Instabilities in supersonic cloud–cloud collisions

Andrew McLeod

A thesis submitted to Cardiff University
for the degree of Doctor of Philosophy

May 2012

Declaration

This work has not been submitted in substance for any other degree at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed ..... Date .....

This thesis is being submitted in fulfillment of the requirements for the degree of PhD.

Signed ..... Date .....

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ..... Date .....

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ..... Date .....

Acknowledgements

Thanks to my fantastic supervisor Ant Whitworth, without whom I would never have started this project. Thanks to Rose, without whom I would never have finished it.
Thanks to my examiners, Richard Nelson and Annabel Cartwright, for being thorough but fair and helping weed out a number of errors.
Thanks to Richard Wünsch, Thomas Bisbas and Jim Dale for being great people to work with, both in Cardiff and in Prague. Many thanks to Jan Palouš who arranged for me to come and work in Prague for eight months; it was one of the best times of my life.
Back in Cardiff, thanks to Steffi Walch and Dimitris Stamatellos for all their help.
Thanks to David Hubber and Chris Batty for helping provide the tools without which the project would not have been possible, and much more. Thanks to Oliver Lomax for making it fun, and Katy Holman for putting up with our endless distractions.
Thanks to Tom for his help immediately before submission and immediately after the viva, and many other times besides – it was an honour.
Thanks to all my other friends who have made my time at Cardiff University a pleasure.
Finally, thanks to my mother Jane, who has invested even more time and energy into me than I have into this.

Abstract

We study the effects of the supersonic collision of molecular clouds using smoothed particle hydrodynamics (SPH) simulations. We review the observational evidence for cloud–cloud collision and previous computational work. We describe the SPH method, the algorithms used in the SPH code SEREN (Hubber *et al.* 2011), and how we have extended the parallelization of SEREN. We review the non-linear thin shell instability (NTSI) and gravitational instability in a shock-compressed layer.

We present the results of two sets of SPH simulations. In the first set of simulations we collide supersonic flows of gas without self-gravity. We impose a range of velocity perturbations, including monochromatic perturbations, white noise perturbations and both subsonic and supersonic turbulence. The colliding flows create a dense shock-compressed layer which is unstable to the NTSI.

We examine the effect of the differing initial perturbations on the NTSI, and calculate rates of growth of both bending modes and breathing modes as a function of time and wavenumber. We compare our results to the time-independent result predicted by Vishniac (1994) for a one-dimensional monochromatic perturbation, and examine how this result can be extended to two-dimensional perturbations and non-monochromatic perturbations.

In our second set of simulations we model the head-on supersonic collision of two identical uniform-density spheres. We include self-gravity, allowing the dense layer to become gravitationally unstable and produce stars. We explore the effect of increasing collision velocity, and show that the NTSI is present only at higher collision velocities. At the highest collision velocities the NTSI severely disrupts the layer, and the collision does not produce stars. Although the global properties of the collision, such as the thickness of the layer, the size of the star-forming region and the time of first star formation, depend on the collision velocity, most individual properties of the stars do not.

Important conventions

We use the following conventions and definitions in this work.

- We use the floor and ceiling operators $\lfloor x \rfloor$ and $\lceil x \rceil$, which return the largest integer less than x and the smallest integer greater than x respectively.
- A tilde above a variable, such as $\tilde{F}$, indicates the result of a Fourier transform, which is a field of complex numbers.
- The wavenumber k is defined so that $k = \lambda^{-1}$ where λ is the wavelength. The magnitude of the wavevector $\boldsymbol{k}$ is similarly defined.
- The *amplitude spectrum* or *Fourier spectrum* of a Fourier transform is the quantity

$$A(k) = \sqrt{\tilde{F}(k)\ \tilde{F}^*(k)}\,, \tag{1}$$

  for a one-dimensional Fourier transform, where $\tilde{F}$ is the complex Fourier field and $\tilde{F}^*$ is its complex conjugate. The *power spectrum* is defined as the amplitude spectrum squared, but we avoid its use in this work.

- A *growth rate* is defined as the inverse of a *growth time*, such that

$$\omega = \frac{1}{\tau_{\text{growth}}}\,, \tag{2}$$

  where τ is the growth time. The growth rate has units of inverse time, such as Myr^{-1}. The growth timescale is the time it takes for the physical system in question to grow. This is usually the time a perturbation has taken to grow or will take to grow to some value.

- A *rate of growth* is the actual rate at which some quantity associated with a perturbation grows. Its units depend on the quantity being measured for that perturbation. For a wave-like perturbation whose amplitude is a length, the rate of growth will have units of length per time, which is simply a velocity such as $\text{pc}\,\text{Myr}^{-1}$. For surface density perturbations, the rate of growth will have units of surface density per time, such as $\text{g}\,\text{cm}^{-2}\,\text{Myr}^{-1}$.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Stars form by the gravitational collapse of dense molecular gas. On large scales gas is not gravitationally unstable, and cannot collapse to form stars without external influence. The rate of star formation therefore depends on the rate of external forcing that raises the density of gas.

A number of processes can increase gas density. Spiral density waves can compress gas, and star formation is preferentially sited along spiral arms. Turbulence in the interstellar medium can create density peaks; however the source of the turbulence must then be explained. Proposed sources include supernova explosions and stellar winds from massive stars; this can be considered as a collision between hot and cold gas. Collisions between gas clouds may also be an important source of gas compression.

We use numerical simulations to explore the effect of cloud–cloud collisions. We examine the major instabilities that can affect the dense shock-compressed layer formed in a cloud–cloud collision. We simulate cloud–cloud collisions at a range of collision velocities to examine the effect on the instabilities produced, and the effect on the properties of the stars that are formed.

In this chapter, we take an overview of star formation, examine the importance of collisions to star formation rates in the Galaxy, and review the observational evidence for cloud–cloud collisions. In Chapter 2 we describe the numerical method we use to perform our simulations. In Chapter 3 we describe the algorithms we use to perform our simulations. In Chapter 4 we describe the parallelization techniques we use to accelerate our simulations. In Chapter 5 we describe the techniques we use to analyse the output of our simulations. In Chapter 6 we examine the gravitational instability in three-dimensions and in a two-dimensional layer, and review previous numerical work studying the gravitational instability. In Chapter 7 we examine the theory of the non-linear thin shell instability, and review previous numerical work studying the instability. In Chapter 8 we describe how we create initial conditions for our simulations, and examine some resolution requirements for our simulations. In Chapter 9 we present the results of simulations of the non-linear thin shell instability, and compare our results to the

theoretical predictions. In Chapter 14 we present the results of simulations of cloud–cloud collisions, examining the effect of collision velocity on the instabilities produced and the properties of any stars formed. Finally in Chapter 15 we present our conclusions.

1.1 Importance of cloud–cloud collisions to star formation

Elmegreen (1998) notes that the average collision timescale for clouds is

$$\tau \sim \frac{1}{\pi R^2 n v}\,, \tag{1.1}$$

where R is the radius of a cloud, n is the number density of clouds and v is the cloud–cloud rms velocity dispersion.

A large fraction of the molecular gas in the Galaxy is confined to a molecular ring between Galactocentric radii of approximately 3.5 kpc and 7.5 kpc (Burton *et al.* 1975; Clemens, Sanders & Scoville 1988; Scoville & Solomon 1975). The full width at half maximum thickness of the disc is approximately 100 pc (Bronfman *et al.* 2000, 1988; Clemens, Sanders & Scoville 1988; Stark & Lee 2005). This gives a total volume of approximately 13.8 kpc^3.

We can estimate the number of clouds in the galaxy from the total molecular mass of the disc and the median mass of a cloud. Estimates of the total mass, adjusted for a solar Galactocentric radius of 8.5 kpc, are of at least $10^9\,\mathrm{M}_\odot$ (Bronfman *et al.* 1988; Clemens, Sanders & Scoville 1988; Solomon *et al.* 1987). Solomon *et al.* (1987) found a median cloud mass of $5 \times 10^5\,\mathrm{M}_\odot$, giving 2000 clouds. Williams & McKee (1997) suggest a model with 6400 clouds above $10^4\,\mathrm{M}_\odot$. We use an intermediate value of 4000 molecular clouds. We therefore obtain a number density of molecular clouds in the molecular ring of approximately 290 kpc^{-3}.

The sizes of molecular clouds vary, but are usually taken to be of the order of 30 pc in radius (Combes 1991; Dame, Hartmann & Thaddeus 2001; Solomon *et al.* 1987). The clouds have an cloud–cloud rms velocity dispersion of approximately 5 $\mathrm{km\,s}^{-1}$ (Liszt, Burton & Xiang 1984; Stark & Brand 1989). The average timescale between collisions is then approximately 240 Myr. Each collision lasts a time

$$\tau_\mathrm{c} \sim \frac{R}{v}\,, \tag{1.2}$$

which for our assumptions is approximately 6 Myr. Each cloud spends only 2% of its time engaged in collisions. Cloud–cloud collisions are therefore unlikely to be the major source of star formation in the Galaxy.

There are a number of factors that may increase the number of cloud–cloud collisions in the Galaxy. At smaller Galactic radii, the number density of clouds may be greater. A possible example of a cloud collision at very small Galactocentric radius is given in

Section 1.2.3. Self-gravity may cause more clouds to collide, and gravitational focussing will tend to increase the significance of each collision by making each collision faster and more head-on. Finally, the density of clouds is likely to be much higher in spiral shocks, leading to increased collision rates (Kenney & Lord 1991; Kwan & Valdes 1983, 1987; Roberts & Stewart 1987).

Our simple estimate of collision times may also be inaccurate; Tasker & Tan (2009) conducted simulations which suggest that collisions may be much more frequent. They suggest clouds are involved in a merger every fifth of an orbit, or approximately 25 Myr for a cloud at a Galactocentric radius of 4 kpc.

However, even ignoring these additional factors then at any given time approximately 80 of our 4000 clouds will be undergoing collisions. In Section 1.2 we show some observational evidence of such collisions.

1.2 Observational evidence

In this section we review some of the observational evidence for star formation triggered by cloud–cloud collisions in the interstellar medium. It would be fortunate, but improbable, to observe a cloud–cloud collision in the process of forming stars. Instead we are more likely to observe the resulting young star cluster embedded in their parent molecular clouds. While young star clusters are often seen in association with the molecular gas from which they formed, it is difficult to determine whether the stars formed from spontaneous gravitational collapse, or whether star formation was externally triggered.

Molecular clouds are chaotic and turbulent, making it difficult to determine even their recent history from observations. Distances to molecular clouds are often uncertain and usually only the radial component of cloud velocities can be measured. Star clusters formed following a collision can cause the clouds to move apart from each other by the rocket effect (Kahn 1954; Oort & Spitzer 1955). This can disguise the previous velocities of the clouds.

Nonetheless there are a few examples, on a range of scales, that are consistent with collision-induced star formation. This includes evidence of collisions between giant molecular clouds of $\sim 10^5\,\mathrm{M}_\odot$ to much smaller collisions between cores of just a few $\mathrm{M}_\odot$.

Loren (1976) proposed a cloud collision in NGC 1333 to explain CO observations; he then proposed a collision in LkHα for the same reason (Loren 1977). Since then, cloud collisions have been proposed to explain observations in the WR75–DR 21 complex (Dickel, Dickel & Wilson 1978), the high latitude cloud MBM 55 (Vallee & Avery 1990), the stellar clusters DR 15 and DR 20 (Odenwald *et al.* 1990), the Orion molecular ridge OMC-1 (Greaves & White 1991; Womack, Ziurys & Sage 1993), the low-mass cloud G100–13 (Odenwald *et al.* 1992), the molecular cloud core W49N (Serabyn, Güsten &

Source	Distance (kpc)	Object	Observations
Westerlund 1961	≥ 1.4	West. 2[1]	Optical[2]
Wilson *et al.* 1970	6.0 ± 1.2	RCW49	Radio recombination lines[3]
Caswell & Haynes 1987	4.9	RCW49	Radio recombination lines[4]
Moffat, Shara & Potter 1991	$7.2^{+1.2}_{-1.0}$	82 stars	UBV photometry
Shara *et al.* 1991	6.3/4.6	WR20a	Optical spectra[5]
Shara *et al.* 1991	6.3/4.5	WR20b	Optical spectra[5]
Brand & Blitz 1993	2.31	West. 2[1]	Spectro-photometric
Piatti, Bica & Clariá 1998	5.7 ± 0.3	4 O7V stars	Photometry[6]
Van der Hucht 2001	5.75	WR20a	Catalogue data
Van der Hucht 2001	2.27	WR20b	Catalogue data
Carraro & Munari 2004	6.4 ± 0.4	∼300 stars	UBVRI photometry
Rauw *et al.* 2007	8	WR20a	Photometry & spectra
Rauw *et al.* 2007	8.3 ± 1.6	12 O stars	Photometry & spectra
Ascenso *et al.* 2007	2.8	West. 2[1]	JHK photometry
Dame 2007	6.0 ± 1.0	Mol. cloud	^{12}CO (J=1→0) velocities[7]
Furukawa *et al.* 2009	$5.4^{+1.1}_{-1.4}$	Mol. cloud	^{12}CO (J=2→1) velocities[7]

[1] Westerlund 2
[2] Cluster assumed to be in Sagittarius arm
[3] Schmidt model galaxy
[4] Flat rotation curve at 250 km s^{-1} outside solar Galactocentric radius of 10 kpc
[5] Different calibrations were used to obtain the two different values
[6] Recalculation using the photometry data of Moffat, Shara & Potter (1991)
[7] Distance calculated using Galactic model of Brand & Blitz (1993)

Table 1.1 – List of distance estimates to RCW49/Westerlund 2. Where not otherwise specified, 'stars' are members of Westerlund 2, and 'photometry' and 'spectroscopy' refer to optical observations.

Schulz 1993), the IRAS sources IRAS 19550+3248 (Koo *et al.* 1994), IRAS 2306+1451 (Vallee 1995) and IRAS 04000+5052 (Wang *et al.* 2004), and the young cluster associated with the star BD+ 40 4124 (Looney *et al.* 2006).

Lepine & Duvert (1994) propose a method of star formation triggered by the infall of high-velocity clouds onto the Galactic disc, and provide a number of possible examples. Higuchi *et al.* (2010) present the results of a mapping study, and suggest that several of the objects they detect are the results of cloud–collisions. We examine a small number of cloud collision candidates in more detail.

1.2.1 RCW49 and Westerlund 2

The HII region RCW49 (also known as Gum 29 and NGC 3247) contains the super star cluster Westerlund 2 (Westerlund 1961). Westerlund 2 contains at least 12 O stars, the Wolf–Rayet star WR20b and the eclipsing binary of Wolf–Rayet stars WR20a (Rauw *et*

al. 2007; Shara *et al.* 1991; van der Hucht 2001). Molecular clouds associated with the cluster are spread over a relatively wide range of velocities (Dame 2007; Furukawa *et al.* 2009; Ohama *et al.* 2010; Wilson *et al.* 1970), and the region is also associated with a pair of gamma-ray sources (Abramowski *et al.* 2011).

Westerlund 2 is an extremely young cluster. Ascenso *et al.* (2007) observe several thousand stars in the near-infrared and estimate a total stellar mass of 7000 $M_\odot$, while Whitney *et al.* (2004) estimate a total stellar mass of $\sim 3 \times 10^4\,M_\odot$ by considering the ionizing flux from radio observations. A wide range of distances from 2 to 8 kpc have been used in the recent literature. Some of these values are shown in Table 1.1.

The cluster is a site of ongoing massive star formation. By observing stars with an infrared excess, Whitney *et al.* (2004) find approximately 300 YSOs of at least 2.5 $M_\odot$, and estimate a total mass of 4500 $M_\odot$ in 7000 YSOs (Whitney *et al.* 2004).

Rauw *et al.* (2007) estimate that there must be $\sim 4500\,M_\odot$ of stars above $1\,M_\odot$ to produce the observed HII region, although Ascenso *et al.* (2007) expects only $\sim 700\,M_\odot$. Piatti, Bica & Clariá (1998) assume an age of 2 Myr to 3 Myr by comparison with the similar cluster NGC 3603, while Ascenso *et al.* (2007) find a cluster age of ≤ 2 Myr. The age of the binary system WR20a has been estimated as 1 Myr to 2 Myr by Rauw *et al.* (2005).

Fukui *et al.* (2009) present ^{12}CO (J=1→0) observations of the region surrounding the cluster, showing an extended jet and arc feature. It is likely that the observations are not seeing the jet itself, but instead molecular gas formed in shocks created by the passage of the much-faster jet. The jet is approximately 100 pc long over a velocity range of 19 km s^{-1} to 30 km s^{-1}, and may be the result of either an anisotropic supernova explosion or an accretion shock from a supernova explosion in a binary system. The arc is centred at a velocity of 26 km s^{-1}, with a relatively small velocity width of 4 km s^{-1}, and appears to be the limb-brightened edge of an expanding shell. With a typical jet velocity of 1000 km s^{-1}, the supernova would have exploded approximately 10^4 yr ago.

If the jet and arc are the result of a supernova in the cluster, this would suggest a minimum cluster age of approximately 3 Myr. This would match the observations of Ascenso *et al.* (2007) who suggested a possible earlier wave of star formation at 3.0 Myr to 4.8 Myr.

Cloud 8 of Grabelsky *et al.* (1988) is a giant molecular cloud (GMC) in the region of RCW49 and Westerlund 2. Recently a range of higher-resolution observations of this cloud and the regions surrounding Westerlund 2 have been performed using the molecular gas tracers ^{12}CO (J=1→0), ^{12}CO (J=2→1), and ^{13}CO (J=2→1).

Dame (2007) referred to this cloud as MC8, and identified clumps, each of a few $10^4\,M_\odot$, in observations using ^{12}CO (J=1→0). The total mass of H_2 was estimated as $7.5 \times 10^5\,M_\odot$. A 14 km s^{-1} clump was described as associated with Westerlund 2, together

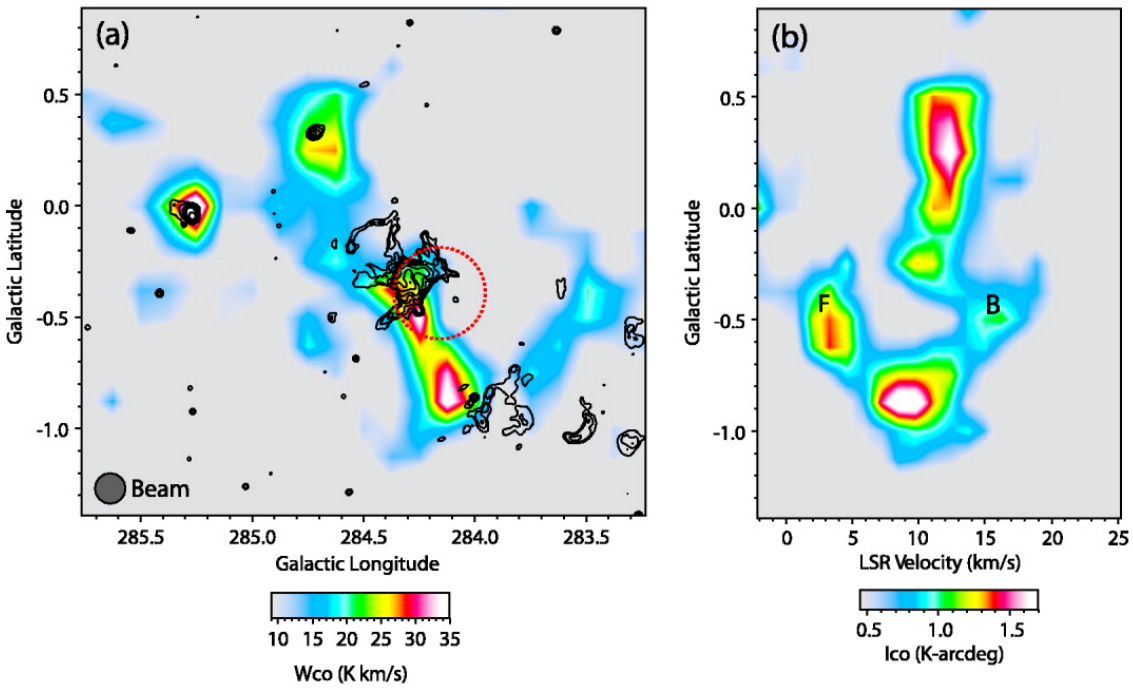


Figure 1.1 – ^{12}CO (J=1→0) observations of the region around RCW49 (Dame 2007). The colour scale indicates CO intensity. In the left figure this is integrated between 0 and 20 km s^{-1}; in the right figure this is integrated over the longitude range 284° to 285°. The contours are 843 MHz continuum emission (Green *et al.* 1999). As described in Dame (2007), the shell-like structure at G 283.9,−0.9 is RCW48 and may be unrelated.

with a 4 km s^{-1} clump whose association was less certain. They conclude that the 4 km s^{-1} cloud, and all gas with a lower velocity, probably lies in front of Westerlund 2, while the 14 km s^{-1} cloud, and all gas with a higher velocity, probably lies behind Westerlund 2.

They found a systematic velocity for the GMC of 11 km s^{-1}, but noted there is relatively little molecular gas at that velocity in the region of RCW49. Figure 1.1 suggests this is because gas is moving away from RCW49, which they suggested is the result of feedback from the cluster.

Furukawa *et al.* (2009) presented observations of ^{12}CO (J=2→1). They identified emission at $\sim$ 4 km s^{-1} and $\sim$ 16 km s^{-1}, corresponding to the 4 km s^{-1} and 14 km s^{-1} clumps of Dame (2007), but also identified a clump with a velocity of ~ -4 km s^{-1}. Figure 1.2 shows the distribution of molecular gas at different velocity ranges. The *Spitzer* image shows both the Westerlund 2 cluster, and the nebula of dust and gas surrounding RCW49, including a bright ridge just to the south-east of the cluster. Figure 1.3 shows CO intensity as a function of latitude, integrated over the longitude range 284.2° to 284.4°.

The 10.8 km s^{-1} to 20.9 km s^{-1} range of Figure 1.2 shows the extent of the 16 km s^{-1} cloud. This cloud extends well beyond the cluster, yet has a relatively uniform velocity,

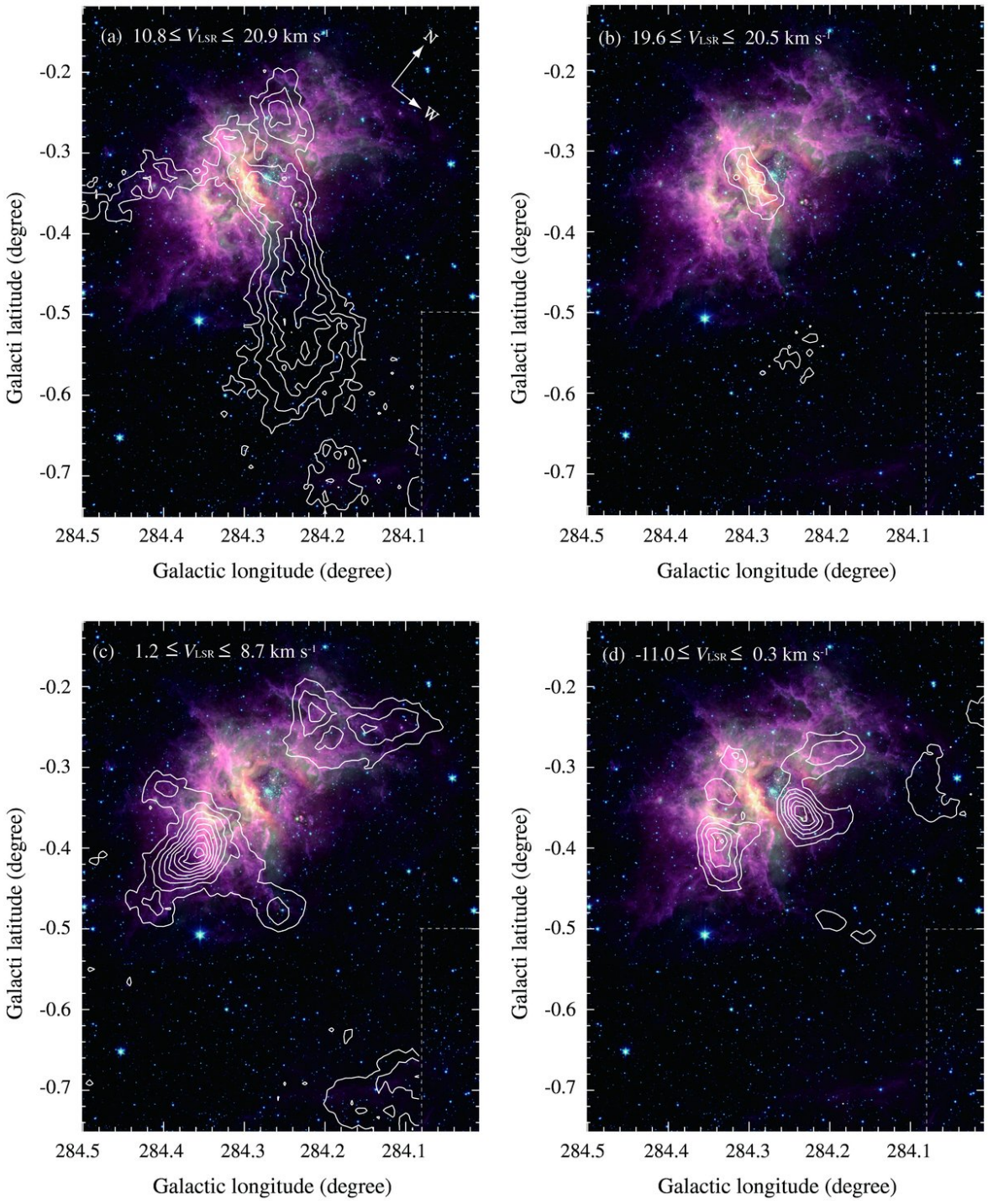


Figure 1.2 – ^{12}CO (J=2→1) observations of the region around RCW49 (Furukawa *et al.* 2009). The contours show CO intensity integrated over different velocity ranges, and overlaid onto a *Spitzer* GLIMPSE three-colour image. Blue, green and red indicate the intensity at 3.6, 4.5 and 8.0 μm.

