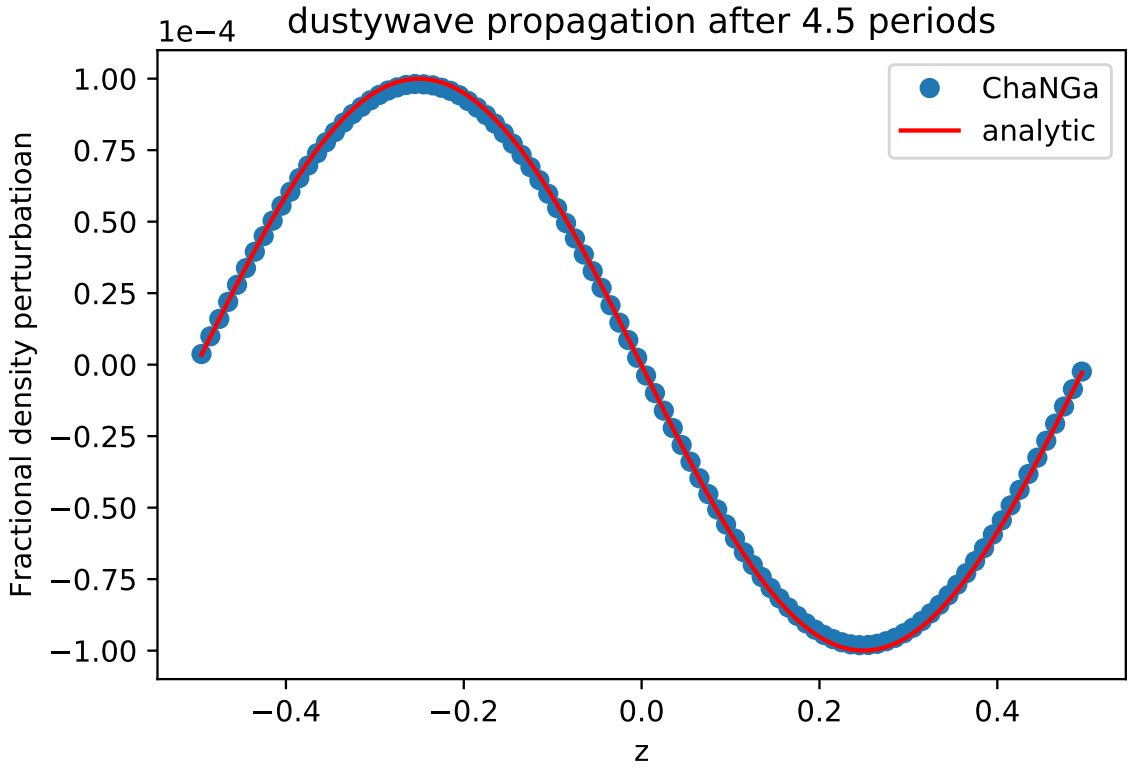



Figure B.3: Results for the **?** dustywave test. Plotted is the fractional density perturbation, defined as $(\rho - \langle\rho\rangle)/\langle\rho\rangle$. All 100 SPH particles are plotted. The dustywave test is a 1D dusty-gas mixture with a small (order $10^{-4}$) wave perturbation propagating along the +z direction. The grain size $s_{\text{grain}}$ is very small such that the dust is strongly coupled to the gas.

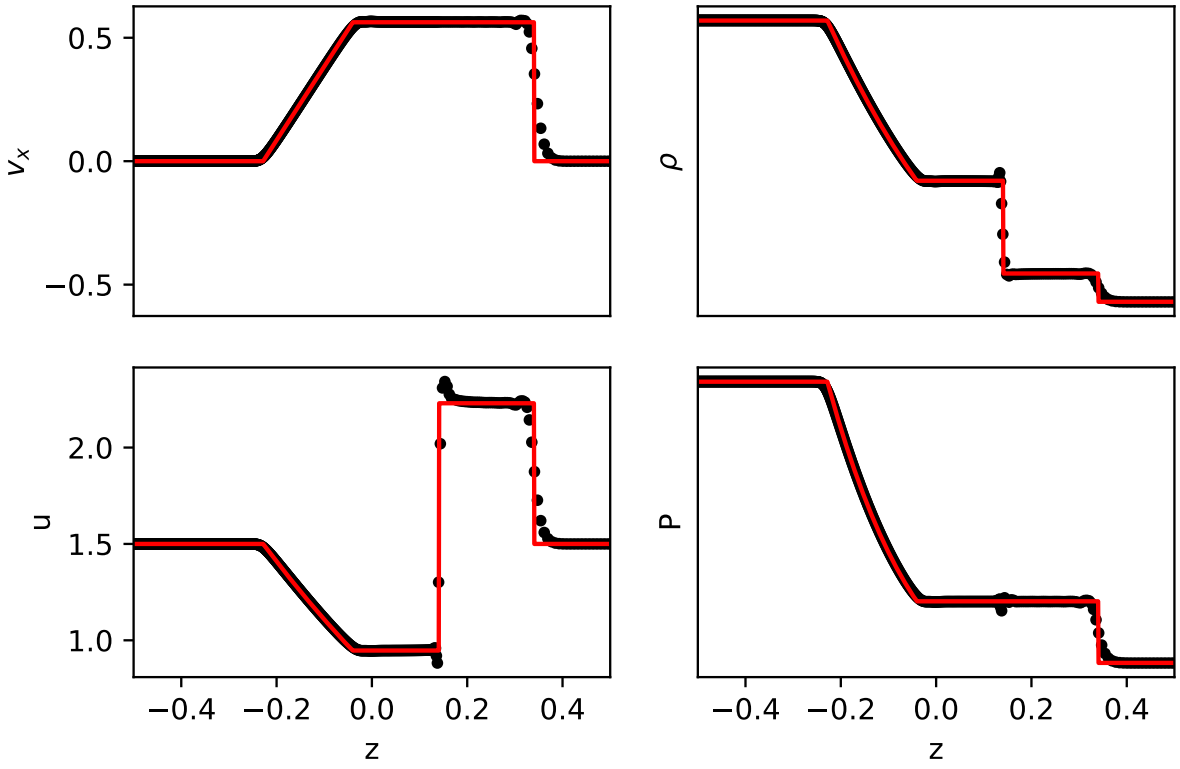



Figure B.4: Results for the dustyshock test of **?**, similar to their figure 2. All particles are plotted (black dots) and the analytic solution is overlaid (red line). The test is a 1D Sod shock tube with $s_{\text{grain}}$ very small such that the dust is strongly coupled to the gas.