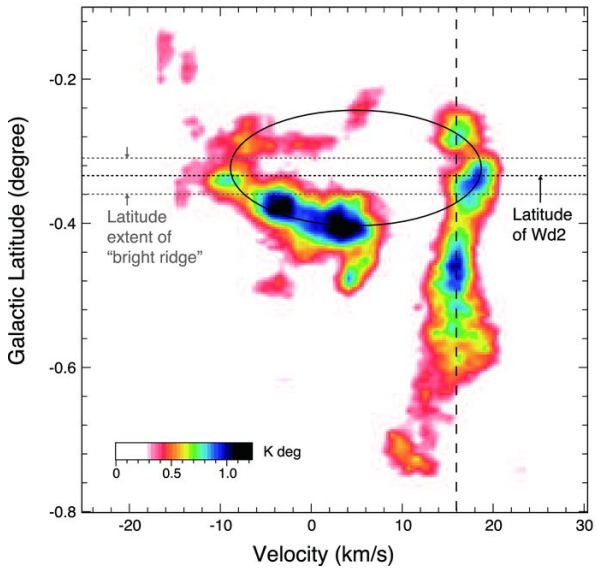


Figure 1.3 – ^{12}CO (J=2→1) position–velocity observations of the region around RCW49 (Furukawa *et al.* 2009). The colours show CO intensity integrated over the longitude range 284.2° to 284.4°. The solid black circle shows the velocities consistent with expansion around the cluster and the central regions of RCW49.

as seen in Figure 1.3. The cloud extends beyond the region in which the cluster would drive the cloud outwards. This is inconsistent with the explanation of Dame (2007), and instead suggests that the velocity of the 16 km s^{-1} cloud is not caused by the cluster.

The 19.6 km s^{-1} to 20.5 km s^{-1} range of Figure 1.2 shows how the 16 km s^{-1} cloud is associated with RCW49. This narrow range represents the highest velocity parts of the 16 km s^{-1} cloud, and clearly traces the bright ridge in the *Spitzer* image. This can also be seen Figure 1.3, where the cloud appears to be shifted to slightly higher velocities in the region of Westerlund 2 and the bright ridge.

The 1.2 km s^{-1} to 8.7 km s^{-1} range of Figure 1.2 shows the extent of the 4 km s^{-1} cloud. This cloud is not observed close to Westerlund 2 itself, but closely matches the extended parts of the nebula in the *Spitzer* image, again suggesting an association with the RCW49 system.

The −11.0 km s^{-1} to 0.3 km s^{-1} range of Figure 1.2 shows the extent of the −4 km s^{-1} cloud. This cloud is not clearly separated from the 4 km s^{-1}. The main peak matches the position of a region of star formation (Whitney *et al.* 2004), showing this is also associated with the RCW49 system. Figure 1.3 shows how the −4 km s^{-1} and 4 km s^{-1} form a continuous arc in velocity around the cluster. These can therefore be considered as the extensions of a single cloud centred on 0 km s^{-1}.

Ohama *et al.* (2010) present new observations of ^{13}CO (J=2→1), shown with the previous CO data in Figure 1.4. By combining this data they calculate temperatures and

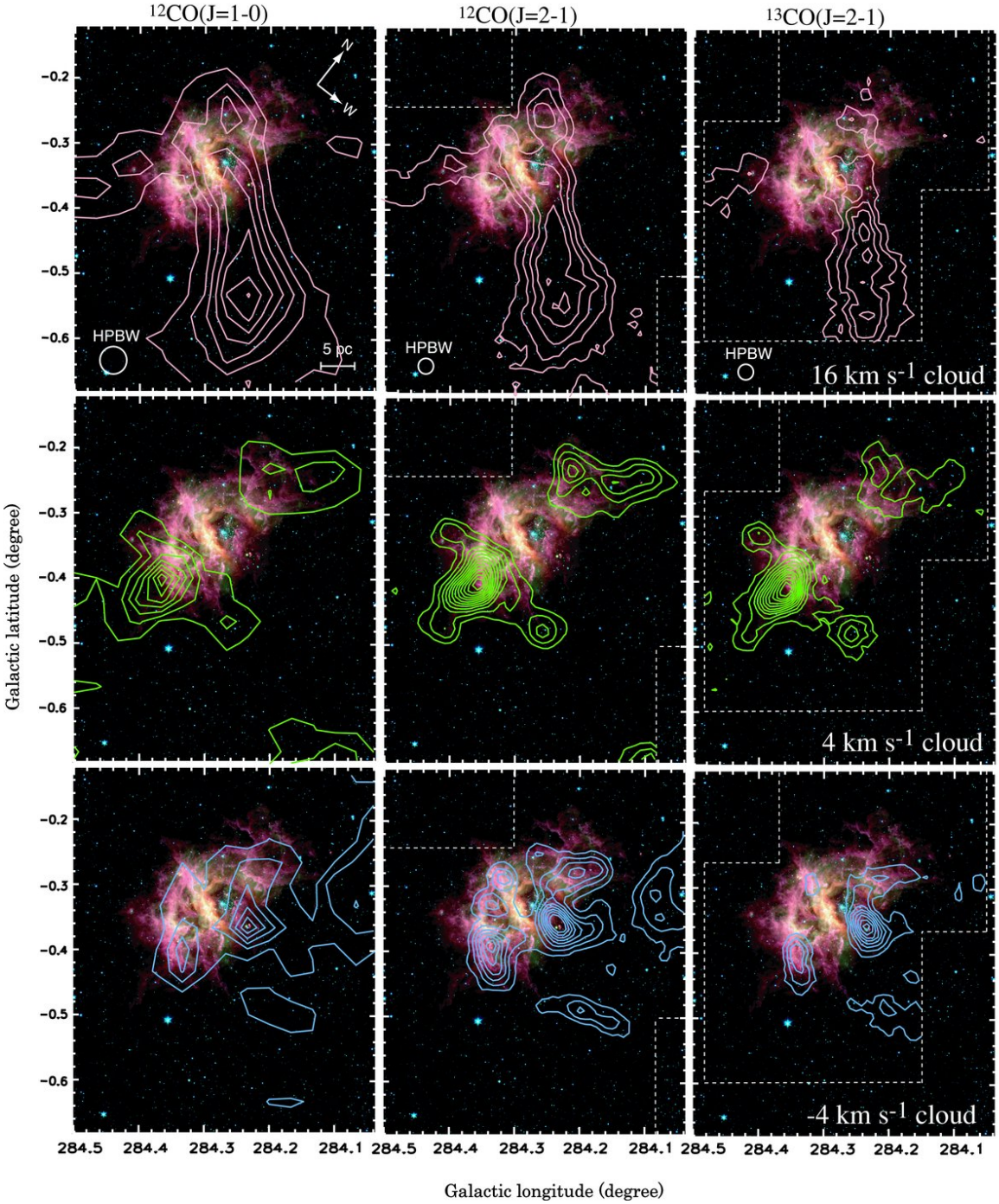


Figure 1.4 – CO observations of the region around RCW49 (Ohama *et al.* 2010). The contours show CO intensity integrated over different velocity ranges, and overlaid onto a *Spitzer* GLIMPSE three-colour image. The left column shows observations of ^{12}CO (J=1→0) data, the centre column ^{12}CO (J=2→1), and the right column ^{13}CO (J=2→1). The top row shows CO intensity integrated over the range 11.0 km s^{-1} to 20.9 km s^{-1}, the middle row 1.2 km s^{-1} to 8.7 km s^{-1}, and the bottom row −11.0 km s^{-1} to 0.3 km s^{-1}. The ^{12}CO (J=2→1) and ^{13}CO (J=2→1) observations have been smoothed for comparison with the ^{12}CO (J=1→0) data.

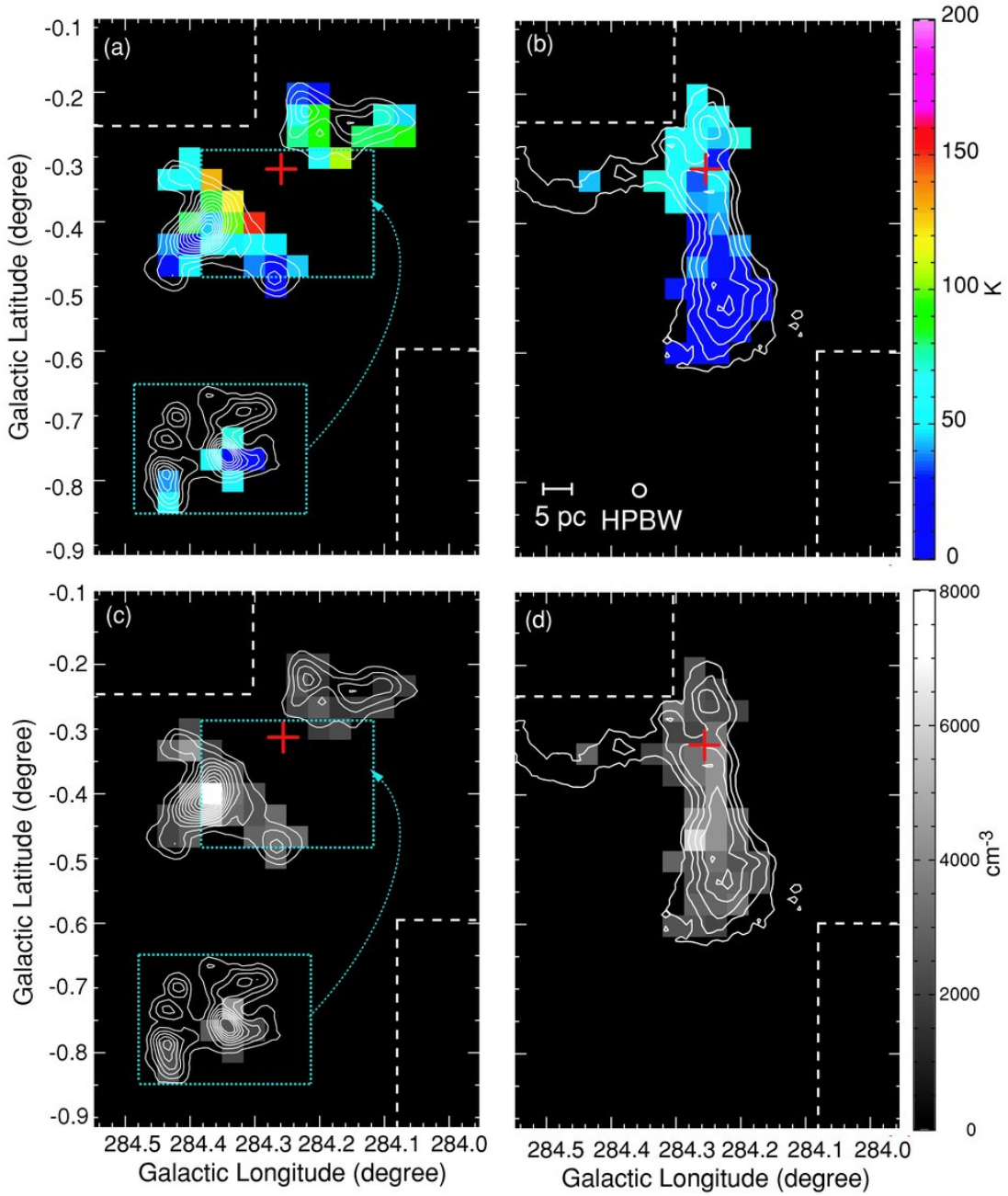


Figure 1.5 – CO temperature and density observations of the region around RCW49 (Ohama *et al.* 2010). The colours show gas temperature (top row) and density (bottom row) for the $4\,\mathrm{km\,s^{-1}}$ cloud (left column), the $16\,\mathrm{km\,s^{-1}}$ cloud (right column) and the $-4\,\mathrm{km\,s^{-1}}$ cloud (inset dotted box). The contours show the ^{12}CO (J=2→1) contours from Figure 1.4. The position of Westerlund 2 is shown with a red cross.

densities for the gas in the various clouds, as shown in Figure 1.5. A number of features show that these clouds are associated with the cluster.

The $16\,\mathrm{km\,s^{-1}}$ cloud shows a clear temperature gradient in Figure 1.5. The top end of the cloud, closer to Westerlund 2, is at approximately 30 K to 70 K, which is significantly warmer than would be expected if the cloud was distant from Westerlund 2. In the $4\,\mathrm{km\,s^{-1}}$ cloud the effect is even more dramatic. The edges of the cloud facing the cluster are significantly warmer than edges facing away from it, reaching temperatures of 70 K to 150 K. Finally, the density peak in the $-4\,\mathrm{km\,s^{-1}}$ cloud closely corresponds to the position of a number of YSOs detected by Whitney *et al.* (2004).

Dame (2007) suggests that the $\sim 4\,\mathrm{km\,s^{-1}}$ and $\sim 16\,\mathrm{km\,s^{-1}}$ clumps are simply parts of the $11\,\mathrm{km\,s^{-1}}$ giant molecular cloud found on larger scales, perturbed by the winds and ionization of the cluster. However, Furukawa *et al.* (2009) suggest this is unlikely for a number of reasons.

Firstly, the $16\,\mathrm{km\,s^{-1}}$ cloud extends well past the cluster, but its velocity is relatively constant, making it difficult to explain its velocity as the result of feedback from the cluster. Secondly, the addition of the new cloud at $-4\,\mathrm{km\,s^{-1}}$ changes the average velocity to $6\,\mathrm{km\,s^{-1}}$, suggesting that the clouds may not be directly associated with the larger GMC at $11\,\mathrm{km\,s^{-1}}$.

Finally, the output of the cluster since its formation may not have been sufficient to accelerate the clouds to their current velocity. Furukawa *et al.* (2009) estimate the mass of the $16\,\mathrm{km\,s^{-1}}$ cloud and the combined $0\,\mathrm{km\,s^{-1}}$ cloud (the combination of the $-4\,\mathrm{km\,s^{-1}}$ and $4\,\mathrm{km\,s^{-1}}$ clouds). They calculate masses of $(9.1 \pm 4.1)\,\mathrm{M_\odot}$ and $(8.1 \pm 3.7)\,\mathrm{M_\odot}$ respectively. By combining this with the velocities of the clouds they calculate a total kinetic energy of $(1.2 \pm 0.5) \times 10^{43}$ J. Rauw *et al.* (2007) estimate the total mechanical energy from the cluster since its formation as 3.6×10^{44} J, assuming no supernovae have yet exploded. This is a reasonable assumption for a cluster younger than 3 Myr, but must be considered against the jet and arc discovered by Fukui *et al.* (2009).

Furukawa *et al.* (2009) point out that in realistic simulations of expanding bubbles, such as those of Arthur (2008), only a few percent of the mechanical luminosity ends up in the gas velocity. Furthermore, the clouds subtend a small solid angle. Furukawa *et al.* (2009) conclude that, although not impossible, it seems unlikely that the velocity of the clouds is entirely due to the action of the cluster. They therefore propose that the cloud velocities are older than the cluster, and suggest that the clouds, now receding from each other, may have collided in the past. At their current velocities and assumed separation of ~ 40 pc, the clouds would have collided 4 Myr ago. This is consistent with the estimated age of the cluster of a few million years.

Source	Distance (kpc)	Object	Observations
Morgan, Whitford & Code 1953	1.4	5 stars	Spectroscopy
Ogura & Ishida 1975	1.4	320 stars	UBV photometry[1]
Stone 1978	2.84	HD 164 492	BV photometry
Lynds, Canzian & O'Neil 1985	1.68	Cluster stars	Optical extinction
Kohoutek, Mayer & Lorenz 1999	2.5/2.8[2]	Cluster stars	UBV photometry
Cambrésy *et al.* 2011	2.7 ± 0.5	Stars	NIR extinction[3]
Cambrésy *et al.* 2011	3.2 to 3.4	Mol. cloud	CO velocities[4]

[1] Fitting of (U-B)/(B-V) colour-colour diagram against ZAMS
[2] Values calculated using two different calibrations
[3] Extinction mapping using 2MASS J, H & K_s and *Spitzer*/IRAC 3.6, 4.5 and 5.8 μm colour/colour diagrams
[4] Velocity of the Trifid cloud taken from the CO survey of Dame, Hartmann & Thaddeus (2001). Distance calculated using Galactic model of McMillan & Binney (2010), with a solar Galactocentric radius of 7.8 kpc and a velocity of 247 km s^{-1}

Table 1.2 – List of distance estimates to M20, the Trifid Nebula. Where not otherwise specified 'photometry' and 'spectroscopy' refer to optical observations.

1.2.2 The Trifid Nebula (M20, NGC 6514)

The Trifid nebula, also known as M20 and NGC 6514, is an unusual combination of an emission nebula, a reflection nebula and a dark nebula. It contains an extremely young stellar cluster, and a HII region ionized by the O7.5 star HD 164 492A. The dark nebula is formed by dust lanes that trisect the nebula, giving it the nebula its name. The HII region has a size of approximately 3 pc, and is surrounded by a ring of denser gas (Cernicharo *et al.* 1998).

The nebula appears to be part of the complex of molecular gas surrounding the supernova remnant W28 (Lefloch, Cernicharo & Pardo 2008). The remnant is believed to be approximately 10^5 yr old and at a distance of approximately 2 kpc, with uncertainties of at least 50% (Kaspi *et al.* 1993).

Lynds, Canzian & O'Neil (1985) estimated an often-quoted distance to the nebula of 1.68 kpc, but as shown in Table 1.2 this distance is disputed and we use a value of 2.7 kpc supported by more modern estimates (Cambrésy *et al.* 2011; Kohoutek, Mayer & Lorenz 1999). The cluster contains approximately 500 $M_\odot$ of stars. It is extremely young, and the HII region is no older than 0.3 Myr (Cernicharo *et al.* 1998; Lefloch & Cernicharo 2000).

The nebula has been estimated to have a total H_2 mass of approximately 2.2×10^5 $M_\odot$ (Torii *et al.* 2011) for the area examined below or 5.8×10^5 $M_\odot$ for a larger area (Cambrésy *et al.* 2011). The nebula is a region of active star formation. Cernicharo *et al.* (1998) observed a number of condensations in 1.3 mm thermal dust emission which they labelled TC0 through to TC4. They observed an optical jet in TC2, and identified TC1 to TC4 as

containing extremely young YSOs with associated protostellar outflows.

Later authors such as Rho *et al.* (2006) and Lefloch, Cernicharo & Pardo (2008) separated the condensations into larger clumps and smaller cores. TC3 and TC4 were identified as clumps which contain multiple cores and associated protostars. Lefloch, Cernicharo & Pardo (2008) extended the TC notation to include 33 condensations. Jets and outflows were observed in some cores using a range of molecular lines, indicating the presence of protostellar objects, while other cores did not contain protostellar objects. The overall positioning of the clumps is consistent with a second generation of star formation triggered by the expansion of the HII region.

Using *Spitzer* IRAC and MIPS data Rho *et al.* (2006) found approximately 160 YSOs in the cluster. Evidence for young stellar objects has also been obtained in the X-ray; Rho *et al.* (2001) identified 44 YSOs and 41 T Tauri stars by comparing X-ray point sources with 2MASS observations. Figure 1.6 shows *Spitzer* imagery of the nebula, showing both YSOs and polycyclic aromatic hydrocarbon emission associated with the HII region.

Torii *et al.* (2011) observed the Trifid nebula in the ^{12}CO (J=1→0), ^{12}CO (J=2→1), ^{13}CO (J=1→0) and ^{13}CO (J=2→1) lines. Figure 1.7 shows the ^{12}CO (J=1→0) emission integrated over the range $0\,\mathrm{km\,s^{-1}}$ to $30\,\mathrm{km\,s^{-1}}$ in the region surrounding M20, including the W28 supernova remnant. It shows that M20 is part of a large filament of molecular gas running southwest to northeast.

Figure 1.8 shows the molecular gas in the region of M20. Torii *et al.* (2011) classify the emission into a number of distinct clouds. The emission between $-1.9\,\mathrm{km\,s^{-1}}$ to $5.6\,\mathrm{km\,s^{-1}}$ is centred on M20, and approximately matches some or all of the three dust lanes. This can be seen more clearly in Figure 1.9, which shows the ^{12}CO (J=2→1) emission overlaid on an optical image. This emission is referred to as the $2\,\mathrm{km\,s^{-1}}$ cloud, and is associated with the TC1, TC8, TC10, TC11, and TC13 cores, of which TC1, TC8 and TC13 are star-forming (Lefloch, Cernicharo & Pardo 2008).

Emission between 5.6 and $13.1\,\mathrm{km\,s^{-1}}$ (the $8\,\mathrm{km\,s^{-1}}$ clouds) is divided into a central cloud C, three clouds to the northeast, northwest and south labelled NE, NW and S, and three smaller clouds to the northeast labelled NE1, NE2 and NE3. The central cloud C peaks in the same location as the $2\,\mathrm{km\,s^{-1}}$ cloud. Figure 1.9 shows the location of these clouds and shows that they corresponds with visible dark features on the optical image. The NW cloud is associated with the clump TC0 which is not forming stars, while the S cloud is associated with the star-forming core TC2. The NE cloud is not associated with the known cold dust cores. It appears towards the end of a filament of molecular gas extending from the central cloud, comprised of the NE2, NE and NE3 clouds (Torii *et al.* 2011). This filament also appears as an optical dust lane visible in Figure 1.9.

As can be seen from Figure 1.8 the majority of the emission is within the range $13.1\,\mathrm{km\,s^{-1}}$ to $26.3\,\mathrm{km\,s^{-1}}$, at or around approximately $18\,\mathrm{km\,s^{-1}}$. The emission at lower

Figure 1.6 – Three colour *Spitzer* image (Rho *et al.* 2006). Blue and green represent IRAC 4.5 and 8.0 μm respectively; red represents MIPS 24 μm. Green shows polycyclic aromatic hydrocarbon emission and hot dust grains appear in red, while YSOs appear as red or yellow point sources.

velocities is the most massive of all the clouds, and extends from the centre of M20 to the south. It is referred to as the 18 $\mathrm{km\,s^{-1}}$ cloud, and is associated with the star-forming core TC9. The emission at slightly greater velocities forms a cloud that is labelled the TC3 and TC4 cloud (Torii *et al.* 2011). These clumps are currently forming stars (Lefloch, Cernicharo & Pardo 2008).

Figure 1.9 shows that the majority of lower velocity clouds are associated with visible dust lanes, showing they are in front of the cloud. By contrast the higher velocity clouds (the TC3 and TC4 cloud and the 18 $\mathrm{km\,s^{-1}}$ cloud) do not show clear association with the visible dust lanes, showing them to be within or behind the nebula. Lefloch & Cernicharo (2000) showed that the TC3 and TC4 clouds are at the boundary of the HII region.

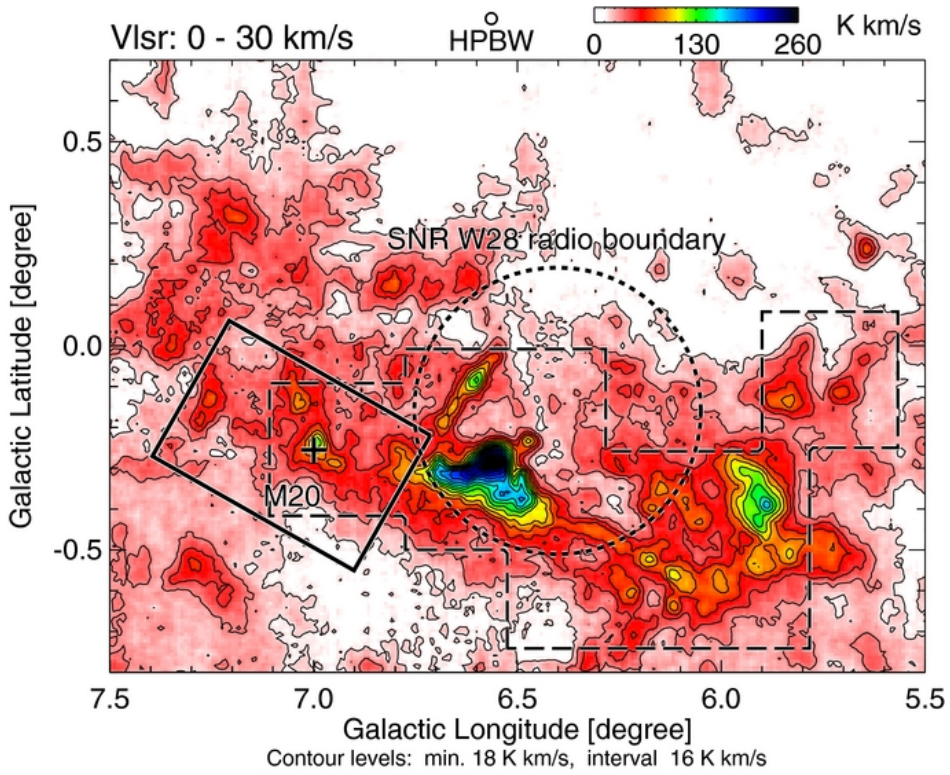


Figure 1.7 – ^{12}CO (J=1→0) observations of the region around M20, integrated between 0 km s^{-1} to 30 km s^{-1} (Torii *et al.* 2011). Dashed lines indicate regions also observed in ^{13}CO (J=2→1). The solid box indicates the region of M20 shown in Figure 1.8.

The exceptions are the central C cloud and the southern S cloud, which are not clearly associated with visible dust lanes but do closely match the 18 km s^{-1} cloud.

Torii *et al.* (2011) also point out that while the extinction to the ionizing star HD 164 492A is only approximately 1.3 (Lynds, Canzian & O'Neil 1985), the total extinction A_V in the central regions of the cluster as measured by Cambrésy *et al.* (2011) is approximately 10 to 15. This suggests that HD 164 492A is towards the closer side of the nebula, while the majority and gas and dust is behind the star.

Figure 1.10 shows a position–velocity maps of the ^{12}CO (J=2→1) emission in M20. The map is integrated over the limits of declination shown in Figure 1.8, and shows the approximate division of emission into the 2 km s^{-1} cloud, a collection of clouds at approximately 8 km s^{-1}, and the 18 km s^{-1} cloud.

Figure 1.11 shows the estimated temperature in each of six main cloud components, calculated using large velocity gradient analysis at an assumed distance of 2.7 kpc. The temperature in the 2 km s^{-1}, C, NW and S clouds is approximately 30 K to 50 K. This is significantly higher than the average background temperature of approximately 10 K, as found in the NE and 18 km s^{-1} clouds.

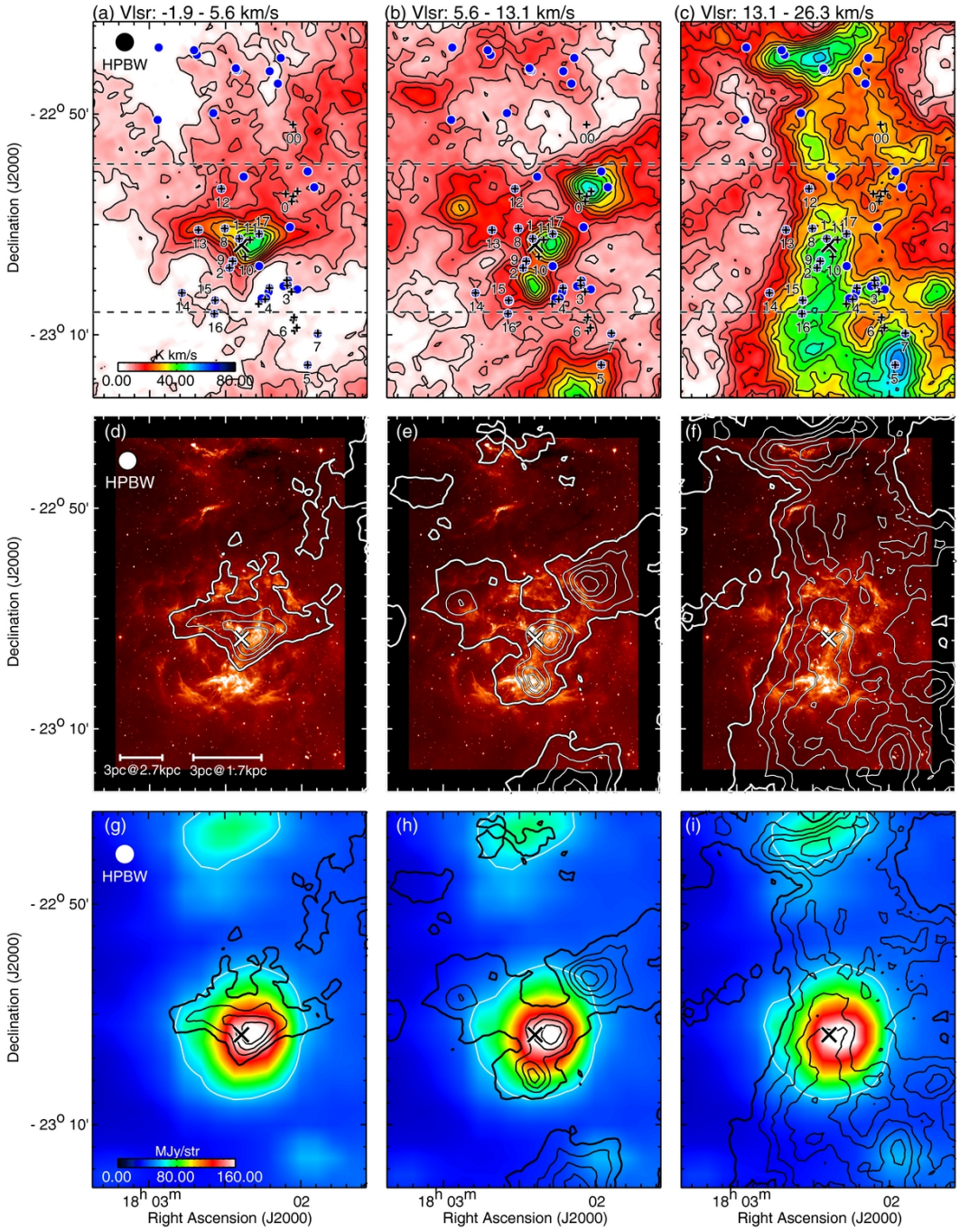


Figure 1.8 – Comparison of ^{12}CO (J=2→1) observations with 8.0 µm *Spitzer* and 25 µm IRAS observations. The top row shows ^{12}CO (J=2→1) observations. The middle row shows an 8.0 µm *Spitzer* image overlaid with ^{12}CO (J=2→1) contours. The bottom row shows a 25 µm IRAS image overlaid with ^{12}CO (J=2→1) contours. The three columns represent velocity ranges in the CO observations of $-1.9\,\mathrm{km\,s^{-1}}$ to $5.6\,\mathrm{km\,s^{-1}}$, $5.6\,\mathrm{km\,s^{-1}}$ to $13.1\,\mathrm{km\,s^{-1}}$ and $13.1\,\mathrm{km\,s^{-1}}$ to $26.3\,\mathrm{km\,s^{-1}}$ respectively. The large × symbol indicates the ionizing star HD 164 492A. The small + symbols indicate the cold dust cores of Lefloch, Cernicharo & Pardo (2008). The filled circles represent Class 0/I YSOs identified from *Spitzer* observations. Dotted lines indicate the integration limits in declination of Figure 1.10.

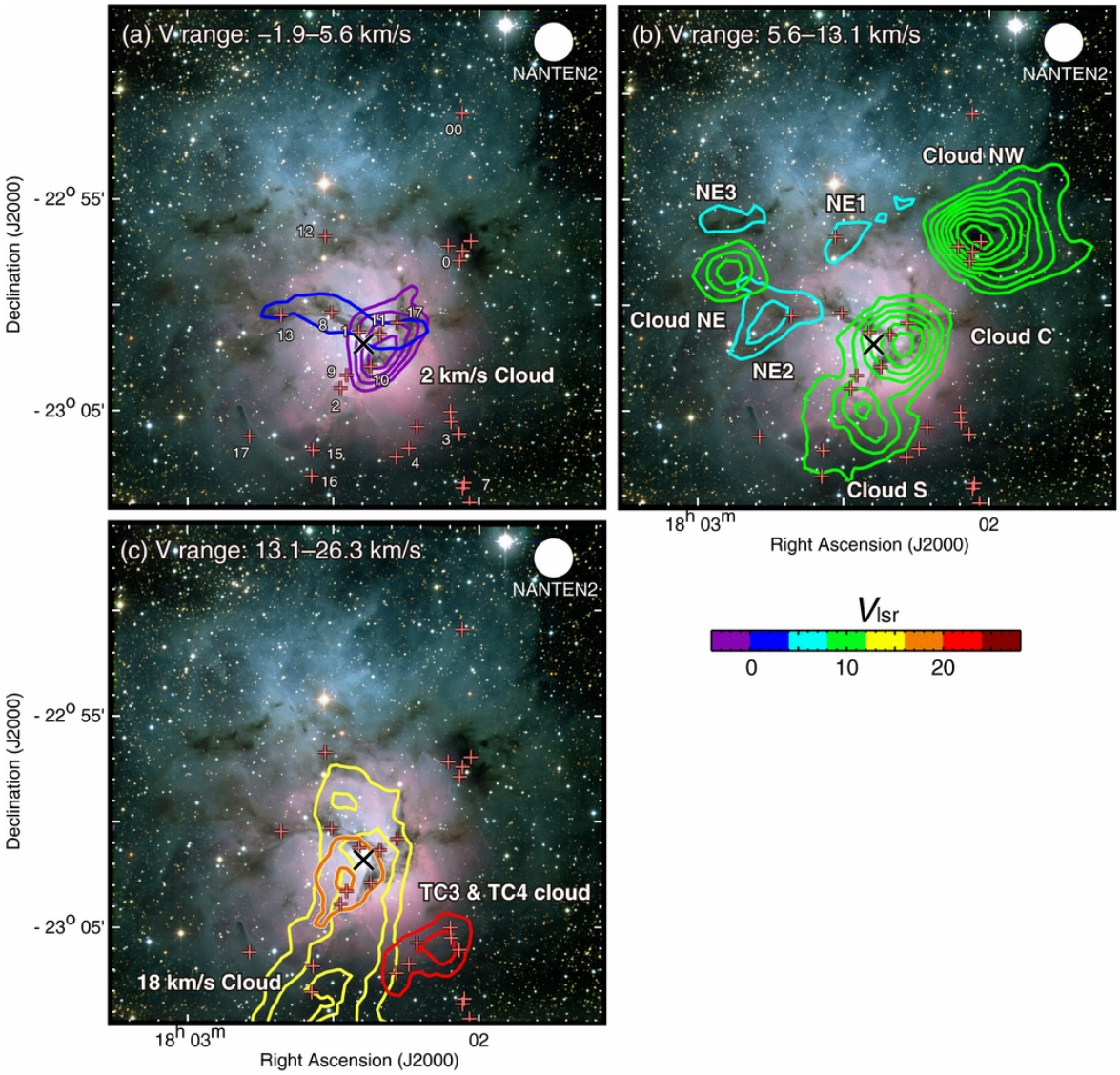


Figure 1.9 – Position of different velocity clouds around M20 (Torii *et al.* 2011). The contours show ^{12}CO (J=2→1) intensity at a range of velocity ranges, indicated by the colour bar, and are overlaid on an optical image (credit: Todd Boroson/NOAO/AURA/NSF). The large × symbol indicates the ionizing star HD 164 492A. The small + symbols indicate the cold dust cores of Lefloch, Cernicharo & Pardo (2008).

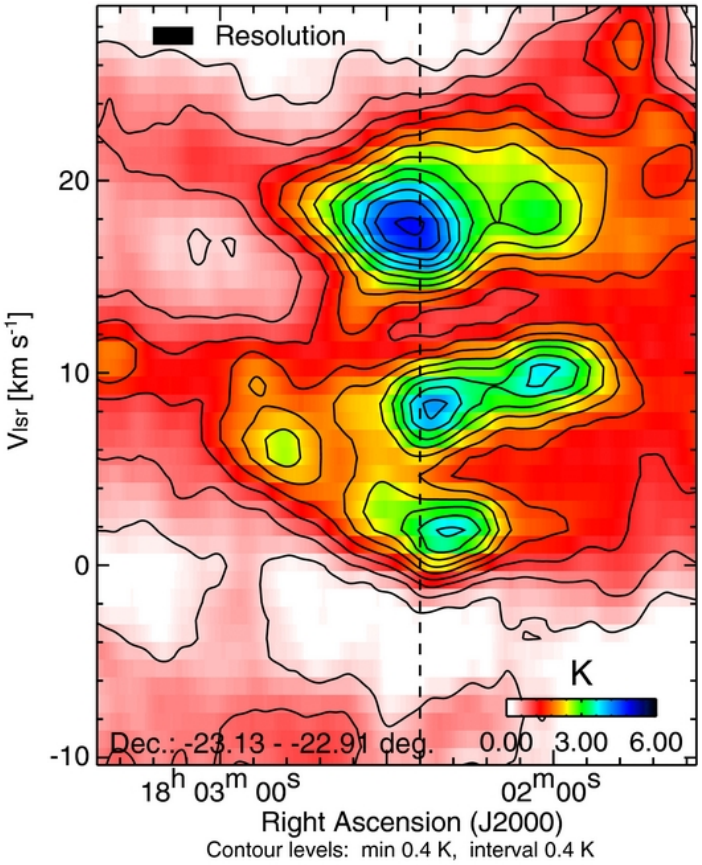


Figure 1.10 – Position-velocity diagram of ^{12}CO (J=2→1) emission across M20 (Torii *et al.* 2011). The dashed line indicates the position of the ionizing star HD 164 492A. The limits of integration in declination are shown as dotted lines in Figure 1.8.

The stars in the cluster can be separated into two generations of star formation. The cluster stars, including the ionizing star HD 164 492A, form the first generation of stars. These stars are no older than 1 Myr, and therefore no stars from this generation should have exploded as supernovae. The second generation of stars consists of the YSOs found in the dense cores. Some of the observed pre-stellar clumps are gravitationally bound and can be expected to collapse and form stars that will also form part of this generation (Lefloch, Cernicharo & Pardo 2008).

The stars of the second generation are likely the result of triggered star formation, either as a result of the expanding HII region (Lefloch & Cernicharo 2000) or due to interaction with the supernova that formed the remnant W28 (Lefloch, Cernicharo & Pardo 2008). This does not provide an explanation for the origin of the first generation of stars. Torii *et al.* (2011) propose that the first generation was formed by the collision of two molecular clouds. This would be similar to the method proposed by Furukawa *et al.* (2009) for the formation of Westerlund 2 and described in Section 1.2.1.

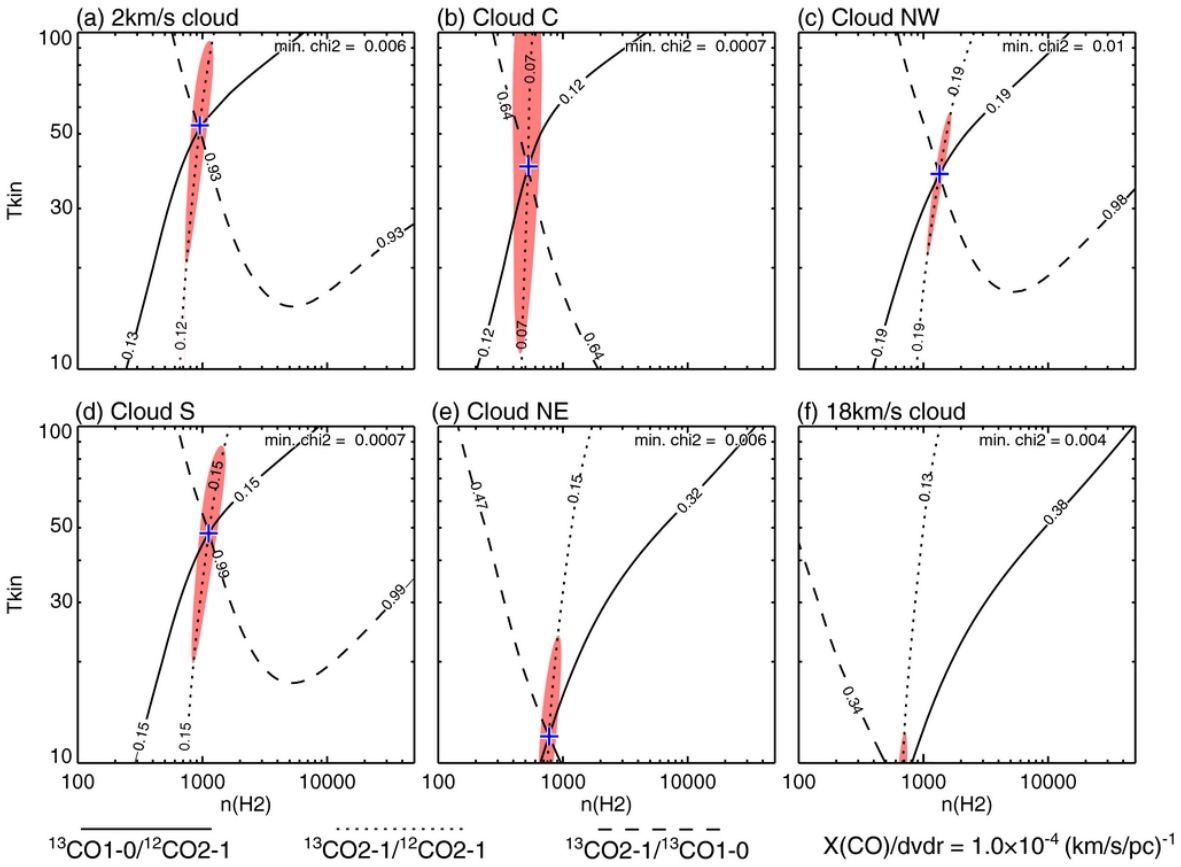


Figure 1.11 – Temperature estimates of the molecular clouds surrounding M20 from large velocity gradient analysis (Torii *et al.* 2011). The clouds are assumed to be at a distance of 2.7 kpc.

Figure 1.9 shows the IRAS 25 µm emission from M20 is concentrated in the centre, close to the ionizing star. They also show that the energy output of the cluster is dominated by the first generation stars. Torii *et al.* (2011) show that only the clouds close to this emission and the ionizing star have elevated temperatures, and that the clouds are warmest close to the centre of M20. This heating is therefore explained as shock heating from the stellar wind of the ionizing star and other first generation stars, showing that these clouds are clearly associated with the cluster. Furthermore, over the estimated lifetime of the HII region of 0.3 Myr, and at an assumed distance of 2.7 kpc, the clouds can only have moved approximately 2 pc.

Torii *et al.* (2011) therefore identify the 2 km s^{-1} and C clouds as potential parent clouds for the first generation stars. The clouds have average velocities of 1.4 km s^{-1} and 8.9 km s^{-1} respectively, giving a current velocity difference of 7.5 km s^{-1}. As previously noted, the 2 km s^{-1} cloud is associated with visible dust lanes, indicating it is in front of M20, while cloud C shows no such association and is likely behind or within M20. The

18 km s^{-1} cloud does not show the heating seen in other clouds, and so is assumed to be separate to the M20 system and not directly involved in the formation of the cluster.

The two clouds are therefore moving away from each other. This would be expected if the clouds have recently collided over a timescale of approximately 1 Myr. Alternatively the clouds may be receding due to the feedback from the stellar cluster. Torii *et al.* (2011) calculate the total kinetic energy of the 2 km s^{-1}, C, NW, S and NE clouds as 4.7×10^{41} J. The total energy output from the mechanical luminosity of the cluster is estimated at 1.5×10^{42} J. This would therefore require 30% of the energy output from the cluster to have been absorbed by the clouds, but as described in Section 1.2.1 realistic simulations of expanding bubbles show that only a few percent of the mechanical luminosity ends up in the gas velocity (Arthur 2008; Furukawa *et al.* 2009; Torii *et al.* 2011). It is also unlikely that the clouds subtend the majority of the solid angle visible from the cluster, further reducing the available energy.

Furthermore Torii *et al.* (2011), show that the pattern shown in Figure 1.10 does not match that expected for a classic shell expansion. In a classic shell expansion the greatest velocities would be along the centre of the shell, whereas here greater velocities are actually observed away from the ionizing star.

Finally, at an assumed separation of 2 pc, Torii *et al.* (2011) estimate that 2.5×10^4 M$_\odot$ would be required to gravitationally bind the clouds. The total estimated mass of the five clouds is only 8.3×10^3 M$_\odot$, and is therefore considerable less than required to bind the M20 system. This increases the likelihood that the first generation of stars was triggered by an external event rather than being solely due to isolated gravitational fragmentation.

However, there are a number of caveats to this model. Torii *et al.* (2011) include the NW, S and NE clouds in their estimation of the kinetic energy. These are at a similar velocity to the C cloud, and so may indeed be part of a larger cloud. While the NW and S clouds are warmer than the background gas, indicating their close association to the first generation stars, this is much less pronounced for the NE cloud, and it may not be directly associated. In some ways, this enhances the case for cloud–cloud collision; if a large collection of gas is moving at similar velocities both close to and away from the cluster then it is more likely that the velocity is systemic and not a result of the feedback from the cluster. However it may be fairer to calculate the kinetic energy requirements using only the 2 km s^{-1} and C cloud, which have a combined mass of 3.2×10^3 M$_\odot$.

Secondly, Torii *et al.* (2011) assume each cloud is receding from the cluster at 7.5 km s^{-1}, which is the velocity difference between the 2 km s^{-1} cloud and the C cloud. If we ignore the mass difference between the two clouds, and assume the two clouds were initially at rest at an intermediate velocity, then each clouds only needs to recede at 3.75 km s^{-1}, leading to a total kinetic energy of 4.8×10^{40} J, which is only a few percent of the total available mechanical energy. This is similar to that calculated for Westerlund

2 as described in Section 1.2.1 (Furukawa *et al.* 2009). Given the poor coupling between the mechanical energy of the winds and the gas, and the small solid angle subtended by the $2\,\mathrm{km\,s^{-1}}$ cloud and the C cloud, this still supports the hypothesis of a cloud–cloud collision leading to a first generation of stars.

1.2.3 Sagittarius B2

The Galactic centre contains more than one hundred molecular clouds and clumps (Miyazaki & Tsuboi 2000). Approximately 10% of the total molecular gas of the Milky Way is believed to reside in these innermost regions (Güsten 1989; Miyazaki & Tsuboi 2000), with a total mass of approximately $10 \times 10^8\,\mathrm{M_\odot}$ within a few hundred parsecs of the Galactic centre (Güsten 1989; Sodroski *et al.* 1995; Tsuboi, Handa & Ukita 1999). Clouds in the Galactic centre often have an average density 5 to 10 times greater than clouds elsewhere (Bally *et al.* 1987; Miyazaki & Tsuboi 2000). A large number of molecules have been observed in these clouds due to their high average densities.

The velocity dispersions of these clouds are also approximately an order of magnitude higher than clouds away from the Galactic centre (Bally *et al.* 1987; Miyazaki & Tsuboi 2000; Tsuboi, Handa & Ukita 1999). For clouds away from the Galactic centre the peculiar velocities of the clouds are typically much larger than the internal velocity dispersion. Galactic centre clouds often have internal velocity dispersions of $20\,\mathrm{km\,s^{-1}}$ to $50\,\mathrm{km\,s^{-1}}$, which is comparable to the velocities of the clouds. These clouds also have strongly non-circular orbits, and may collide as they pass repeatedly through the inner disc of molecular gas (Bally *et al.* 1988).

Sagittarius B2 (Sgr B2) is the dominant molecular cloud in the region (Güsten 1989). Located approximately 120 pc from the Galactic centre, the cloud has a size of approximately 20 pc to 50 pc (Lis & Goldsmith 1989; Scoville, Solomon & Penzias 1975). The average density of molecular hydrogen in the cloud is a few thousand atoms per $\mathrm{cm^3}$, which is higher than in typical giant molecular clouds (Lis & Goldsmith 1990; Scoville, Solomon & Penzias 1975). Molecular emission is observed over a very large velocity range of more than $60\,\mathrm{km\,s^{-1}}$. The total molecular mass in the cloud is estimated as approximately $3 \times 10^6\,\mathrm{M_\odot}$ to $10 \times 10^6\,\mathrm{M_\odot}$ (Lis & Goldsmith 1989, 1990; Scoville, Solomon & Penzias 1975; Tsuboi, Handa & Ukita 1999).

High-mass star formation is on-going in Sgr B2. Three dense clumps are observed in Sgr B2; a central main clump and smaller north and south clump, labelled Sgr B2(M), Sgr B2(N) and Sgr B2(S) respectively. All three regions contain HII regions (Gaume *et al.* 1995), and high-velocity massive molecular outflows have been observed from both Sgr B2(N) and Sgr B2(M). Such outflows are a strong indicator of star formation (Lis *et al.* 1993; Mehringer 1995; Vogel, Genzel & Palmer 1987).

Maser emission is also observed from both Sgr B2(M) and Sgr B2(N). Masers are believed to be tracers of high-mass star formation (Minier *et al.* 2003). Sgr B2 contains

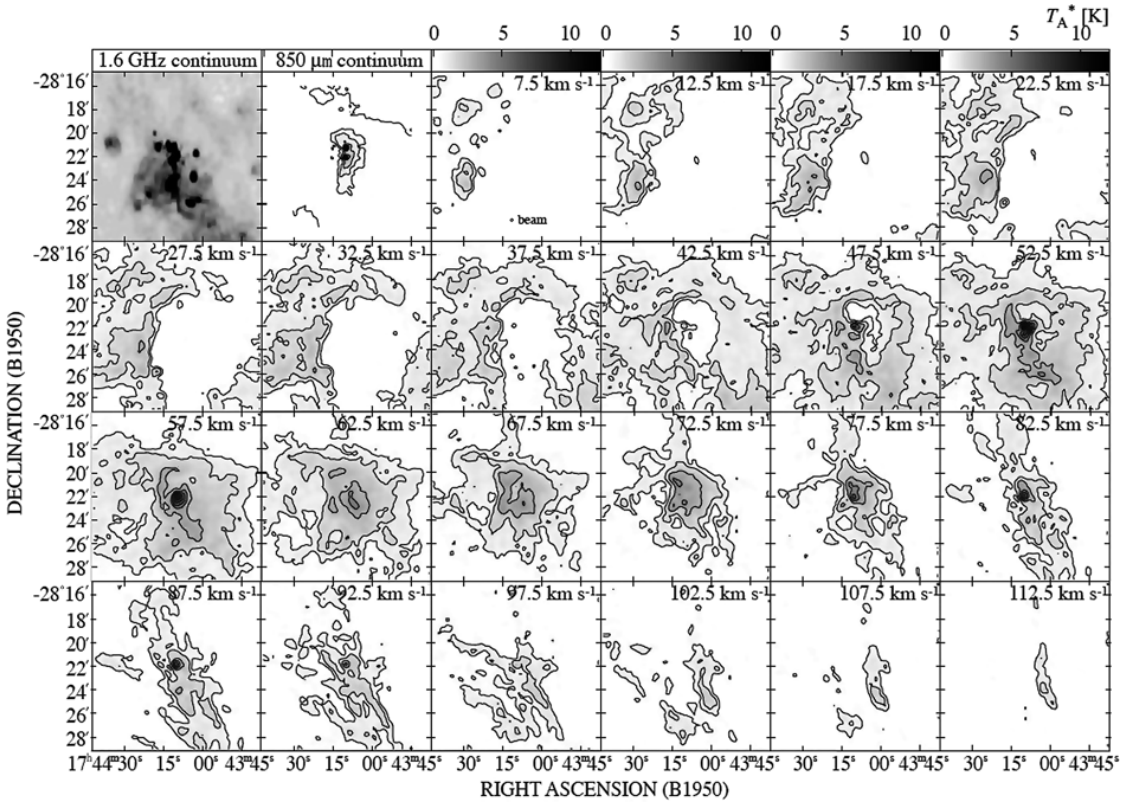


Figure 1.12 – ^{13}CO (J=3→2) observations of Sgr B2 (Hasegawa *et al.* 2008). The top left panel shows 1.616 GHz continuum emission (Liszt 1992). The second panel shows 850 μm continuum emission observed with *SCUBA* (Pierce-Price *et al.* 2000). The remaining images show ^{13}CO (J=3→2) emission averaged over 5 km s^{-1} intervals.

over a hundred H_2O and OH masers (Benson & Johnston 1984; Gaume & Claussen 1990; Gaume & Mutel 1987). Two classes of CH_3OH masers have been observed at 6 GHz (Houghton & Whiteoak 1995) and 44 GHz (Mehringer & Menten 1997), together with less common H_2CO masers (Gardner *et al.* 1986; Mehringer, Goss & Palmer 1994) and at least one SiO maser (Hasegawa *et al.* 1986; Morita *et al.* 1992; Shiki, Ohishi & Deguchi 1997). Sgr B2 also contains more than 50 compact, ultra-compact and hyper-compact HII regions (Benson & Johnston 1984; Gaume & Claussen 1990; Gaume *et al.* 1995; Mehringer *et al.* 1993).

Hasegawa *et al.* (1994) observed ^{13}CO (J=1→0) emission from Sgr B2 and identified a number of components. As shown in Figure 1.12, extended emission surrounding Sgr B2 is observed at velocities of approximately 20 km s^{-1} to 60 km s^{-1}. A large shell structure, labelled the Shell, surrounds Sgr B2 at velocities of approximately 20 km s^{-1} to 40 km s^{-1}. Larger scale maps, such as those of Bally *et al.* (1987), show a large complex of emission surrounding this shell. At higher velocities the shell structure disappears except for a small hole at velocities of 40 km s^{-1} to 50 km s^{-1}. This feature is labelled the Hole. At velocities

of approximately $70\,\mathrm{km\,s^{-1}}$ to $80\,\mathrm{km\,s^{-1}}$, the extended emission has disappeared leaving only a small region of emission labelled the Clump. This is positioned at the approximate centre of both the Shell and the Hole. Hasegawa *et al.* (1994) estimate the mass of the Clump to be $10^6\,\mathrm{M_\odot}$.

Figure 1.13 shows both the region in more detail. The Hole and the Clump, shown in panels (*a*) and (*b*) respectively, show an obvious morphological similarity. The two shapes approximately match if the Clump is shifted approximately $30''$ to the south and $40''$ to the west, as shown more clearly in panels (*c*) and (*d*).

Hasegawa *et al.* (1994) compare the positions of the shell, hole and clump to the simulations of Habe & Ohta (1992) (described in more detail in Section 1.3). Figure 1.14, reproduced from their Figure 5f, shows the result of a small dense cloud colliding with a larger less dense cloud. The figure shows the density and velocity of the gas after 0.6 freefall times of the initial small cloud. The collision of a smaller, more dense clump into a larger, less dense cloud results in a conical bow shock. This is consistent the observations of the shell, hole and clump.

In this model, the Shell represents the lack of gas observed at low velocities due to the large bow shock behind the colliding cloud. Emission is seen at these velocities only outside of this bow shock. The Clump represents the small, dense colliding cloud moving at a velocity significantly different to the cloud it is colliding with. The lack of emission in the Hole represents the interface of the collision, explaining the similar shapes. The Clump has removed the gas at these velocities, accelerating them to higher velocities. The offset between the Clump and the Hole is then due to velocity of the Clump parallel to the sky. Hasegawa *et al.* (1994) conclude that the star formation in Sgr B2 may have been triggered by the collision of the Clump into the larger cloud surrounding the Shell, and creating the Hole and Shell in the process.

The star-forming regions Sgr B2(N), Sgr B2(M) and Sgr B2(S) are aligned along one edge of the Hole, suggesting they have formed along the interface of the collision. This suggests that star-formation in these clumps may have been triggered by the interaction between the Clump and the gas surrounding the Hole.

Figure 1.15 shows the Hole in more detail using emission from ^{13}CO (J=3→2). The hole is surrounded by a ring of clumps. These clumps represent gas which is warmer, denser, or both. This is consistent with shock-induced compression and heating at the interface between the Clump and the surrounding gas. This can also be compared with the 'hot ring' observed by de Vicente, Martin-Pintado & Wilson (1997). They found a hot ring of radius 2 pc of a similar size to the Hole. This ring consisted of gas with a kinetic temperature of 100 K to 120 K. This could also be shock heating at the collision interface, although the hot ring is not centred precisely on the Hole. The hot ring is instead centred on the HII regions of the star-forming clumps, and so could be the edge of an expanding

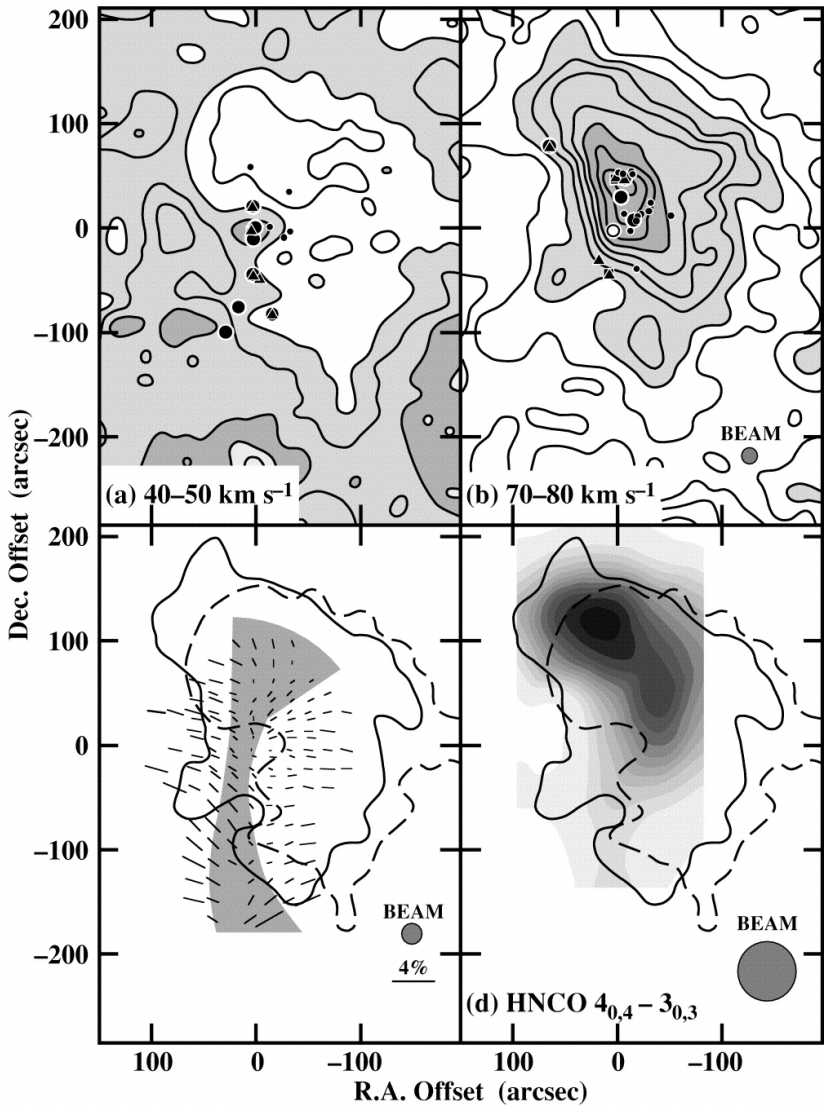


Figure 1.13 – ^{13}CO (J=1→0) observations of the 'hole' and 'clump' (Sato *et al.* 2000). Panels (*a*) and (*b*) show emission integrated over the velocity ranges 40 km s^{-1} to 50 km s^{-1} and 70 km s^{-1} to 80 km s^{-1} respectively. Light and dark shading indicate regions with an integrated intensity above 15 and 25 K km s^{-1} in panel (*a*) and above 20 and 40 K km s^{-1} in panel (*b*). Molecular masers with a velocity of less than 65 km s^{-1} are shown in panel (*a*). Molecular masers with a velocity of more than 65 km s^{-1} are shown in panel (*b*). Squares show OH masers (Gaume & Claussen 1990); triangles show H_2CO masers (Mehringer, Goss & Palmer 1994). Large and small filled circles show CH_3OH masers observed at 6.7 GHz (Houghton & Whiteoak 1995) and 44 GHz (Mehringer & Menten 1997) respectively. The open circle shows an SiO maser (Hasegawa *et al.* 1986; Morita *et al.* 1992). Panel (*c*) shows polarization at 350 μm (Dowell *et al.* 1998). Assuming a constant background field, and that the polarization tracks the local magnetic field, the shaded region indicates areas where the local magnetic field appears to be different to the background field. Panel (*d*) shows the integrated intensity of HNCO (Minh *et al.* 1998). In panels (*c*) and (*d*), the dashed contour shows the 15 K km s^{-1} contour from panel (*a*) and the solid contour shows the 20 K km s^{-1} contour from panel (*b*). The contours illustrate the position of the Hole and the Clump respectively.

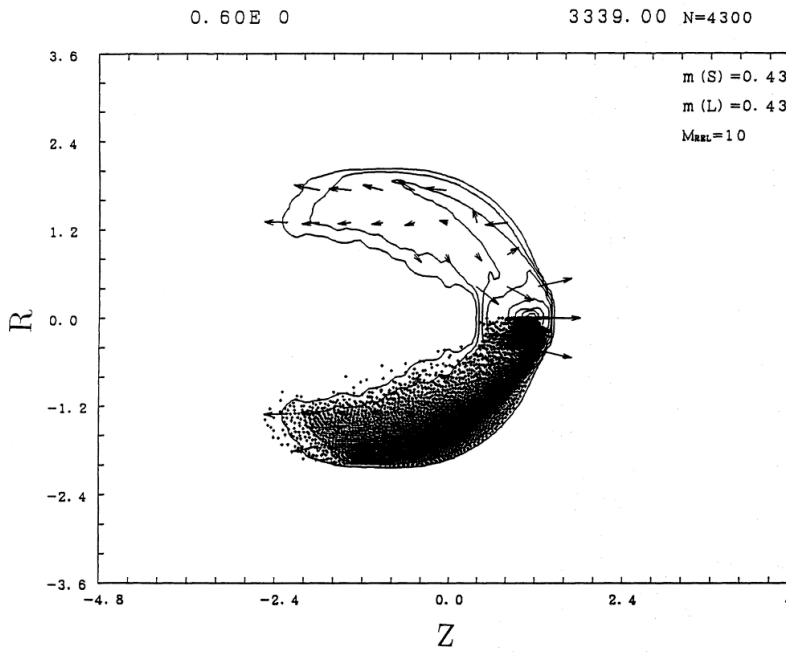


Figure 1.14 – Results of a collision between a small dense cloud and a large less dense cloud (Figure 5f of Habe & Ohta 1992). The simulations of Habe & Ohta 1992 are described in more detail in Section 1.3. The large cloud has a mass four times larger than the small cloud. The contours show density; arrows show gas velocities. The time after the initial contact is 0.6 in units of the freefall time of the initial small cloud. SPH particles are plotted on and below the $z = 0$ plane.

ionized bubble (de Vicente, Martin-Pintado & Wilson 1997).

Sato *et al.* (2000) note a selection of other observations consistent with a cloud–cloud collision, which we examine below.

Sato *et al.* noted that at velocities of approximately 65 km s^{-1}, observations of ^{13}CO show an absorption minimum while C^{18}O is observed in emission. The self-absorption is observed over a large extended plateau, and suggests a region of relatively high density. The velocity of this dense gas is intermediate between the velocities of the Clump and the velocity of the gas surrounding the Hole.

Sato *et al.* compare this extended emission to the simulations of Nagasawa & Miyama (1987). These simulations show that a cloud–cloud collision leads to the production of an extended component at a velocity between that of the two clouds. The extended plateau of absorbing gas may be this extended emission, collected in a shock-compressed layer by the cloud-cloud collision.

Observations of SiO seem to trace high-velocity shocked gas (Martin-Pintado, Bachiller & Fuente 1992; Mikami *et al.* 1992). They suggest that SiO is heavily depleted onto dust grains in quiescent gas, as emission from SiO is often undetectable in such gas. Shocks partially destroy the grains, dramatically increasing the abundance in the gas phase.

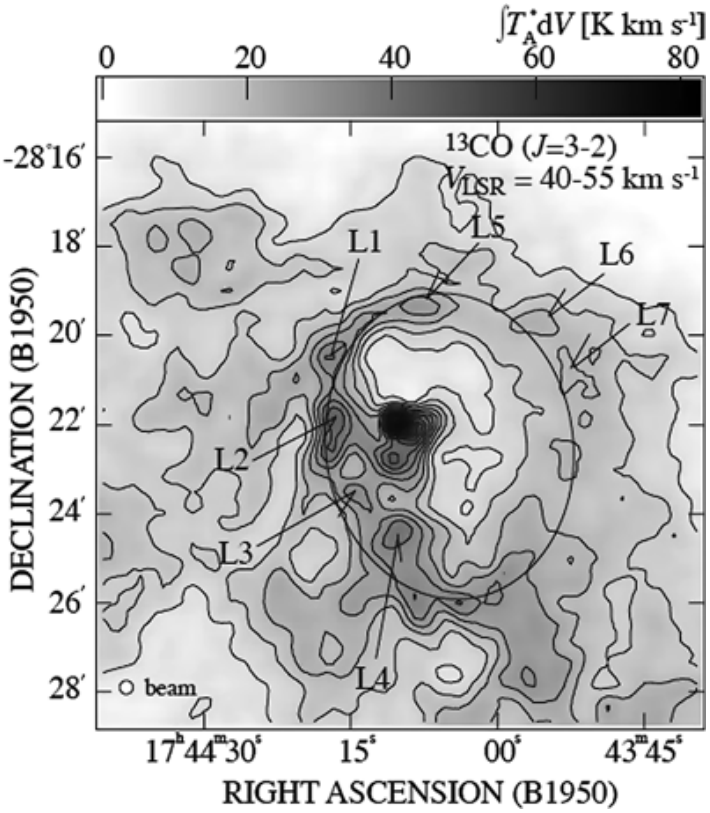


Figure 1.15 – ^{13}CO (J=3→2) observations of the 'hole' and brightened rim in Sgr B2 (Hasegawa *et al.* 2008). The contours and greyscale both show emission integrated over the range 40 km s^{-1} to 55 km s^{-1}. 'Lumps' are indicated with labels L1 to L7.

Ziurys, Friberg & Irvine (1989) studied SiO emission from a number of star formation regions. They observed an apparent activation energy, suggesting the production of SiO involves a gas phase reaction, and not grain destruction. Furthermore they suggest velocities of approximately 40 km s^{-1} are required to destroy SiO grains, and that this would be expected to release other volatiles which are not observed. Instead they suggest SiO is an indicator of high temperature chemistry. Emission from SiO might therefore be more local and due to feedback from massive star formation rather than a cloud collision.

However, for Sgr B2 the shock velocities in a cloud collision may be similar to or even greater than 40 km s^{-1}, so it is possible that partial grain destruction is taking place. Only a small percentage of grains need to be destroyed to noticeably increase the SiO abundance (Mikami *et al.* 1992).

Hüttemeister *et al.* (1995) observed SiO emission from Sgr B2, suggesting the presence of strong shocks. Martin-Pintado *et al.* (1997) also observed SiO emission from Sgr B2. The emission is fragmented into large condensations. The large-scale envelope has a systematic velocity trend from approximately 0 km s^{-1} to 70 km s^{-1}. The broad line profiles and large velocity gradient suggest that the Sgr B2 is undergoing or has undergone energetic events.

This is consistent with the cloud collision model, as the large-scale motions of a cloud collision would create shocks that would create SiO.

They also observed metastable inversion lines of NH_3, and found a hot gas component of approximately 600 K. This gas is at velocities of approximately $65\,\mathrm{km\,s^{-1}}$, which is similar to the gas observed in absorption of ^{13}CO by Sato *et al.* (2000). The hot gas is at low density and so is not heated in the hot molecular cores. Instead it is in more diffuse gas, away from obvious sources of heating. The hot NH_3 gas may have been liberated from dust grains in shocks in a similar way to that suggested for SiO. This is consistent with the hot gas being collected in the shock-compressed layer of the cloud–cloud collision.

Minh *et al.* (1998) observed line emission from the species HCO_2^+ and HNCO. Figure 1.13 shows the emission from HNCO in panel (*d*). The emission is concentrated at the north end of the Clump and Hole in a region labelled the $2'$ N cloud. This lies to the north of the Sgr B2(N) clump. The mass of the cloud is estimated as $10^5\,M_\odot$. Minh *et al.* find that the abundances of HCO_2^+ and HNCO in this cloud are enhanced by at least an order of magnitude compared to the main molecular cloud in Sgr B2. They suggest that, since CO_2 is rare in the gas phase but is common on grains, CO_2 is first liberated from grains by shocks, and then reacts with H_3^+ to form HCO_2^+. It is also possible that HNCO forms in a similar process after OCN^- is liberated from dust grains.

The emission from HNCO has significant red and blue shifted components of approximately $\pm 30\,\mathrm{km\,s^{-1}}$ affecting a mass of approximately $10^4\,M_\odot$. Since the energy requirements of an outflow are unreasonably high, Minh *et al.* (1998) propose the components show material undergoing gravitational collapse.

Sato *et al.* re-examine this in terms of their cloud collision model. In this model the abundances are indeed enhanced by shocks liberating molecules from dust grains. The high velocity components are then simply due to the large velocity differences of the colliding Clump and the gas surrounding the Hole. These observations are therefore consistent with a cloud collision.

Other observations have reported enhanced abundances in this region. These could be due to shock-enhanced chemistry but could also be due to higher gas densities in Sgr B2 compared with molecular clouds outside the Galactic centre. Kuan & Snyder (1996) found that HNCO was pervasive throughout Sgr B2, and matched the eastern outline of the Hole. This is the region containing the clumps Sgr B2(N), (M) and (S), together with most of the masers and HII regions. Mehringer (1995) studied a molecular outflow using CS observations. Using standard abundances, they derived a mass of the outflow that was excessively large, and so concluded they the abundance of CS in the outflow was probably higher than typical values.

Polarization has been observed in far-infrared and sub-mm observations of Sgr B2 (Dowell 1997; Dowell *et al.* 1998; Novak *et al.* 1997). Panel (*c*) of Figure 1.13 shows

polarization measurements of the Sgr B2 region. This polarization is due to the orientation of dust grains, and reflects the underlying magnetic field. The polarization can be due to emission from the oriented grains or, as likely in Sgr B2, due to absorption of a background continuum by the grains. The envelope contains a large-scale, well-ordered and relatively strong magnetic field that does not appear to contain small-scale variations (Novak *et al.* 1997). However the central regions contains a clearly disordered magnetic field that deviates from that on larger scales.

The apparent distortion of the magnetic field near to the star-forming regions may be due to the cloud collision. The region is unusual in that polarization strength correlates with optical depth (Dowell 1997). If the collision rapidly collects gas with a uniform magnetic field into a dense layer, then (if magnetic field lines are frozen into the gas) the magnetic field strength may be increased. Alternatively the collapsing gas of star formation, and the feedback from newly-formed stars, may be responsible for this local disruption.

Although there is no conclusive evidence to show that the Clump–Hole collision model is correct, it is consistent with the currently available observations, and provides a framework that explains many other observations of the region. Sgr B2 is therefore a possible candidate for star formation triggered by cloud collision.

1.2.4 Barnard 68

On smaller scales Burkert & Alves (2009) invoke a core–core collision to resolve some apparently contradictory observations of the dense core Barnard 68. B68 is often described as a classical Bok globule (Bok & Reilly 1947). It is relatively isolated and surrounded by warm gas, but forms part of the Pipe nebula complex of approximately 200 objects. Figure 1.16 shows an extinction map of B68, together with a larger optical image showing other nearby dense cores.

Observations suggest it has a Bonnor–Ebert density profile (Alves, Lada & Lada 2001a). It has a chemical profile which suggests it has experienced a long period remaining quasi-static; it also has a complex velocity field, which has been interpreted as evidence for stable oscillations (Redman, Keto & Rawlings 2006). Both imply a relatively long lifetime, possibly several million years or more.

However, the cloud is clearly non-spherical, whereas a stable Bonnor–Ebert sphere would be expected to reach a roughly spherical configuration within a few periods of oscillation. Furthermore, the best fit to the surface density profile for the cloud suggests a dimensionless Bonnor–Ebert parameter ξ of 6.9 ± 0.2. While this is technically in equilibrium, all values of ξ greater than 6.451 are unstable to small perturbations and thus should collapse (Bonnor 1956).

Burkert & Alves (2009) propose that the observed extension to the south-east is actually a 'bullet' of approximately $0.2\,M_\odot$, which is 10% of the mass of the total system.

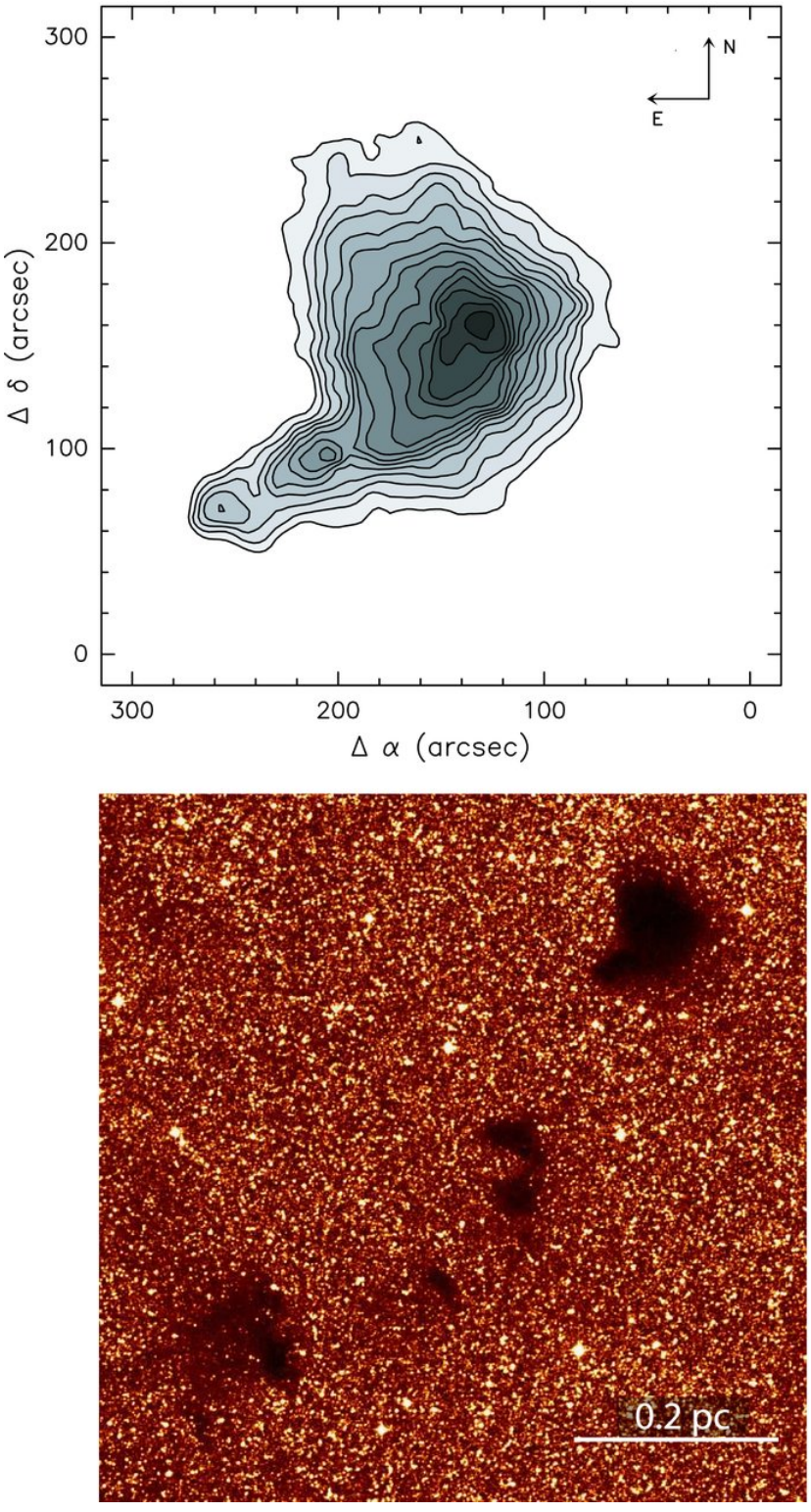


Figure 1.16 – Upper plot: Dust column density map of B68, constructed by measuring extinction to approximately 1000 background stars (Alves, Lada & Lada 2001b). Contours start at and are separated by 2 magnitudes of optical extinction. Lower plot: the surroundings of B68; an optical image from the Digitized Sky Survey.

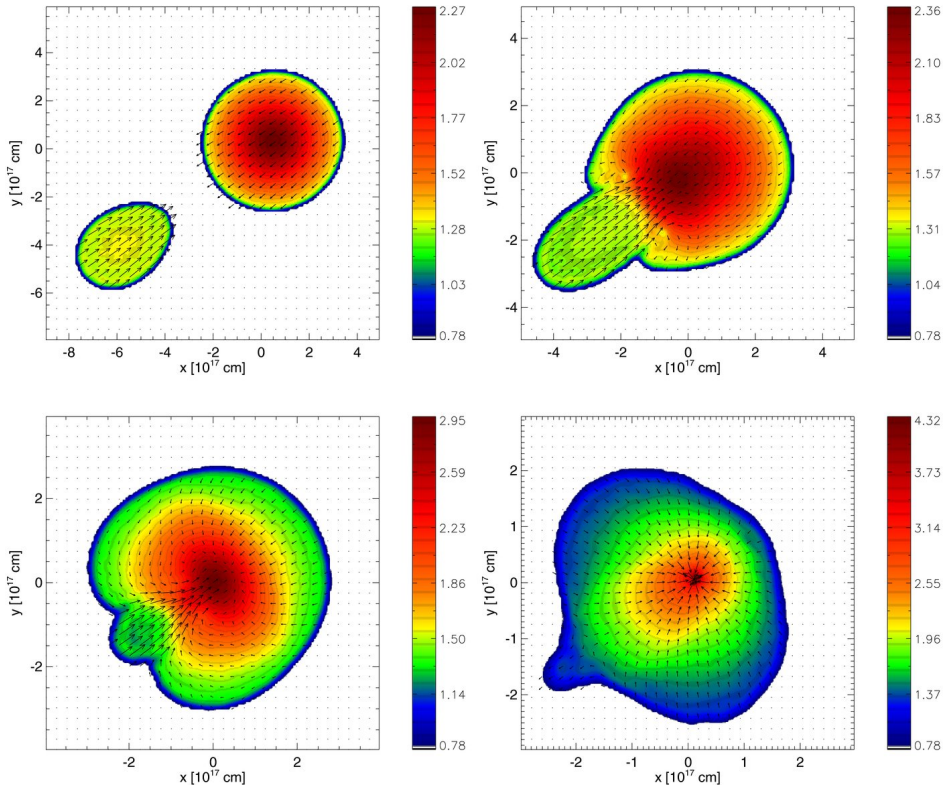


Figure 1.17 – Density cross-sections, at different times, of the SPH simulation of the B68 system. Colours show the logarithm of gas density in units of $1.2 \times 10^{-21}\,\mathrm{g\,cm^{-3}}$. Arrows indicate the velocity field. Times of the snapshots from top left to bottom right with respect to the top right snapshot, which represents the current state of B68, are $-0.9\,\mathrm{Myr}$, $0\,\mathrm{Myr}$, $0.25\,\mathrm{Myr}$ and $0.4\,\mathrm{Myr}$.

To demonstrate this, they perform an SPH simulation using 40 000 SPH particles for the $2.1\,\mathrm{M_\odot}$ main cloud and 4000 particles for the $0.21\,\mathrm{M_\odot}$ bullet. The simulation includes hydrodynamics, self-gravity and an external pressure to account for the warm intercloud medium.

Both the main core and bullet are stable Bonnor–Ebert spheres at the start of the simulation. The dimensionless parameter ξ for the main cloud is set to 6, and the external pressure set accordingly. This fixes ξ for the bullet at 1.16. The main core and bullet are placed at rest at a distance of 0.3 pc from each other. They are drawn together by their mutual gravitational attraction, and collide 1.7 Myr after the start of the simulation at a supersonic velocity of $0.37\,\mathrm{km\,s^{-1}}$.

Figure 1.17 shows the results of the simulation. During the first stages of the collision, where the bullet first makes contact with the main core, the system appears very similar to that observed. Figure 1.18 shows the surface density profile of the observations and the simulation. Both the simulation and the observations are a good fit for a dimensionless parameter of approximately 7, which is unstable. Following the collision, the combined

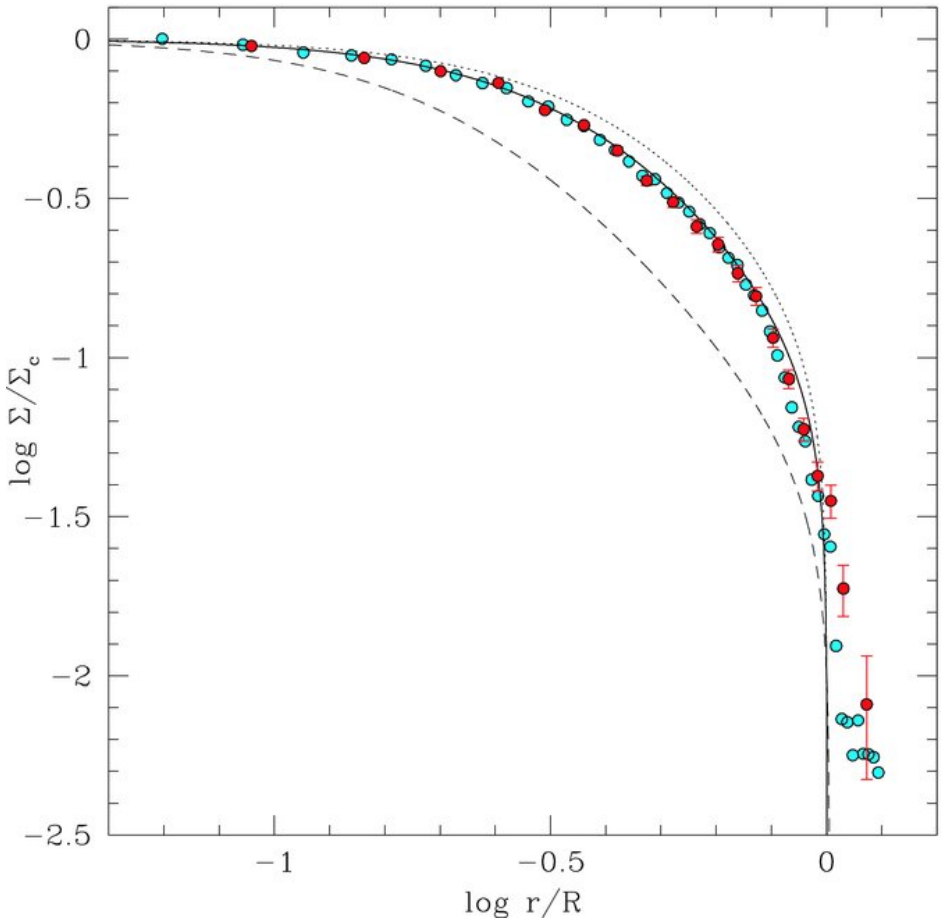


Figure 1.18 – Surface density profile $\Sigma(r)$ for B68. Cyan points are taken from the simulation at the onset of merging. Red points with error bars are taken from near-IR dust extinction observations of B68. The dotted, solid and dashed lines show the profile for theoretical Bonnor–Ebert spheres with dimensionless parameters ξ of 6, 7 and 10 respectively. The effect of the infalling bullet can be seen at large radii.

cloud collapses on a short timescale.

Burkert & Alves (2009) suggest that an observed coherent infall of approximately $120\,\mathrm{m\,s^{-1}}$ in the south-east regions of B68 is evidence for the start of the collapse of B68. They propose that B68 will collapse gravitationally within 10^5 yr, and that within another 10^5 yr a new protostar will have formed, eventually becoming an isolated star like the Sun.

1.3 Previous numerical work

Cloud–cloud collisions have been studied extensively through numerical simulations. Early authors were interested in calculating a mass spectrum of clouds built by collisions and coalescence of smaller clouds, as suggested by Oort & Spitzer (1955). Due to limited computational power, early simulations were limited to studying the most basic properties

of collisions, such as whether clouds were merged or disrupted. The gravitational stability of clouds following a collision was also explored. Some simulations followed the motion of clouds in a galactic disc as they collided, merged and fragmented, and thus were able to determine a mass spectrum.

More recent authors have been able to use higher resolutions to study cloud–cloud collisions. These simulations can show the details of collisions, such as the effect of inhomogeneities in the clouds and instabilities triggered during the collision. They can also follow the formation of protostars from dense gas.

It is now recognized that cloud–cloud collisions are probably rarer than previously thought, and collisions between spherical isolated clouds do not represent the dominant mode of star formation in the Galaxy. However, turbulence on large scales can be modelled as the collision of atomic gas flows. These collisions may be the major mechanism by which molecular gas clouds are produced, with critical importance for the Galactic star formation rate. The collision creates dense gas which can cool efficiently, allowing the rapid formation of molecular hydrogen.

1.3.1 The first cloud–cloud collision simulations

The first simulations of supersonic cloud–cloud collisions were performed by Stone (1970a,b). He simulated both one-dimensional collisions and a two-dimensional collision. Since the initial conditions were assumed to be anti-symmetric, only one side of the collision was simulated, with a solid boundary at the origin. In the one-dimensional case, the cloud was assumed to be infinite in the directions perpendicular to the collision, and finite in length along the collision axis. In the two-dimensional case the clouds were assumed to be spherical and initially touching.

The equations of motion were solved numerically using a Runge-Kutta method for a small number of Lagrangian points in the one-dimensional case, and Lagrangian circular discs in the two-dimensional case. The one-dimensional simulations included optically-thin cooling, artificial viscosity, magnetic fields, initial inhomogeneities and self-gravity. The two-dimensional simulation included optically-thin cooling, artificial viscosity and self-gravity, but did not include magnetic fields or initial inhomogeneities. Molecular cooling from H_2 was included in both cases, but the clouds were assumed to be primarily atomic.

The one-dimensional simulations (Stone 1970b) collided gas flows of length 17 pc and density $2.8 \times 10^{-23}\,\mathrm{g\,cm^{-3}}$ with an equilibrium temperature of 25 K. The collision velocity was $\pm 10\,\mathrm{km\,s^{-1}}$ relative to the centre-of-mass frame. When using artificial viscosity and optically-thin cooling, but not including self-gravity, magnetic fields or initial inhomogeneities, the layer reached a density approximately 2000 times the original density due to the efficient cooling. Without cooling the layer rapidly heat to approximately 6000 K, and so only reached a density approximately four times greater

than the background density. The inclusion of self-gravity had only a very minor effect on the initial collision, but after the collision it slowed and then stopped the expansion of the coalesced cloud. Including initial inhomogeneities produced slightly larger differences than self-gravity, but the overall result of the collision was not significantly altered. Including a magnetic field strongly inhibited the compression of the layer. As the gas is compressed, the magnetic pressure in the layer increased dramatically, and the layer reached a density only 25 times higher than the original density.

The two-dimensional collision (Stone 1970a) collided homogeneous spheres of mass $350\,\mathrm{M}_\odot$, radius $7\,\mathrm{pc}$ and density $1.65 \times 10^{-23}\,\mathrm{g\,cm^{-3}}$. The collision velocity is again $\pm 10\,\mathrm{km\,s^{-1}}$. Each sphere, although not in equilibrium, should be gravitationally stable; however, a merger between the two spheres would produce a sphere that would be gravitationally unstable. The clouds are compressed into a dense slab, which later re-expands. Approximately 25% of the original mass of the clouds is ejected from the collision, which together with the re-expansion of the cloud following the collision means that the coalesced cloud is probably gravitationally stable.

Despite the extremely low resolution of the simulations, Stone was able to produce some accurate results. He showed that colliding clouds would form a dense, shock-compressed layer with a high density contrast, and that cooling was both highly efficient and significant to the development of the layer. He showed that the layer would not expand significantly perpendicular to the layer during the duration of the original collision. He showed that the layer is never in approximate hydrostatic equilibrium and would not be stable; it would either collapse under gravity or re-expand, with some fraction of the mass being ejected. During re-expansion, he showed that the surface of the layer should be Rayleigh–Taylor unstable.

Smith (1980) used a finite difference code to make very similar simulations to the one-dimensional simulations of Stone (1970b). He also considered the collision of a plane-parallel slab, using mirror symmetry to reduce the computational cost. However more physical effects were included; as well as hydrodynamics, self-gravity and optical thin heating and cooling, Smith included thermal conductivity and the formation and destruction of molecular hydrogen.

He was therefore able to simulate the collision of both atomic and molecular clouds. The collision of molecular clouds was approximately isothermal, while the collision of atomic clouds led to a thermal instability. This thermal instability allowed the rapid formation of molecular hydrogen. The collision of atomic clouds thus led to the creation of a diffuse molecular clouds.

Hausman (1981) studied the collision of stable isothermal spheres of masses between 150 and $8000\,\mathrm{M}_\odot$. These spheres represented atomic gas clouds embedded in the warm, low-density ISM. He used a finite particle method with 200 equal mass particles (Hausman

1979; Larson 1978). The particles were treated as individual portions of gas, and as such experienced a pressure force from their nearest neighbours. Self-gravity was included, together with artificial viscosity and optically-thin radiative heating and cooling. External pressure from the warm ISM was also included, as most of the clouds were not gravitationally bound.

The density of clouds was approximately $20\,\mathrm{cm}^{-3}$, and the clouds had an initial temperature of 80 K. Hausman (1981) found that clouds were not typically in pressure balance with the warm ISM. The pressure in a cloud during a collision would rise rapidly, but following the collision would fall back to or below the external pressure of the ISM. The temperature of the clouds did not change much; the rapid cooling ensured that the collisions remained approximately isothermal.

Although he was able to simulate off-centre and unequal mass collisions with this method, particles tended to heavily interpenetrate in shocks. Hausman found that clouds were unlikely to merge, except for very low-velocity collisions or very unequal mass ratios. They would either form a collapsing object or dissipate into the general ambient medium.

Later, Gilden (1984) used a two-dimensional grid code with cylindrical coordinates (Gilden 1982) to simulate the collision of molecular clumps. Hydrodynamics, self-gravity and an optically-thin radiative cooling and heating was included. A constant density was used at all points in the initial conditions; clumps were differentiated from the background by their uniform collision velocity. Unlike the clumps, the background was not self-gravitating.

He collided a single clump with a reflecting boundary and therefore only simulated head-on collisions of equal mass clouds. All of the simulated clumps either collapsed gravitationally or dispersed, with no clumps forming a stable hydrostatically supported clump. The collisions increased the density in the clumps, while the efficient cooling made the collisions approximately isothermal, causing the Jeans mass to be lower in the combination of the two clumps during the collision. He found that the clumps collapsed gravitationally if the combined mass of the two clumps was greater than the Jeans mass of the combination during the collision.

He also simulated the ‘plowing’ collisions of a small clump into an ambient medium, representing the collision of a small clump with a much larger cloud. The small clump ablated material into the surrounding medium as it slowed, and he found that gravitational instability was much rarer in these collisions. The small clump would coalesce with the larger cloud if the column density of the large cloud through the cross-sectional area of the small clump was larger than the column density of the small clump, and if the extra momentum deposited by the small clump did not disrupt the larger cloud.

Hunter *et al.* (1986) used a two-dimensional grid code with a cylindrical coordinate system to simulate the supersonic collision of streams of dusty hydrogen gas. The method

included hydrodynamics, self-gravity, radiative cooling, followed both dust and gas, and accounted for the ionization fraction of the gas. The simulations were limited to less than 5000 computational zones, due to limited computational resources.

The gas in the simulations had a density of either $10^{-21}\,\mathrm{g\,cm^{-3}}$ or $2 \times 10^{-21}\,\mathrm{g\,cm^{-3}}$, and a temperature of 10 K. Mirror symmetry was not used, unlike earlier simulations, allowing instabilities to form in the dense layer formed by the collision. Such an instability did appear and excited strong bending modes in the layer. Hunter *et al.* (1986) attributed this to a Rayleigh–Taylor type instability, but it was more likely due to the non-linear thin shell instability first described by Vishniac (1994) and described in Chapter 7.

Sabano & Tosa (1985) performed a one-dimensional simulation of a head-on supersonic collision between clumps at low Mach number. As with previous one-dimensional calculations, they used mirror symmetry to represent the second clump. The initial density of the clumps was $100\,\mathrm{cm^{-3}}$, and the initial temperature 25 K. They included hydrodynamics, self-gravity and optically thin cooling and heating. Similarly to previous workers, they found the collision produces a dense shock-compressed layer which undergoes rapid cooling, making the collision approximately isothermal.

They also modelled the effect of a magnetic field. A suitable magnetic field strongly hindered the compression of the layer, which would make it less amenable to gravitational collapse. The layer was found to be supported primarily by magnetic pressure. Magnetic fields may therefore prevent or suppress star formation in cloud collisions.

1.3.2 Early SPH simulations of cloud–cloud collisions

Following the development of smoothed particle hydrodynamics (Gingold & Monaghan 1977; Lucy 1977), many workers used it to simulate cloud–cloud collisions. Early simulations were limited to very low numbers of particles, but as better computational resources became available the resolution of simulations improved. These simulations were able to resolve the global evolution of the clouds, but were not able to resolve the formation of protostars within them.

One advantage of these SPH simulations was that they were inherently three-dimensional and were not tied to a particular coordinate system or boundary conditions. They could therefore resolve a full range of instabilities both along the collision axis and across it. SPH simulations have not usually included full radiative cooling or chemistry, instead assuming an isothermal or barotropic equation of state. However, earlier work shows that cloud–cloud collisions are usually close to isothermal for a range of conditions, and so such simplifications are justified.

Lattanzio *et al.* (1985) used an SPH code (Monaghan & Lattanzio 1985) to perform 3D simulations of a comprehensive range of head-on and off-centre cloud collisions. Hydrodynamics and self-gravity were included; the gas was assumed to be isothermal. The external ISM was modelled as an external pressure. The use of artificial viscosity

(Lattanzio *et al.* 1986) in the SPH method prevented particle interpenetration in shocks, solving a major problem of the similar simulations of Hausman (1981). They were able to use a few thousand particles in these simulations.

They used uniform density atomic clouds of density $\sim 30\,\mathrm{cm}^{-3}$, temperature $\sim 80\,\mathrm{K}$, and a range of cloud masses from a few hundred solar masses up to $8000\,\mathrm{M}_{\odot}$. They collide clouds in mass ratios of 1, 2.5, 5 and 10. They found a wide range of cloud morphologies following collisions. The less massive and initially expanding clouds tended to merge into a single expanding cloud. The more massive and initially collapsing clouds collapsed more rapidly following the collision than if left unperturbed. Initially marginally stable intermediate mass clouds merged into one gravitationally unstable cloud following collisions.

In general they found that the collision did not lead to the formation of multiple clouds, even in off-centre collisions. Instead, clouds that were gravitationally unstable or only marginally unstable would coalesce and collapse, while gravitationally stable and unconfined clouds would expand as a single diffuse cloud following the collision. Off-centre collisions produced a large variety of cloud forms which were not static, but continued to evolve following the collision.

Nagasawa & Miyama (1987) used three-dimensional SPH simulations to simulate the collision of spherical clouds which were initially in hydrostatic equilibrium. Some collisions resulted in the coalescence of the two clouds to form a single stable cloud. This new cloud underwent prolate and oblate oscillations following the collision until it reached a new hydrostatic equilibrium. Other collisions resulted in the central regions of the cloud becoming gravitationally unstable and collapsing.

Habe & Ohta (1992) used two-dimensional axisymmetric SPH to simulate head-on collisions of non-identical clouds. The simulations were isothermal, and self-gravity was included. They used a mass ratio of 1:4 between a small, dense cloud and a larger, less dense cloud. These clouds were either both hydrostatically supported clouds confined by an external pressure, or uniform density clouds. All of their clouds were initially stable to gravitational collapse, and they used approximately 4300 SPH particles per simulations.

They found that the large cloud was disrupted by the bow shock of the small cloud. The small cloud was compressed, and material ablated away from it. For higher cloud masses or lower collision velocities, the small cloud became gravitationally unstable and collapsed. This occurred even when the small cloud contained less than a Jeans mass in its unperturbed state.

1.3.3 Star-forming SPH simulations of cloud–cloud collisions

Later SPH simulations of cloud–cloud collision had sufficient resolution to begin modelling the formation of protostars. The earliest simulations were limited to the formation of just a few systems. In other areas of study, the introduction of sink particles (Bate, Bonnell

& Price 1995), described in Section 3.7, raised this limit considerably by replacing the dense protostars with a single sink particle which accreted inflowing gas. However, this technique appears to have been neglected in simulations of cloud–cloud collisions until recently.

A number of workers have studied the formation of binary and multiple star systems following the off-centre collision of equal-mass clumps, including Chapman *et al.* (1992), Turner *et al.* (1995) and Whitworth *et al.* (1995b). These workers all used the same computational method, which is described below.

Chapman *et al.* (1992) perform three-dimensional simulations of off-centre supersonic cloud–cloud collisions. They include self-gravity and an equation of state which includes implicit radiative cooling. Their spherical clouds contain $75\,\mathrm{M}_\odot$, but they only use 2000 particles per cloud and so are only able to resolve protostars of mass greater than approximately $2\,\mathrm{M}_\odot$. The cloud radius is 2×10^5 au, and they collide with an impact parameter of 8×10^4 au and a relative collision velocity of $1.6\,\mathrm{km\,s^{-1}}$.

They collide both stable spheres in hydrostatic balance, confined by an external pressure, and gravitationally stable but transient uniform density clouds. The results are similar; the transient clouds produce a protobinary system while the hydrostatic clouds produce a protoquadruple system. They find that the formation of multiple systems is a feature of highly compressive radiative shocks. All of their protostars are massive, which is probably due to the comparatively low resolution preventing the formation of smaller systems.

Turner *et al.* (1995) continued the work of Chapman *et al.* (1992), using the same methods to simulate the collision of hydrostatically-balanced clouds of mass $225\,\mathrm{M}_\odot$, radius 1.94 pc and with individual collision velocities of $\pm 1.7\,\mathrm{km\,s^{-1}}$. They used 6272 particles per clump, together with 42 088 particles to simulate the hot low-density ISM and 16 194 fixed boundary particles. The clouds collide at an impact parameter of 1.55 pc.

They find that the shock-compressed dense layer fragments into filaments, and the filaments fragment into beads. They produce more than ten protostellar discs in the layer, which interact to create multiple star systems through dissipative tidal and viscous interactions. This only produces long period binary systems with orbital periods of greater than approximately 3000 yr. As with the simulations of Chapman *et al.* (1992), they are only able to resolve the formation of massive stars.

Whitworth *et al.* (1995b) provide a final set of simulations using the same methods as Chapman *et al.* (1992) and Turner *et al.* (1995). They simulate two collisions of hydrostatic spherical clouds. Each cloud has a velocity of $\pm 1.7\,\mathrm{km\,s^{-1}}$.

The first simulation collides identical clouds of mass $750\,\mathrm{M}_\odot$ and radius 6.45 pc, with an impact parameter of 7.73 pc. They use 15 408 particles for each clump, 101 988 particles to simulate a hot, low-density background, and 23 080 boundary particles. The

shear in the collision causes rapid accretion onto protostellar discs. These then form a bar which becomes unstable and fragments, forming a massive binary system with a separation of approximately 8000 au.

The second simulation collides identical clouds of mass $75\,M_\odot$ and radius 0.446 pc, with an impact parameter of 0.268 pc. They use 1736 particles for each clump, 5492 particles to simulate a hot, low-density background, and 5072 boundary particles. The shear in the collision is less violent than for the first simulation, and accretion onto protostellar discs is slower. These discs form a system of spiral arms, which subsequently fragment, creating an unstable Trapezium-like system with four massive stars as average distances of approximately 7000 au.

Byleveld & Pongracic (1996) perform a three-dimensional simulation of colliding Jeans-stable isothermal spheres of mass $100\,M_\odot$. These spheres, simulated with approximately 2000 particles each, are embedded in a warm intercloud medium of approximately 8000 particles. Self-gravity is also included.

They include the effects of magnethydrodynamics, using the prescription presented by Monaghan (1992). They collide the clouds with and without an initial magnetic field and compare the results. After a time of 1.2 Myr, the layer in the unmagnetized collision reaches a density 10^6 times greater than the original density, while the layer in the magnetized collision only reaches a density of 10^3 times greater. This agrees with the work of Sabano & Tosa (1985), who found a magnetic field reduced the compression of the layer. However, in the magnetized collision the layer takes only another 0.1 Myr to reach a density contrast of 10^6, so the effect on the collision appears to be small.

Bhattal *et al.* (1998) returned to the collision of $75\,M_\odot$ isothermal spheres, using a different three-dimensional SPH code to Chapman *et al.* (1992) and Whitworth *et al.* (1995b), but using similar initial conditions. Bhattal *et al.* use a more complicated equation of state which is isothermal at 100 K below a density of $1.2 \times 10^{-20}\,\mathrm{g\,cm^{-3}}$, and then reduces smoothly to an isothermal temperature of 10 K at higher densities. At a density of $10^{-14}\,\mathrm{g\,cm^{-3}}$, the equation of state becomes adiabatic with an index of 1.4, and the temperature begins to rise with increasing density. This equation of state reflects warm gas at low densities, efficient radiative cooling at intermediate densities, and adiabatic heating at high densities when the cloud become optically thick.

Each clump is represented by 2000 particles, with additional simulations carried out at 8000 particles per clump to examine the effect of numerical resolution. The clouds are embedded in between 16 000 and 24 000 non-self-gravitating inter-cloud particles to provide the external pressure that confines the isothermal clouds. These are contained within approximately 14 000 boundary particles.

Clouds are collided at a range of impact parameters from 0.1 to 0.7, in terms of the cloud radius. At low impact parameters the results are similar to those found by

Whitworth *et al.* (1995b). A central condensed object forms which rotationally fragments due to accretion. At lower impact parameters the primary protostellar object becomes more massive than at higher impact parameters. Multiple systems tend to have less equal mass ratios at these impact parameters.

At higher impact parameters, two equal mass protostellar objects are formed at different positions in the colliding clouds. These protostellar objects do not interact during the simulated time. At lower impact parameters these objects are bound, while they are unbound at higher impact parameters. However, Bhattal *et al.* (1998) point out that these objects are still accreting, and the objects may become bound at later times.

The higher resolution simulation shows that the width of the shock is overestimated in the low resolution simulations, as expected. The masses of protostellar objects are also overestimated, and multiple systems are less tightly bound than in the higher resolution simulations due to the smoothing of gravity and pressure forces in the lower resolution simulations.

Marinho & Lépine (2000) also used SPH, simulating the collision of 2222.5 $M_\odot$ clouds at individual cloud velocities of $\pm 5\,\mathrm{km\,s^{-1}}$. They include self-gravity, optically thin cooling and heating, and a basic chemistry network including the formation and destruction of molecular hydrogen. They simulated head-on and off-centre collisions. Each clump was simulated using 4096 SPH particles.

They initially set up their clouds with a density profile of r^{-1} and a cut-off radius of 10 pc. They found that temperatures in the shock initially rose to a few thousand Kelvin. However cooling was rapid, and the gas quickly returned to 10 K to 20 K, forming cold dense clumps. These clumps reached densities of $10^4\,\mathrm{cm^{-3}}$.

They found that head-on collisions were highly dissipative. Due to the higher densities reached in head-on collisions, they tender to be cooler than off-centre collisions despite head-on collisions being more dissipative of kinetic energy.

Kitsionas & Whitworth (2007) used SPH to simulate collisions between pairs of 10 $M_\odot$ and 75 $M_\odot$ isothermal spheres, extending on the work of Pongracic *et al.* (1992) and Bhattal *et al.* (1998). They used particle splitting (Kitsionas 2000) to avoid violating the Jeans condition (Bate & Burkert 1997; Hubber, Goodwin & Whitworth 2006) and so could resolve gas at a density of $10^{-12}\,\mathrm{g\,cm^{-3}}$. For one simulations, they used sink particles (Bate, Bonnell & Price 1995) to continue the simulation beyond the creation of the first collapsing object, and so allow protostellar statistics to be obtained.

The 10 $M_\odot$ clumps were simulated with 15 000 particles, which increased as required due to particle splitting. Similarly, the 75 $M_\odot$ clumps began with 110 000 particles. They use a similar equation of state to Bhattal *et al.* (1998), except that for simulations with 10 $M_\odot$ clumps the temperature of low-density gas is only 35 K instead of 100 K for the 75 $M_\odot$ clumps. For their lower-mass clumps, they explored the effect of increasing collision

velocity and impact parameter. The higher-mass clumps were included for comparison with the results of Bhattal *et al.* (1998).

The collisions produced a dense shock-compressed layer as expected. This layer fragmented into a network of filaments, and protostars formed within these filaments. These filaments were not resolved in the simulations of Bhattal *et al.* (1998) due to the low number of particles used, the fixed and relatively high gravitational smoothing length, and the failure to resolve the Jeans condition at all times. Kitsionas & Whitworth (2007) therefore argue that the results of Bhattal *et al.* (1998) are not reliable at the later stages of their simulations.

They find that the total mass in the shock-compressed layer decreases with increasing impact parameter, becoming negligible above an impact parameter of 0.5 in terms of the clump radius. Increasing impact parameter therefore decreases the number of protostellar objects produced and the final star formation efficiency. High impact parameters also tend to create two large independent and unbound protostars, similar to that seen in Bhattal *et al.* (1998).

Increasing the collision velocity has the reverse effect, and increases the star formation efficiency. The layer becomes denser, which decreases the fragmentation scale of the layer. The collision therefore produces more filaments and more protostars. However, a combination of high Mach number and high impact parameter destroys the layer through excessive shear before fragmentation can occur.

Increasing the clump mass increases the area of the layer, and therefore creates a more complex network of filaments. Each filament tends to fragment into 4 to 6 stars for the higher-mass clouds compared with only one or two stars per filament for the lower-mass clouds. The overall star formation efficiency of the higher-mass collisions is estimated at 10–15%, whereas the lower-mass collisions have an estimated star formation efficiency of 20–30%. The star formation efficiency therefore appears to decrease with increasing clump mass.

One simulation was continued using sink particles. A number of further protostars were formed, resulting a total of nine sink particles. Some of these formed multiple systems, while others were ejected at a few $\mathrm{km\,s^{-1}}$. Accretion rates of protostars in all simulations were in the range $10^{-5}\,\mathrm{M_\odot\,yr^{-1}}$ to $5\times10^{-5}\,\mathrm{M_\odot\,yr^{-1}}$, and the accretion phase tended to last for $10^4\,\mathrm{yr}$ to $3\times10^4\,\mathrm{yr}$. The rates of accretion did not depend on the parameters of the collision, such as the collision velocity or the impact parameter.

Duarte-Cabral *et al.* (2011) used SPH to model the formation of the Serpens Main cluster. Serpens Main is composed of two main compact protoclusters along a 0.6 pc filament. Both protoclusters are embedded in approximately $30\,\mathrm{M_\odot}$ of gas each, and are believed to be both extremely young and of similar ages, suggesting a common triggering event. However the two clusters have complex and dissimilar velocity structures.

Emission surrounding the NW cluster has only a single velocity component, and a single temperature of 10 K. All the protostars appear to be at a similar evolutionary stage. By contrast, the emission surrounding the SE cluster has a more complex velocity structure with two components, and a more varied temperature distribution between 10 K and 20 K. There is also evidence of on-going star formation in this cluster.

They use an SPH code derived from that of Benz *et al.* (1990), which has been modified to include sink particles. An isothermal equation of state is used for simplicity. The collision is modelled as the collision of two gravitationally stable cylinders of approximate length 1 pc at a 55° angle to each other. External pressure is used to keep the cylinders stable, which are at the isothermal temperature of 30 K and have masses of 30 $M_\odot$ and 45 $M_\odot$. Although this temperature is higher than what is observed, at lower temperatures the cylinder would not be gravitationally stable. Standard length cylinders are modelled with 250 000 SPH particles each.

Duarte-Cabral *et al.* (2011) explored the effect of on-centre and off-centre collisions, extended cylinders, and a turbulent velocity field with the observed average velocity. Simulated column density observations are produced of the system at the time of first sink formation and compared to the real observations. The SE cluster is modelled as a direct result of the collision of the two cylinders, while the formation of the NW cluster is modelled as the more quiescent collapse of one of the cylinders away from the point of collision. The off-centre collisions are found to best match the observations.

A model with extended cylinders and no turbulence produces both the SE and NW clusters, but the addition of any significant turbulence prevents the formation of the NW cluster. The best match to the observations, excluding the formation of the NW cluster, is an off-centre collision with turbulence and with standard length cylinders.

1.3.4 Collisions of clumpy clouds

There has been very little work exploring the effects of initial inhomogeneities on cloud–cloud collision. Kimura & Tosa (1996) create clumpy clouds by imposing substructured random velocities on the clouds. They create a giant molecular cloud (GMC) of radius 25 pc and average density 100 cm^{-3}, together with three small molecular clouds (SMC) of radii 2.5, 5 and 12.5 pc and average density 1000, 500 and 200 cm^{-3} respectively. They use a two-dimensional grid code to simulate a GMC–GMC collision, and the collisions of a GMC with the three SMC. They compare collisions of uniform density clouds with collisions of clumpy clouds. The clouds are surrounded by a low-density high-temperature medium.

They do not simulate a uniform density GMC–GMC collision, but for uniform density GMC–SMC collisions they find their results are similar to those of Habe & Ohta (1992). Even a small SMC creates a bow shock that sweeps through the GMC, collecting material and significantly disrupting the GMC.

The collision of two clumpy GMCs leads to a shock compressed layer, as expected, but this layer is strongly perturbed and distorted by the internal structure of the clouds. The layer fragments into several massive clumps, and eventually disperses.

The collision of a small clumpy SMC with a clumpy GMC is significantly different to that for uniform clouds. The SMC is compressed to an average density 5 to 6 times greater than before the collision. No significant bow shock forms during the collision; instead the SMC penetrates the GMC. As it moves through the GMC it interacts with the clumps in the GMC, which are of a similar size to the compressed SMC. The SMC breaks up into fragments; the largest of these fragments emerges from the far side of the GMC. This fragment has more mass than the original SMC, implying it has accreted material from the GMC during the collision.

A larger clumpy SMC is unable to penetrate a clumpy GMC. Instead, a large shocked region forms due to the larger cross-section of the SMC. This region is larger than the clumps in the GMC, but it is perturbed and distorted by them. Eventually, the layer fragments into a number of dense clumps which do not escape the GMC.

The dense clumps formed in these simulations are typically much larger than the Jeans mass. Such collisions could therefore trigger star formation in clumpy GMCs. Kimura & Tosa (1996) demonstrate that the behaviour of clumpy clouds in collisions is significantly different to that for more homogeneous clouds.

1.3.5 Colliding atomic flows

There have been few recent simulations of cloud–cloud collision, as current thinking suggests that molecular clouds are more dynamic and less long-lived that previously thought. This would tend to reduce the frequency and significance of cloud–cloud collisions. Instead clouds are continually being formed and destroyed through large-scale turbulent flows of atomic gas driven on large scales. The collision of such flows can rapidly produce molecular hydrogen, and have been studied in a range of simulations (Heitsch, Hartmann & Burkert 2008; Hennebelle *et al.* 2008; Vázquez-Semadeni *et al.* 2007b, 2006) using both adaptive mesh refinement (AMR) and SPH codes.

These simulations have included realistic cooling functions to follow the thermodynamics of the warm atomic medium as it is compressed. They find the shock-compressed layer becomes turbulent and expands, while the compressed atomic gas becomes thermally unstable and rapidly cools. This cold gas then gravitationally collapses to high densities. Although chemistry was not simulated, the gas would then be expected to rapidly form molecular hydrogen. However, the addition of magnetic fields strongly suppresses this fragmentation and collapse (Vázquez-Semadeni *et al.* 2007a).

Chapter 2

Smoothed Particle Hydrodynamics

Stars form through the gravitational collapse of interstellar gas. Simulating star formation therefore inevitably involves hydrodynamics. The equations of hydrodynamics can be solved analytically for only a few simple problems, and so a number of methods have been devised to solve them numerically. One method is to define fluid properties on a grid and calculate the flow of gas between grid cells (an Eulerian approach). These *grid* methods have a long history and are commonly used in a wide range of applications, including star formation.

Simulating star formation provides a number of unusual difficulties that can be challenging for grid methods. Flows in astrophysics are often highly compressible and can be either subsonic or supersonic. Shocks are therefore common, as is turbulence. These effects, together with gravity, lead to gas densities from less than $10^{-22}\,\mathrm{g\,cm^{-3}}$ to greater than $10^{1}\,\mathrm{g\,cm^{-3}}$. The simulations in this thesis cover approximately ten orders of magnitude in density. This enormous range of densities corresponds to a large range of physical length scales. Grid codes require complex adaptive-mesh refinement techniques to deal with this range of scales without excessive computational cost.

Smoothed particle hydrodynamics is a Lagrangian *particle* method for calculating hydrodynamics, first introduced by Lucy (1977) and Gingold & Monaghan (1977), which naturally contains this adaptivity (see reviews by Monaghan 1992 and Price 2012). The continuous distribution of fluid properties is replaced by interpolation onto a set of particles. Each particle has a fixed mass and a smoothing kernel which overlaps its neighbouring particles. The properties of the fluid, such as density, velocity and temperature, are interpolated onto the particles. Each particle does not represent a single sampling point in the fluid, but a weighted contribution over its smoothing kernel.

SPH is a conceptually simple yet efficient method which is Galilean invariant and has no preferred axes. As particles move with the flow of the fluid, resolution is naturally concentrated in the regions of highest density. It is usually simple to add additional physics, and there is an extensive range of well-developed particle-gravity techniques used in n-body codes which can easily be adapted to an SPH code. There is now a wide

variety of SPH codes used in star formation, other areas of physics and industry.

2.1 SPH quantities

The basic equation of SPH, describing the connection between particles and the continuous fluid, is the SPH interpolant (Monaghan 1992)

$$A(\boldsymbol{r}) \approx \sum_{b=1}^{N} m_b \frac{A_b}{\rho_b} W\big(|\boldsymbol{r} - \boldsymbol{r}_b|, h\big) \ . \tag{2.1}$$

At some interpolation point $\boldsymbol{r}$, a quantity of the continuous fluid $A(\boldsymbol{r})$ can be calculated by summing over N neighbouring particles. We use b to refer to the indices of neighbouring particles; m_b and ρ_b are the mass and density of each neighbouring particle. The function W is the *kernel function*, which depends only on (*i*) the distance between the interpolation point and the neighbouring particle and (*ii*) a *smoothing length*, h . The purpose of the kernel function is to weight the contributions of neighbouring particles, typically so that nearby particles contribute more than distant ones. The exact form of the kernel function is not critical, and is discussed in Section 2.1.1, while the choice of smoothing length is discussed in Section 2.1.2.

This relation is valid at all points in space, but for most purposes SPH quantities are evaluated only at the positions of particles. Therefore for any particle a, the SPH value of the quantity A is given by

$$A_a \approx \sum_{b=1}^{N} m_b \frac{A_b}{\rho_b} W_{ab} \ , \tag{2.2}$$

where $\boldsymbol{r}_{ab} = \boldsymbol{r}_b - \boldsymbol{r}_a$, and $W_{ab} \equiv W\big(|\boldsymbol{r}_a - \boldsymbol{r}_b|, h\big)$.

We can define the first spatial derivative as simply ∇A, but for increased accuracy we use

$$\nabla A \approx \frac{\nabla(\rho A) - A\nabla\rho}{\rho} \ . \tag{2.3}$$

As described in Monaghan (1992) and Price (2004), this removes the first error term which is present even for constant functions. Substituting in equation 2.2 as appropriate yields

$$\rho_a \nabla A_a \approx \nabla \left(\sum_{b=1}^{N} m_b \frac{A_b \rho_b}{\rho_b} W_{ab} \right) - A_a \nabla \left(\sum_{b=1}^{N} m_b \frac{\rho_b}{\rho_b} W_{ab} \right) \ , \tag{2.4}$$

which simplifies to

$$\nabla A_a \approx \sum_{b=1}^{N} m_b \frac{A_b - A_a}{\rho_a} \nabla W_{ab} \ , \tag{2.5}$$

where $\nabla_a W_{ab}$ indicates that the gradient of the kernel is calculated at particle a.

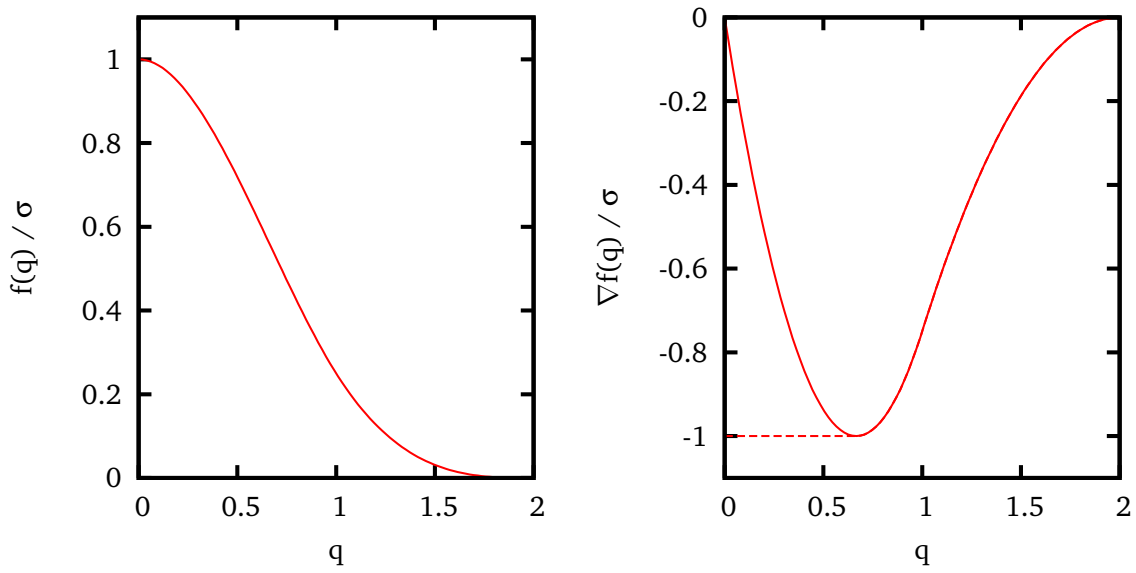


Figure 2.1 – M_4 cubic spline function (left) and gradient (right). The dashed line shows the modification of Thomas & Couchman (1992).

We can also define the SPH approximation for the dot product of a vector using

$$\nabla \cdot \boldsymbol{u} \approx \frac{\nabla \cdot (\rho \boldsymbol{u}) - \boldsymbol{u} \cdot \nabla \rho}{\rho} \,. \tag{2.6}$$

Substituting in equations 2.2 and 2.5 yields

$$\rho_a \nabla \cdot \boldsymbol{u}_a \approx \nabla \left(\sum_{b=1}^{N} m_b \frac{\boldsymbol{u}_b \rho_b}{\rho_b} W_{ab} \right) - \boldsymbol{u}_a \cdot \nabla \left(\sum_{b=1}^{N} m_b \frac{\rho_b}{\rho_b} W_{ab} \right) \,, \tag{2.7}$$

which simplifies to

$$\nabla \cdot \boldsymbol{u}_a \approx \sum_{b=1}^{N} m_b \frac{\boldsymbol{u}_b - \boldsymbol{u}_a}{\rho_a} \cdot \nabla W_{ab} \,. \tag{2.8}$$

Although equation 2.1 depends on particle density, this dependence disappears when density is the desired quantity, as in

$$\rho_a \approx \sum_{b=1}^{N} m_b W_{ab} \,, \tag{2.9}$$

which allows densities to be calculated knowing only the mass and distribution of particles, and the choice of smoothing length used in the smoothing kernel function.

2.1.1 Kernel functions

The smoothing kernel function defines the relative contributions of near and distant particles. As it represents a smoothing of physical properties onto the particles, it must

reduce to a Dirac delta function as the smoothing length tends to zero, so that

$$\lim_{h\to 0} W(|\boldsymbol{r}-\boldsymbol{r}_b|,h) = \delta(|\boldsymbol{r}-\boldsymbol{r}_b|) \ . \tag{2.10}$$

It is normalized so that

$$\int W(|\boldsymbol{r}-\boldsymbol{r}_b|,h)\,\mathrm{d}^3 r = 1 \ . \tag{2.11}$$

Although it is possible to use an infinite kernel, such as a Gaussian kernel, this is computationally inefficient since this implies all particles have to be considered in every calculation. Instead, kernels with compact support are typically used, where the value of the kernel function is zero outside some multiple of the smoothing length.

For the simulations in this thesis, we use the common M_4 cubic spline kernel as shown in Figure 2.1 (Monaghan & Lattanzio 1985). The *extent* of this kernel is twice the smoothing length, or $2h$. We define

$$W(r,h) = \frac{1}{h^D} f(q) \tag{2.12}$$

where D is the number of spatial dimensions, and $q = r/h$. The cubic spline kernel is then

$$f(q) = \sigma \begin{cases} 1 - {}^3\!/_2 q^2 + {}^3\!/_4 q^3 & 0 \le q < 1 \\ {}^1\!/_4 (2-q)^3 & 1 \le q < 2 \\ 0 & q \ge 2 \end{cases} , \tag{2.13}$$

where σ is

$$\sigma = \begin{cases} {}^2\!/_3 & D = 1 \\ {}^{10}\!/_{7\pi} & D = 2 \\ {}^1\!/_\pi & D = 3 \end{cases} . \tag{2.14}$$

We define the kernel gradient, used for calculating spatial derivatives, as

$$\nabla W(r,h) = \frac{1}{h^{D+1}} \frac{\mathrm{d}f}{\mathrm{d}q}\, \widehat{\boldsymbol{r}} \ , \tag{2.15}$$

where $\widehat{\boldsymbol{r}} = \dfrac{\boldsymbol{r}}{|\boldsymbol{r}|}$, and we refer to $\dfrac{\mathrm{d}f}{\mathrm{d}q}$ as the *kernel derivative*.

The kernel derivative is modified from the analytic value as suggested by Thomas & Couchman (1992). Particle forces are calculated using the first derivative of the kernel, as explained in Section 2.2. If the kernel is not modified, then as particle separations decrease the kernel gradient reduces to zero, and the repulsive force between the particles disappears. This leads to the tensile or clumping instability in SPH, where particles tend

to clump together in extremely tight groups. To prevent this, the kernel derivative is set to a constant for values of q below $2/3$. The modified kernel derivative is

$$\frac{\mathrm{d}f(q)}{\mathrm{d}q} = -\sigma \begin{cases} 1 & 0 \geq q < 2/3 \\ 3q - 9/4\, q^2 & 2/3 \geq q < 1 \\ 3/4\,(2-q)^2 & 1 \geq q < 2 \\ 0 & q \geq 2 \end{cases} . \tag{2.16}$$

This modification has the disadvantage, described by Price (2012), that the kernel gradient is no longer correctly normalised. The modified region is oversampled and if there are particles in this region calculated quantities will be incorrect. The derivation from a Lagrangian described in Section 2.2 assumes that the kernel derivative is the analytic derivative of the kernel function, and breaking this assumption breaks the exact conservation of energy.

It is sometimes useful to have symmetrized equations which involve the kernel. For this purpose we define an averaged kernel

$$\overline{W}_{ab} = \frac{W\big(|\boldsymbol{r} - \boldsymbol{r}_b|, h_a\big) + W\big(|\boldsymbol{r} - \boldsymbol{r}_b|, h_b\big)}{2} , \tag{2.17}$$

and a corresponding averaged kernel gradient

$$\nabla_a \overline{W}_{ab} = \frac{\nabla_a W\big(|\boldsymbol{r} - \boldsymbol{r}_b|, h_a\big) + \nabla_a W\big(|\boldsymbol{r} - \boldsymbol{r}_b|, h_b\big)}{2} . \tag{2.18}$$

2.1.2 Smoothing lengths

The choice of smoothing length is of great importance in an SPH simulation. In the first SPH simulations all particles shared a single fixed smoothing length, of the order of the mean particle separation. This works for relatively uniform simulations where the density does not change much, but for star forming simulations it is essential to allow the smoothing length to vary.

There are a number of methods used to determine smoothing lengths. Although smoothing lengths can be defined at all points in space, typically they are calculated only at the positions of the particles as this is where SPH quantities are needed. A simple method is to set a smoothing length so that a certain number of *neighbours* falls within the extent of the smoothing kernel. Many simulations have used a value of $50 \pm N_{\mathrm{TOL}}$ neighbours for 3D simulations. To avoid excessive numerical diffusion, N_{TOL} should always be set to zero, as noted by Nelson & Papaloizou (1994) and Attwood, Goodwin & Whitworth (2007).

We instead use the method suggested by Monaghan (2002) and set smoothing lengths according to the relation

$$h_a = \eta \left(\frac{m_a}{\rho_a} \right)^{1/D} , \tag{2.19}$$

where η is a constant. This maintains the consistency of smoothing lengths and densities. Since the density as calculated in equation 2.9 depends on the smoothing length, these two equations must be solved iteratively. We use $\eta = 1.2$ as suggested by Price (2004).

2.2 Hydrodynamics

2.2.1 Continuity equation

As described in Section 2.1, particle densities in SPH are given by

$$\rho_a = \sum_{b=1}^{N} m_b W_{ab} . \tag{2.20}$$

We wish to take the time derivative of the particle density, which will be a Lagrangian time derivative as SPH particles move with the fluid flow. We first need to calculate the time derivative of the kernel function. From equation 2.12 we note that the kernel is given by

$$W_{ab} = W(\boldsymbol{r}_a - \boldsymbol{r}_b, h) . \tag{2.21}$$

For now we assume that the smoothing length is time independent, giving the time derivative of the kernel as

$$\frac{\mathrm{d}W_{ab}}{\mathrm{d}t} = \frac{\mathrm{d}W_{ab}}{\mathrm{d}\boldsymbol{r}_{ba}} \frac{\mathrm{d}\boldsymbol{r}_{ba}}{\mathrm{d}t} . \tag{2.22}$$

This simplifies to

$$\frac{\mathrm{d}W_{ab}}{\mathrm{d}t} = \nabla_a W_{ab} \boldsymbol{v}_{ba} , \tag{2.23}$$

giving us the derivative of density with respect to time,

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \sum_{b=1}^{N} m_b \boldsymbol{v}_{ba} \nabla_a W_{ab} . \tag{2.24}$$

Expanding this we can obtain

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \boldsymbol{v}_a \cdot \sum_{b=1}^{N} \frac{m_b}{\rho_b} \rho_b \nabla_a W_{ab} - \sum_{b=1}^{N} \frac{m_b}{\rho_b} \rho_b \boldsymbol{v}_b \cdot \nabla_a W_{ab} , \tag{2.25}$$

and by comparison to equation 2.8, this is equivalent to

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \boldsymbol{v} \cdot \nabla\rho - \nabla \cdot (\rho \boldsymbol{v}) . \tag{2.26}$$

Using the identity

$$\nabla \cdot (\phi \boldsymbol{A}) = \boldsymbol{A} \cdot \nabla \phi + \phi \nabla \cdot \boldsymbol{A} \,, \tag{2.27}$$

equation 2.26 reduces to

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = -\rho \nabla \cdot \boldsymbol{v} \,, \tag{2.28}$$

which is the Lagrangian continuity equation (Price 2012). This demonstrates that the SPH density equation is the SPH version of the continuity equation.

2.2.2 Lagrangrian for hydrodynamics

As described by Monaghan (2002), the equations of motion for SPH can be derived from the Lagrangian for hydrodynamics (Eckart 1960),

$$\mathcal{L} = \int \left(\frac{1}{2} \rho \boldsymbol{v}^2 - \rho u \right) \mathrm{d}V \,. \tag{2.29}$$

In SPH form, replacing the mass element $\rho \, \mathrm{d}V$ with the particle mass m_b, this becomes

$$\mathcal{L} = \sum_{b=1}^{N} m_b \left(\frac{1}{2} \boldsymbol{v}_b^2 - u_b(\rho_b, s_b) \right) \,, \tag{2.30}$$

where u_b is the internal energy per unit mass and s_b is the particle entropy. The Euler–Lagrange equation is

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{v}_a} \right) - \frac{\partial \mathcal{L}}{\partial \boldsymbol{r}_a} = 0 \,, \tag{2.31}$$

and its components in terms of SPH quantities are

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \mathcal{L}}{\partial \boldsymbol{v}_a} = m_a \frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} \,, \tag{2.32}$$

and

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{r}_a} = -\sum_{b=1}^{N} m_b \left. \frac{\partial u_b}{\partial \rho_b} \right|_s \frac{\partial \rho_b}{\partial \boldsymbol{r}_a} \,, \tag{2.33}$$

assuming the entropy is constant, and noting that the local density is a function of particle position. The first law of thermodynamics gives us

$$\left. \frac{\partial u}{\partial \rho} \right|_s = \frac{P}{\rho^2} \,, \tag{2.34}$$

and taking the spatial derivative of equation 2.9 at particle b with respect to particle a, we obtain

$$\frac{\partial \rho_b}{\partial \boldsymbol{r}_a} = \sum_{c=1}^{N} m_c \frac{\partial}{\partial \boldsymbol{r}_a} W(|\boldsymbol{r}_b - \boldsymbol{r}_c|, h) \,. \tag{2.35}$$

The spatial derivative of the kernel is zero if a is equal to neither b nor c, and is otherwise equal to either ∇W_{bc} or $-\nabla W_{bc}$ for $a = b$ and $a = c$ respectively. We therefore have

$$\frac{\partial \rho_b}{\partial \boldsymbol{r}_a} = \sum_{c=1}^{N} m_c \frac{\partial W_{bc}}{\partial \boldsymbol{r}_a} \left(\delta_{ab} - \delta_{ac}\right) \,, \tag{2.36}$$

where δ_{ab} is a Kronecker delta. Together with equations 2.33 and 2.34 we have

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{r}_a} = -\sum_{b=1}^{N} m_b \frac{P_b}{\rho_b^2} \sum_{c=1}^{N} m_c \nabla_a W_{bc} \left(\delta_{ab} - \delta_{ac}\right) \,. \tag{2.37}$$

This is zero except when $a = b$ or $a = c$ giving two components

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{r}_a} = \begin{cases} -m_a \dfrac{P_a}{\rho_a^2} \displaystyle\sum_{c=1}^{N} m_c \nabla_a W_{ac} & \text{when } a = b \\ \displaystyle\sum_{b=1}^{N} m_b \frac{P_b}{\rho_b^2} m_a \nabla_a W_{ba} & \text{when } a = c \end{cases} \tag{2.38}$$

The kernel function is anti-symmetric such that $\nabla_a W_{ab} = -\nabla_a W_{ba}$. By rearranging and reusing the index b instead of c for the $a = b$ case we obtain

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{r}_a} = -m_a \sum_{b=1}^{N} m_b \left(\frac{P_a}{\rho_a^2} + \frac{P_b}{\rho_b^2}\right) \nabla_a W_{ab} \,. \tag{2.39}$$

Finally equating this with the other half of the Euler–Lagrange equation, equation 2.32, gives the standard SPH equation of motion

$$\frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} = -\sum_{b=1}^{N} m_b \left(\frac{P_a}{\rho_a^2} + \frac{P_b}{\rho_b^2}\right) \nabla_a W_{ab} \,. \tag{2.40}$$

As this is derived from the Euler–Lagrange equation, in the absence of dissipation and with constant smoothing lengths SPH will conserve energy to the accuracy of the integration scheme. This form gives explicitly symmetrical forces between particles and therefore conserves momentum exactly. A similar derivation can be used to show that angular momentum is also conserved (Price 2004).

2.2.3 Energy equation

The first law of thermodynamics in the absence of dissipation gives

$$\frac{\mathrm{d}u_a}{\mathrm{d}t} = \frac{P_a}{\rho_a^2}\frac{\mathrm{d}\rho_a}{\mathrm{d}t} \,. \tag{2.41}$$

Inserting the Lagrangian time derivative of particle density, as given in equation 2.24, we obtain the energy equation in SPH in terms of the specific internal energy (Price 2004)

$$\frac{\mathrm{d}u_a}{\mathrm{d}t} = \frac{P_a}{\rho_a^2}\sum_{b=1}^{N} m_b(\boldsymbol{v}_a - \boldsymbol{v}_b)\cdot\nabla_a W_{ab} \,. \tag{2.42}$$

Other formulations are possible using the total energy or the specific entropy.

2.2.4 Additional terms

The equations of SPH given above have been consistently derived from variational principles, as shown by Monaghan (2002), Springel & Hernquist (2002) and Price (2004, 2012). The density summation (equation 2.9) is derived from only the continuity equation and the SPH interpolant as shown in Section 2.2.1. The equation of motion (equation 2.40) is similarly derived, including only the SPH interpolant, the Euler–Lagrange equation (equation 2.31) and the first law of thermodynamics (equation 2.34). Finally the energy equation (equation 2.42) is derived from only the SPH interpolant and the first law of thermodynamics. For these reasons in the absence of dissipation and with constant smoothing lengths SPH conserves, to the accuracy of the integration scheme, energy, momentum and angular momentum.

However, this derivation ignores the variation of smoothing lengths. As first noted by Nelson & Papaloizou (1994), additional terms (referred to as ∇h terms) should be added to the momentum equation to account for the dependence of the kernel function on smoothing lengths. We define a new shorthand:

$$W_{ab,h=a} \equiv W\big(|\boldsymbol{r}_{ab}|, h_a\big) \,. \tag{2.43}$$

The SPH density summation is then

$$\rho_a \approx \sum_{b=1}^{N} m_b W_{ab,h=a} \,. \tag{2.44}$$

Taking the spatial derivative of density at particle b with respect to particle a,

equation 2.36 should be replaced by (Monaghan 2002; Price 2012)

$$\frac{\partial \rho_b}{\partial \boldsymbol{r}_a} = \frac{1}{\Omega_b} \sum_{c=1}^{N} m_c \frac{\partial W_{bc,h=b}}{\partial \boldsymbol{r}_a} \left(\delta_{ab} - \delta_{ac}\right) , \tag{2.45}$$

where Ω is a term which accounts for the dependence of the smoothing kernel on smoothing length, and is given by (Monaghan 2002)

$$\Omega_a = 1 - \frac{\partial h_a}{\partial \rho_a} \sum_{b=1}^{N} m_b \frac{\partial W_{ab,h=a}}{\partial h_b} . \tag{2.46}$$

Since, as described in Section 2.1.2, we have defined smoothing lengths to have a simple relationship to density,

$$h_a = \eta \left(\frac{m_a}{\rho_a}\right)^{1/D} , \tag{2.47}$$

we have

$$\frac{\partial h_a}{\partial \rho_a} = \frac{-\eta}{D\,\rho_a} \left(\frac{m_a}{\rho_a}\right)^{1/D} = -\frac{h_a}{D\,\rho_a} . \tag{2.48}$$

Finally, as shown by Price (2012), we can replace the SPH momentum equation as given in equation 2.40 with

$$\frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} = -\sum_{b=1}^{N} m_b \left(\frac{P_a}{\Omega_a \rho_a^2} \nabla_a W_{ab,h=a} + \frac{P_b}{\Omega_b \rho_b^2} \nabla_a W_{ab,h=b} \right) . \tag{2.49}$$

Similarly, the energy equation becomes

$$\frac{\mathrm{d}u_a}{\mathrm{d}t} = \frac{P_a}{\Omega_a \rho_a^2} \sum_{b=1}^{N} m_b (\boldsymbol{v}_a - \boldsymbol{v}_b) \cdot \nabla_a W_{ab,h=a} . \tag{2.50}$$

2.3 Dissipation

As cautioned by Monaghan (1997) and Price (2008), most formulations of SPH include implicit assumptions that quantities such as velocity and internal energy are differentiable. For example the use of Euler–Lagrange equations (equation 2.31) in the derivation of the equations of motion implies differentiability. The outcome is that SPH is not accurate at sharp discontinuities, such as in shocks. The usual compromise to mitigate this is to apply dissipation to each quantity that may be discontinuous, smoothing the discontinuities over a few smoothing lengths and allowing SPH to resolve them. Although all conservative quantities in SPH may require a similar treatment, we use only artificial viscosity.

2.3.1 Artificial viscosity

We use the method of artificial viscosity described by Monaghan (1997), together with the time-dependent switch of Morris & Monaghan (1997).

The equation of motion, equation 2.49, is modified to include an artificial viscosity term, such that

$$\frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} = -\sum_{b=1}^{N} m_b \left(\frac{P_a}{\Omega_a \rho_a^2} \nabla_a W_{ab,h=a} + \frac{P_b}{\Omega_b \rho_b^2} \nabla_a W_{ab,h=b} + \Pi_{ab} \nabla_a \overline{W}_{ab} \right) . \tag{2.51}$$

This extra term is given by

$$\Pi_{ab} = \begin{cases} -\dfrac{\overline{\kappa}_{ab} v_{\mathrm{SIG}}\, \boldsymbol{v}_{ab} \cdot \boldsymbol{r}_{ab}}{\overline{\rho}_{ab} |\boldsymbol{r}_{ab}|} & \text{if } \boldsymbol{v}_{ab} \cdot \boldsymbol{r}_{ab} \leq 0 \\ 0 & \text{if } \boldsymbol{v}_{ab} \cdot \boldsymbol{r}_{ab} > 0 \end{cases} , \tag{2.52}$$

where v_{SIG} is a signal velocity between particles, $\boldsymbol{v}_{ab} = \boldsymbol{v}_b - \boldsymbol{v}_a$ and $\overline{\rho}_{ab} = {}^1\!/_2(\rho_a + \rho_b)$. $\nabla_a \overline{W}_{ab}$ is the average kernel gradient as defined in equation 2.18. The parameter $\overline{\kappa}_{ab}$ is of value ~ 1, and can be either equal and constant for all particles or time-varying as described in Section 2.3.2. We define the signal velocity as proposed by Monaghan (1997) but with an additional β parameter, similar to that used by Price (2004), such that

$$v_{\mathrm{SIG}} = c_a + c_b - \beta \frac{\boldsymbol{v}_{ab} \cdot \boldsymbol{r}_{ab}}{|\boldsymbol{r}_{ab}|} , \tag{2.53}$$

where c_a and c_b are the sound speeds at particles a and b respectively.

We need to include heating from this artificial viscosity into the energy equation. Equation 2.50 is replaced by

$$\frac{\mathrm{d}u_a}{\mathrm{d}t} = \sum_{b=1}^{N} m_b \left(\frac{P_a}{\Omega_a \rho_a^2} \boldsymbol{v}_{ab} \cdot \nabla_a W_{ab,h=a} + \Lambda_{ab}\, \widehat{\boldsymbol{r}_{ab}} \cdot \nabla_a \overline{W}_{ab} \right) , \tag{2.54}$$

where $\widehat{\boldsymbol{r}_{ab}} = \dfrac{\boldsymbol{r}_{ab}}{|\boldsymbol{r}_{ab}|}$, and

$$\Lambda_{ab} = \begin{cases} -\dfrac{\overline{\kappa}_{ab} v_{\mathrm{SIG}} (\boldsymbol{v}_{ab} \cdot \boldsymbol{r}_{ab})^2}{2\overline{\rho}_{ab} |\boldsymbol{r}_{ab}|^2} & \text{if } \boldsymbol{v}_{ab} \cdot \boldsymbol{r}_{ab} \leq 0 \\ 0 & \text{if } \boldsymbol{v}_{ab} \cdot \boldsymbol{r}_{ab} > 0 \end{cases} . \tag{2.55}$$

2.3.2 Time-varying artificial viscosity

Artificial viscosity is useful in shocks, as it allows SPH to resolve the shock and prevents particle interpenetration, but away from shocks it causes unphysical dissipation. It is

therefore desirable to reduce viscosity where it is not needed. For this reason viscosity is set to zero when $\boldsymbol{v}_{ab} \cdot \boldsymbol{r}_{ab} > 0$ and particles are receding from each other, as in equations 2.52 and 2.55.

We also use the method proposed by Morris & Monaghan (1997). Viscosity is controlled by the parameter $\overline{\kappa}_{ab}$ in equations 2.52 and 2.55, which we define as

$$\overline{\kappa}_{ab} = \frac{\kappa_a + \kappa_b}{2} \,, \tag{2.56}$$

where κ_a and κ_b are the values of the time-varying parameter κ for particle a and b respectively. The time-varying value κ_a is not given directly, but evolved using the equation

$$\frac{\mathrm{d}\kappa_a}{\mathrm{d}t} = \frac{\kappa_{\mathrm{MIN}} - \kappa_a}{\tau_a} + S_a \,, \tag{2.57}$$

where τ_a is a smoothing time, S_a is a source term and κ_{MIN} is the minimum permitted value of κ (typically ~ 0.1). There is a corresponding κ_{MAX} (typically ~ 1). The smoothing time τ_a is defined by

$$\tau_a = \frac{h_a}{\mathscr{C}_1 c_a} \,, \tag{2.58}$$

where the parameter $\mathscr{C}_1 = 0.1$, following Morris & Monaghan (1997). The source term S is then given by

$$S_a = \left(\kappa_{\mathrm{MAX}} - \kappa_a\right) \begin{cases} -(\nabla \cdot \boldsymbol{v}_a) & \text{if } (\nabla \cdot \boldsymbol{v})_a \leq 0 \\ 0 & \text{if } (\nabla \cdot \boldsymbol{v})_a > 0 \end{cases} . \tag{2.59}$$

These quantities are time-evolved with the Euler integration scheme. In strong shocks, κ_a rises rapidly from $\sim \kappa_{\mathrm{MIN}}$ to $\sim \kappa_{\mathrm{MAX}}$, then decays back to $\sim \kappa_{\mathrm{MIN}}$. This ensures viscosity is included only where required.

2.4 Gravity in SPH

SPH particles do not represent point masses. Instead they sample the continuous fluid, and so it is necessary to include gravitational softening. The simplest choice is to use Plummer sphere smoothing with a fixed softening length, but this loses the inherent adaptivity of SPH and does not represent the local density. Gravity can also be smoothed using a kernel function. If this is the same kernel as used for SPH, then it benefits from compact support, which means that outside the kernel gravitational forces reduce to exactly Newton's law for point particles.

2.4.1 Resolution requirements

Gravity in SPH has no fixed spatial resolution limit. Instead, resolution in SPH is mass-limited, as a smoothing kernel always contains an approximately fixed mass. When conducting self-gravitating simulations, the critical mass scale is the Jeans mass. A number of authors have shown it is important to always resolve the Jeans mass in SPH to resolve real fragmentation and prevent artificial fragmentation.

Bate & Burkert (1997) caution against using a gravitational softening length different to the particle smoothing length. They find that when the gravitational softening length is larger than smoothing length, pressure forces tend to dominate. When the gravitational softening length is smaller than the smoothing length, gravitational forces tend to dominate. This can disrupt the equilibrium of a condensation of approximately the Jeans mass, causing a stable clump to collapse or preventing the collapse of an unstable clump. It is therefore critical to use a softening length equal to the smoothing length; however, the use of variable softening lengths can lead to non-conservation of energy.

These results were confirmed by Whitworth (1998) and Hubber, Goodwin & Whitworth (2006). They showed that SPH neither causes artificial fragmentation, nor prevents the collapse of objects with greater than the Jeans mass, as long as the gravitational softening length is set to match the smoothing length, and the Jeans length is always resolved. Bate, Bonnell & Bromm (2003) showed that the Jeans mass was adequately resolved by $1.5\mathcal{N}_{\rm NEIB}$ SPH particles, where $\mathcal{N}_{\rm NEIB}$ is the number of neighbours in a smoothing kernel. The Jeans mass, described in Section 6.1, is given by

$$M_{\rm J} \approx \sqrt{\frac{375}{4\pi G^3 \rho}} c_{\rm s}^3 \, , \tag{2.60}$$

where ρ is the gas density and $c_{\rm s}$ is the sound speed. This should never be smaller than the minimum resolvable mass,

$$M_{\rm MIN} \approx 1.5 \, \mathcal{N}_{\rm NEIB} \, m_{\rm SPH} \, , \tag{2.61}$$

where $m_{\rm SPH}$ is the mass of each SPH particle.

For a given total simulation mass, and without changing the number of neighbours, this gives a minimum required number of particles for a desired minimum resolvable mass. Equivalently, for a given total simulation mass, and a maximum number of particles (typically limited by computational constraints), this places an upper limit on permissible density, although this depends on the equation of state used. When using an isothermal equation of state, the Jeans mass decreases monotonically, giving a simple maximum density. However when using a more realistic equation of state which accounts for the opacity limit, such as the barotropic equation of state used by Bate, Bonnell & Bromm

(2002) and others, the Jeans mass reaches a minimum at a critical density. Above this density, the gas is no longer isothermal and rises in temperature. This rise in temperature raises the Jeans mass. For these equations of state it is sufficient to resolve the Jeans mass at the critical density.

When using a more complicated equation of state, where there no simple relationship between density and temperature, calculating the required resolution is more complicated. This is discussed further in Section 8.3.

2.4.2 A conservative formulation of SPH gravity

To address the non-conservation of energy with simple kernel-smoothed gravity and variable smoothing lengths, Price & Monaghan included the gravitational potential into the Lagrangian described in Section 2.2.2. They derived a new formulation which conserves energy even with variable smoothing lengths. They defined a new gravitational potential smoothing kernel

$$\phi(r,h) = 4\pi \left[-\frac{1}{r}\int_0^r W(r',h)\, r'^2\, \mathrm{d}r' + \int_0^r W(r',h)\, r'\, \mathrm{d}r' - \int_0^{2h} W(r',h)\, r'\, \mathrm{d}r' \right] \quad (2.62)$$

and a gravitational force smoothing kernel,

$$\phi'(r,h) = \frac{4\pi}{r^2}\int_0^r W(r',h)\, r'^2\, \mathrm{d}r' \,. \quad (2.63)$$

We define an averaged smoothing kernel,

$$\overline{\phi}_{ab} = \frac{\phi(|\boldsymbol{r}_{ab}|,h_a) + \phi(|\boldsymbol{r}_{ab}|,h_b)}{2} \,, \quad (2.64)$$

and a quantity ζ given by

$$\zeta_a = \frac{\partial h_a}{\partial \rho_a}\sum_{b=1}^{N} m_b \frac{\partial \phi(|\boldsymbol{r}_{ab}|,h_a)}{\partial h_a} \,. \quad (2.65)$$

Not including artificial viscosity, the equation of motion (equation 2.49) is then replaced with

$$\begin{aligned}\frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} = -\sum_{b=1}^{N}\Bigg[G\overline{\phi}_{ab}\frac{\boldsymbol{r}_a - \boldsymbol{r}_b}{|\boldsymbol{r}_{ab}|} + \frac{G}{2}\left(\frac{\zeta_a}{\Omega_a}\nabla_a W_{ab,h=a} + \frac{\zeta_b}{\Omega_b}\nabla_a W_{ab,h=b}\right) \\ + \left(\frac{P_a}{\rho_a^2\Omega_a}\nabla_a W_{ab,h=a} + \frac{P_b}{\rho_b^2\Omega_b}\nabla_a W_{ab,h=b}\right)\Bigg] \,.\end{aligned} \quad (2.66)$$

The first term represents the standard kernel-smoothed gravity, while the third term is the standard SPH pressure force. The second term is a new, additional force that increases gravity at small distances. It is this new term that corrects the conservation of energy.

Chapter 3

Simulation algorithms

It is not enough to have an accurate method for solving a problem. In order to be useful, it must be practical to apply the method to the problem of interest. In the context of computer simulations, this means that the method must have an associated *algorithm* which produces results without excessive computational work. This usually means that a computer *code* written using the algorithm produces results without using an excessive amount of storage or computer time.

It is therefore not surprising that research in many fields that rely heavily on computational simulations has kept pace with the development of faster computing facilities and improved algorithms. Increases in computing power typically allow problem sizes and accuracy to be increased. The development of a new algorithm can allow research in entirely new fields.

In this chapter, we review some of the algorithms used in the SEREN code. Many of these algorithms date back more than 20 years, but have been continuously improved, expanded and repurposed.

3.1 The SEREN code

We use an MPI-parallelized version of the SPH code SEREN. The standard (non-MPI) version of SEREN is described in detail by Hubber *et al.* (2011) and was written primarily by David Hubber and Chris Batty, with contributions by Thomas Bisbas, Andrew McLeod, Krisada Rawiraswattana, Dimitrios Stamatellos, Stefanie Walch and Anthony Whitworth.

The SEREN code was designed primarily for simulations of star formation in the local universe. It is a computationally efficient code, making use of OpenMP parallelization (Section 4.1.1) to allow greater efficiency on shared-memory machines. We use the following features in the SEREN code:

- Smoothed particle hydrodynamics (Section 2)
- Self-gravity (Section 2.4)
- Solving the energy equation (Section 3.4)

	Hydro tree	Gravity tree
Link to the *next* cell on the same level	✓	✓
Link to the *if-open* cell of this cell	✓	✓
Number of particles (0 for non-leaf cells)	✓	✓
List of links to any particles (leaf cells only)	✓	✓
Centre of mass (COM) of particles ($\boldsymbol{r}_c$)	✓	✓
Maximum distance of particles from COM (r_{MAX})	✓	
Largest particle smoothing length (h_{MAX})	✓	
Total mass of particles in cell		✓
Quadrupole moment terms of cell		✓

Table 3.1 – List of quantities stored in each Barnes–Hut tree. The first two properties define the structure of the tree; the next two relate to particles for leaf cells.

- Barnes–Hut trees (Section 3.2)
- Block timestepping (Section 3.6.2)
- Leapfrog integration scheme (Section 3.6.3)
- Sink particles (Section 3.7)

3.2 Barnes–Hut trees

The SEREN code uses spatial oct-trees extensively, both to find neighbouring particles and as Barnes–Hut gravity trees (Barnes & Hut 1986). We refer to these spatial oct-trees as BH trees. Each *cell* of the tree has up to eight *child* cells, which correspond to the eight spatial octants of a cube. The tree is constructed so that cells are divided until there are less than a chosen number of particles in any cell. By *walking* this tree, a complete list of all particles in a chosen region of space can be found efficiently. The tree structure is also used for tree-gravity, as described in Section 3.5.

3.2.1 Building trees

Only one parameter must be chosen when building a BH tree. This is the maximum number of particles in a cell at the bottom of the tree, which we set to eight.

Building a tree begins with the *root cell*. The cubic root cell is set to encompass all particles in the simulation, and is then divided into its eight octants. Each of the eight octants are considered in turn. A *cell* is created for each octant which contains particles. Any cell with eight or fewer particles is set as a *leaf* cell. These leaf cells contain references to the particles within them. Any cells which contain more than eight particles is set as a *link* cell, and is further subdivided into eight octants. Unlike leaf cells, link cells do not contain the indices of individual particles. The repeated subdivision of space is illustrated in Figure 3.1.

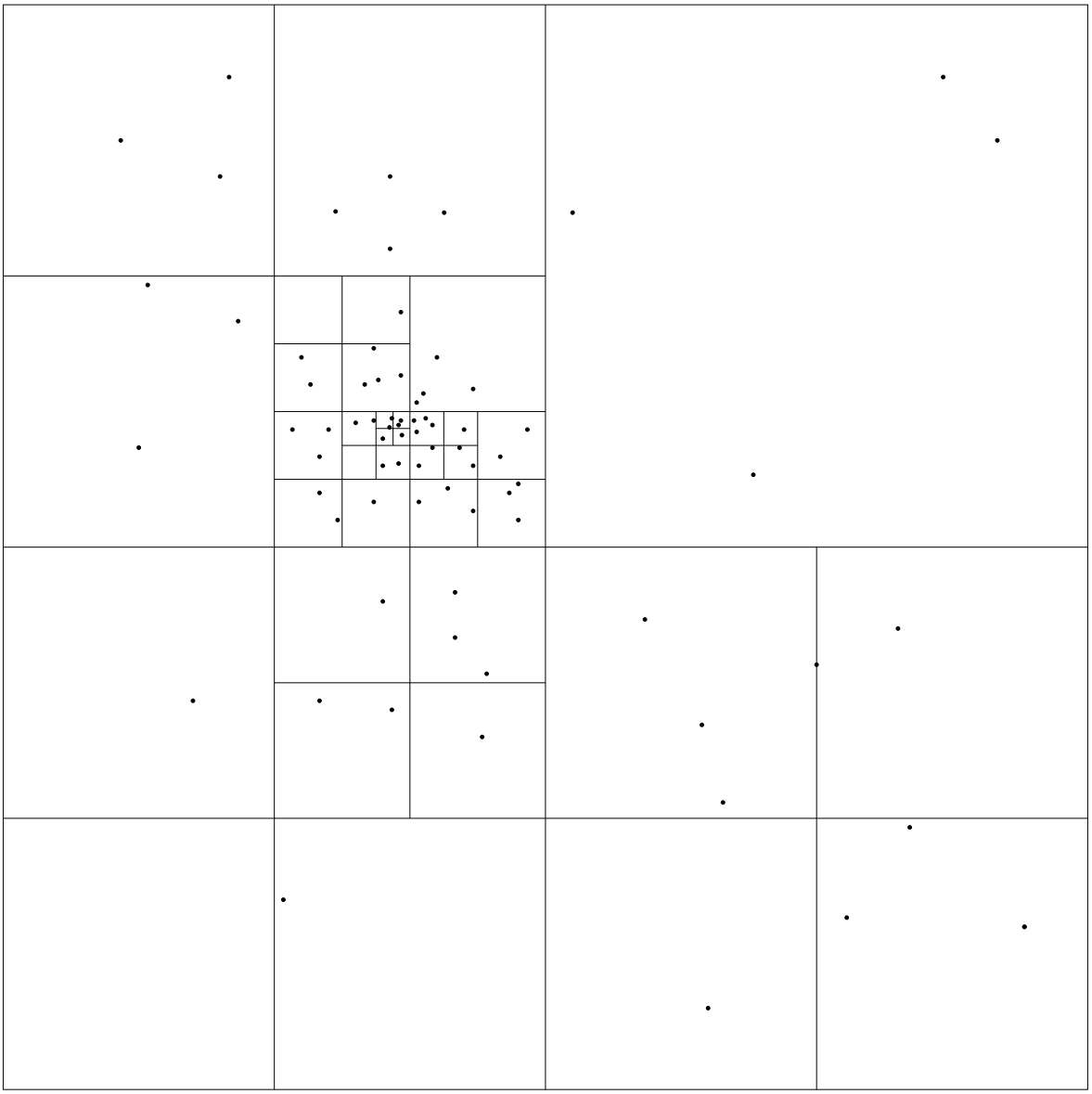


Figure 3.1 – A 2D version of a Barnes–Hut tree. Each cell splits into four child cells until no cell contains more than four particles.

The tree-building then proceeds to the next *level*. Each link cell on this level is considered in turn in the same way as the original root cell. The link cells are divided into octants, a new cell is created for each occupied octant, and any cells with eight or fewer particles is set as a leaf cell. The process is repeated until all cells on the current level are leaf cells, or equivalently all particles have been assigned to a leaf cell. Finally once the complete tree has been constructed, it needs to be stocked (as described in Section 3.2.2).

The BH tree has two sets of links per cell, as shown in Figure 3.2. The first is the *next* cell link. In any set of sibling cells (child cells of the same parent cells), the *next* cell links will point from the first sibling cell to the second sibling cell, and so on until the last sibling cell. The *next* cell link of the last sibling cell points to the *next* cell of its parent cell.

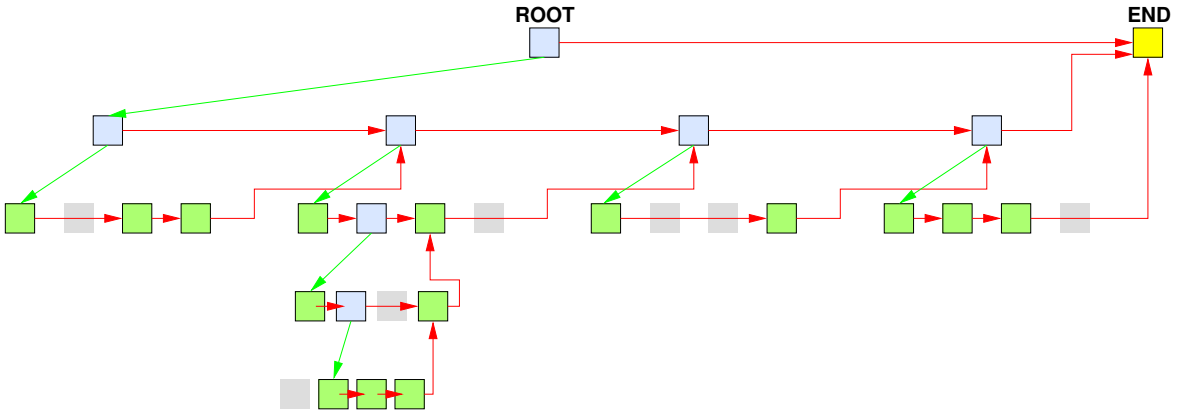


Figure 3.2 – A typical tree structure for a 2D variant of a Barnes–Hut tree. Link cells are shown in blue, leaf cells in green. A greyed-out box is shown where a quadrant contains no particles and thus no cell exists. The *next* cell link is shown with a red arrow, and the *if-open* cell link with a green arrow. The virtual *end cell* is shown in yellow.

The root cell is given a special *next* cell link to a virtual *end cell* to indicate that there is no more tree to follow, and this value is inherited by all cells on each level at the end of the tree. The second set of links is the *if-open* links, which simply point to the cell's first child cell. If the cell is a leaf cell, then it has no child cells and this value is not set or used.

In SEREN, two trees are built: one for hydrodynamics (the 'hydro' tree) and one for gravity. The hydrodynamics tree includes all particles and is used when searching for particles. The gravity tree contains only particles that feel gravity and is used for tree-gravity as described in Section 3.5. Table 3.1 shows the quantities stored in each tree. For our simulations, all particles are self-gravitating, and therefore both trees contain the same particles.

Tree-building is relatively expensive, so the trees are only occasionally rebuilt. As the particles in the simulation move around, the cells of the tree must be expanded and begin to overlap. This reduces the efficiency of a tree walk, as a search for particles in a given region will require opening more cells. The frequency of tree-building is therefore a compromise between the incurred expense and this reduced efficiency.

3.2.2 Stocking trees

Unlike tree-building, which begins at the top of the tree and builds downwards, tree-stocking begins at the very deepest leaf cell and works upwards, one level at a time. First, each leaf cell on the level is updated. For the hydrodynamics tree, this means that the centre of mass of each leaf cell is calculated. The maximum distance of any particle in that cell from the centre of mass (r_{MAX}) and the maximum smoothing length of any particles in the cell (h_{MAX}) is also calculated. For the gravity tree the mass of particles in a cell does not usually change, but the centre of mass and the quadrupole moments described in Section 3.5.2 are updated.

Once each leaf cell has been updated on a level, the link cells are updated. The centre of mass of a link cell is the centre of mass of all its child cells. For each child cell, a distance $s = r + r_{\rm MAX}$ (where r is the distance to the child cell) is calculated. The largest value of s from the child cells is then used as an approximate value for $r_{\rm MAX}$ for the parent cell. The maximum smoothing length is simply the maximum $h_{\rm MAX}$ of the child cells. For the gravity tree the quadrupole moments are calculated as described in Section 3.5.2.

The tree needs to be restocked every time particle information changes. In an SPH simulation this is typically every step. Fortunately tree-stocking is not as computationally expensive as tree-building.

3.2.3 Walking trees

Trees are usually 'walked' to find neighbouring particles within a certain distance of a particle. We define $\boldsymbol{r}_c$ as the position of the current cell c, and $r_{\rm MAX}$ as the maximum distance of particles from the centre of mass for that cell. The following procedure builds a list of particles within some distance $2h$ of a particle p at $\boldsymbol{r}_p$, starting at the root cell.

- Calculate the distance between the particle and the closest possible particle in the cell, defined as $s = |\boldsymbol{r}_p - \boldsymbol{r}_c| - r_{\rm MAX}$.
- If the current cell is a leaf cell:
  - Test whether $s \leq 2h$. If it is, add any particles within $2h$ to the list.
  - Move to the *next* cell.
- If the current cell is a link cell:
  - Test whether $s \leq 2h$. If it is, child cells of this cell may include particles within $2h$. Move to the *if-open* cell.
  - Otherwise, move to the *next* cell.
- Repeat this process for the new cell until we are directed to the *end* cell.

For hydrodynamic forces it is necessary to find not only the particles that fall within $2h$ of a particle, but also particles whose smoothing kernels overlaps the particle. For this, the same procedure is used but substituting $r_{\rm MAX} + 2h_{\rm MAX}$ instead of $r_{\rm MAX}$.

3.3 SPH in SEREN

The basics of SPH are described in detail in Section 2. In order to calculate SPH quantities at each particle, it is necessary to know the particle's smoothing length and density. These must be found simultaneously: SEREN uses a simple fixed-point iteration scheme (Price 2004). If this fails to converge, it falls back to a binary chop scheme.

We begin with the previous value of the smoothing length h. A potential neighbour list containing the positions of all particles within a radius $h_{\rm search} = 1.1 \times 2h$ of the particle is obtained using the tree walk described in Section 3.2.3. The potential neighbours in this list within $2h$ of the particles are then used to calculate a particle density using equation 2.9. Using this density, it is possible to calculate a new estimate of the smoothing length using equation 2.19. This allows a new estimate of the smoothing length. If this increases the smoothing length above $h_{\rm search}$, it is necessary to do another tree walk with a larger $h_{\rm search}$ to return a new list of potential neighbours. The iteration is repeated until smoothing lengths converge to some tolerance value. If smoothing lengths do not converge within a fixed number of iterations, the iteration is abandoned. The smoothing length and density are then found by a binary chop.

Once smoothing lengths and densities are calculated all other SPH quantities, such as velocity divergence, can be calculated. The sound speed, pressure and temperature depend on the equation of state, which can be isothermal, barotropic or use the method described in Section 3.4.

Hydrodynamic forces are relatively straightforward. Equation 2.51 calculates particle accelerations by an SPH sum. In order to symmetrize forces, it is necessary that if a particle includes another particle in its calculation of forces, the reverse must also be true. Therefore for force calculations a particle b is considered a *neighbour* of a particle a if either it falls within $2h_a$ of particle a or particle a falls within $2h_b$ of particle b. The neighbour search, performed by a tree walk, is carried out using the alternative procedure involving $h_{\rm MAX}$ described in Section 3.2.3.

3.4 Solving the energy equation

We use the method of Stamatellos *et al.* 2007 to solve the energy equation. This method follows the energy equation (equation 2.54) with the addition of a term representing radiative heating and cooling. The particle is assumed to be embedded in a spherical polytrope of index $n = 2$; the exact choice of index is not important. For any given position in the polytrope, the density and gravitational potential of the particle uniquely fix the central density and scale length of the polytrope. The column density to infinity can then be calculated by integrating through the polytrope. Since the exact position of the particle in the polytrope is not known, a mass-weighted average of all possible positions in the polytrope is taken. This gives an estimated column density for the particle.

The column density is converted to a *pseudo-opacity* which accounts for surrounding material in the polytrope. This allows the cooling rate to be calculated using the Rosseland-mean and Planck-mean opacities. These use the parametrization proposed by Bell & Lin (1994). The tabulated opacity tables include ice mantle melting, the sublimation of dust, molecular lines, H^-, bound–free and free–free processes and electron scattering. Finally a

background temperature is set, representing a background radiation field. No particle can cool radiatively below this background temperature.

The method also includes a more accurate equation of state. It uses pre-calculated tables which relate the specific internal energy to the particle temperature and mean molecular mass. These tables are calculated considering the rotational and vibrational degrees of freedom of H_2, H_2 dissociation and ionization of hydrogen.

Since the estimated column density depends only on the gravitational potential and the density, the column density can be pre-calculated in a table. The equation of state is also tabulated, and the internal energy is integrated using an implicit exponential Euler integrator. This makes the method much faster than a full treatment of radiative transfer.

The method works best when there is spherical symmetry, although it has been shown to work well in other geometries. Our primary aim for the method is to accurately follow the evolution of the gas as it forms dense, self-gravitating objects undergoing Kelvin–Helmholtz contraction. For these objects spherical symmetry should be a reasonable assumption and we expect the method to perform well. As these objects collapse they will form sink particles that represent a single protostar, as long as the sink density is set high enough that further fragmentation is unlikely.

The method is designed for relatively dense environments, following the collapse of a dense core from the end of the isothermal phase at approximately $10^{-13}\,\mathrm{g\,cm^{-3}}$ up to stellar densities at approximately $10^{1}\,\mathrm{g\,cm^{-3}}$. For this reason, it assumes efficient coupling between the gas and dust. However, below densities of approximately $10^{-19}\,\mathrm{g\,cm^{-3}}$ this is no longer true and the gas begins to heat up as it is unable to transfer heat effectively to the dust, which remains at low temperatures. This could lead to an overestimate of cooling at low densities.

3.5 Tree gravity

The calculation of gravity for a set of point masses is a common problem, and a wide range of methods has been developed to solve it efficiently. A direct approach of considering the force between each pair of particles requires N^2 pairs to be considered, where N is the number of particles. This has a computational cost which scales as $O(N^2)$ and becomes rapidly unfeasible for large numbers of particles. SEREN uses the tree-gravity method originally proposed by Barnes & Hut (1986), which scales as $O(N \log N)$.

For any given point in space, the particles closest to that point need to be considered individually to get an accurate estimate of gravity. However, a collection of particles at a distance, where the spacing between the particles is much less than the distance from the point in space to the particles, can be easily approximated without excessive error. The simplest approximation is to consider the collection of particles as a single mass at the centre of mass (the *monopole* moment). This can be extended to higher order

moments. Since the dipole moment in the centre of mass frame is always zero, we use the quadrupole moments to improve on the estimate.

We use the gravity tree built and stocked as described in Section 3.2. We can consider any cell of the tree as a single collection of particles, and replace them with a single monopole and quadrupole contribution. We calculate gravity for a particle by beginning at the root cell of the tree, and considering whether we can approximate the gravitational force from the contents of the cell using the multipole expansion of that cell. Using a cell that is too large, contains too much mass or is too close will not produce a sufficiently accurate answer. We use a *multipole acceptance criterion* (MAC) to decide if the multipole expansion of a cell can be used or whether the cell needs to be 'opened' and each child cell considered in turn. When opening leaf cells, each particle is considered in turn as described in Section 2.4.

3.5.1 Multipole acceptance criteria

The simplest MAC is the original geometric MAC proposed by Barnes & Hut (1986). We define the angle subtended by a cell c as seen from a particle p as

$$\theta_{cp} = \frac{r_{\mathrm{MAX}}}{|\boldsymbol{r}_c - \boldsymbol{r}_p|} \,. \tag{3.1}$$

A cell must be opened, and each of its child cells examined in turn, if the angle θ_{cp} is greater than the parameter θ_{MAC}. Otherwise, we can use the multipole expansion.

We use the 'GADGET MAC' method proposed by Springel, Yoshida & White (2001), based on an estimate of the next terms in the multipole expansion. The next largest term after the quadrupole moment is the octopole moment; however, Springel, Yoshida & White suggest that this is relatively low for a smooth density distribution. They instead calculate an estimated error based on the hexadecapole moment. The acceleration due to the hexadecapole moment is estimated as

$$|\boldsymbol{a}_{\mathrm{HEX}}| \sim \frac{G M_c r_{\mathrm{MAX}}^4}{|\boldsymbol{r}_p - \boldsymbol{r}_c|^6} \,, \tag{3.2}$$

where M_c is the total mass in cell c. Cells must be opened, and each child cell examined in turn, if $|\boldsymbol{a}_{\mathrm{HEX}} > \eta_{\mathrm{MAC}}|\boldsymbol{a}_p|$, where η_{MAC} is an error parameter which we set to 0.05, and $\boldsymbol{a}_p$ is the total gravitational acceleration of the particle p. Since this is not known in advance, the value of the gravitational acceleration for the previous timestep is used.

Unlike the geometric MAC, the GADGET MAC does not guarantee a maximum fractional force error. We therefore use the GADGET MAC method in conjunction with the geometric MAC, but we set $\theta_{\mathrm{MAC}} = 1.0$, which is relatively high and so is usually less stringent than the GADGET MAC. We also use only the geometric MAC prior to the first

timestep, as a value of the gravitational acceleration is not available.

One final check is made before a multipole expansion is used: a leaf cell will always need to be opened if it is so close that the gravitational forces between particles might be softened. This is because the multipole expansion does not include any gravitational softening.

3.5.2 Quadrupole moments

The quadrupole moment tensor for a leaf cell c, summing over particles in the leaf b and with Cartesian indices i and j is given by

$$\boldsymbol{Q}_{ij,c} = \sum_{b=1}^{N} m_b \left[3(\boldsymbol{r}_b)_i (\boldsymbol{r}_b)_j - |\boldsymbol{r}_b|^2 \delta_{ij} \right] , \tag{3.3}$$

where $\boldsymbol{r}_b$ is the vector from cell c to particle b, $(\boldsymbol{r})_i$ is the ith cartesian component of $\boldsymbol{r}$, and δ_{ij} is a Kronecker delta. For a link cell, the summation of particles b is replaced by a summation of daughter cells d and is given by

$$\boldsymbol{Q}_{ij,c} = \sum_{d=1}^{N} m_d \left[3(\boldsymbol{r}_d)_i (\boldsymbol{r}_d)_j - |\boldsymbol{r}_d|^2 \delta_{ij} \right] + \sum_{d=1}^{N} \boldsymbol{Q}_{ij,d} \, . \tag{3.4}$$

When using the multipole expansion, we do not use any gravitational softening (see Section 2.4). For this reason, as described in Section 3.5.1, we always open cells that could contain particles whose gravity would be softened. For those cells that pass the multipole acceptance criteria, the gravitational acceleration on the particle due to that cell c is given by

$$\boldsymbol{a}_{\text{GRAV}} = -\frac{GM_c}{|\boldsymbol{r}|^3}\boldsymbol{r} + \frac{G\boldsymbol{Q}_{ij,c}(\boldsymbol{r})_i}{2|\boldsymbol{r}|^5}\widehat{\boldsymbol{e}}_j - \frac{5G\boldsymbol{Q}_{ij,c}(\boldsymbol{r})_i(\boldsymbol{r})_j}{2|\boldsymbol{r}|^7}\boldsymbol{r} \, , \tag{3.5}$$

where we use the Einstein summation convention over repeated Cartesian indices i and j, $\widehat{\boldsymbol{e}}_j$ is the unit vector in the j direction, M_c is the total mass of the cell and $\boldsymbol{r}$ is the vector from the cell to the particle.

3.6 Time integration

SEREN uses hierarchical block timesteps. Instead of the simulation being advanced with a single global timestep, typically only a subset of particles have force calculations in any given timestep. This allows those particles that require more frequent force calculations to be calculated with short timesteps, while saving the computational cost of calculating forces for particles which do not need such frequent timesteps. The particles are assigned an optimal timestep as described in Section 3.6.1, and placed in a hierarchy as described in Section 3.6.2. Finally for each timestep, all particles are advanced using the Leapfrog

integration scheme described in Section 3.6.3.

3.6.1 Timestep criteria

We define an optimal time step for each particle that we do not use directly, but use to place the particle into the hierarchy of block timesteps described in Section 3.6.2. The optimal timestep τ_p is defined as the minimum of a modified Courant timescale τ_{COUR} and an acceleration timescale τ_{ACC}.

The modified Courant timescale includes a contribution for particles in shocks, and therefore uses the particle's current viscosity factor κ_p. It is defined as

$$\tau_{\text{COUR}} = \frac{\gamma_{\text{COUR}} h_p}{(1+0.6\kappa_p)\, c_p + (1+1.2\kappa_p)\, h_p |\nabla \cdot \boldsymbol{v}|_p} \,, \tag{3.6}$$

where γ_{COUR} is the Courant timestep multiplier (which we set to 0.3), c_p is the sound speed at particle p and $|\nabla \cdot \boldsymbol{v}|_p$ is the magnitude of the velocity divergence at particle p. The velocity divergence is used instead of the absolute particle velocity as this keeps the method Galilean invariant.

The acceleration timescale is defined as

$$\tau_{\text{ACC}} = \gamma_{\text{ACC}} \left(\frac{h_p}{|\boldsymbol{a}_p| + \eta} \right)^{1/2} \,, \tag{3.7}$$

where γ_{ACC} is the acceleration timestep multiplier (which we set to 0.4), $\boldsymbol{a}_p$ is the total acceleration of particle p, and η is a small number to avoid singularities.

3.6.2 Block timesteps

SEREN uses a binary hierarchy of timestep levels as shown in Figure 3.3, which means that the timestep of each level is twice the timestep of the level below it. We set SEREN to use up to 30 levels of the hierarchy, with 25 of those initially spare.* Initially, the smallest particle timestep is used to set the timestep of the lowest non-spare level. This fixes the timesteps of every level. First the timestep of the smallest timestep level, t_{min}, is determined. Then the timestep of each level is simply $t_{\text{level}}(l) = t_{\text{min}} 2^{l-1}$, where 2^{l-1} is the *integer timestep* shown in Figure 3.3. This process is updated each longest timestep level as all timestep levels are aligned at this point.

Calculating forces for particles is computationally expensive, whereas it is relatively cheap to 'drift' particles along using the previous value of their acceleration. It is this property that block timesteps exploit, as particles have their forces calculated only when necessary and are otherwise extrapolated as described in Section 3.6.3.

*When using an integration scheme that uses 'half-steps', such as the Leapfrog scheme used in this work, the lowest spare level is unavailable, as it is reserved for the half-steps of any particles on the second lowest level.

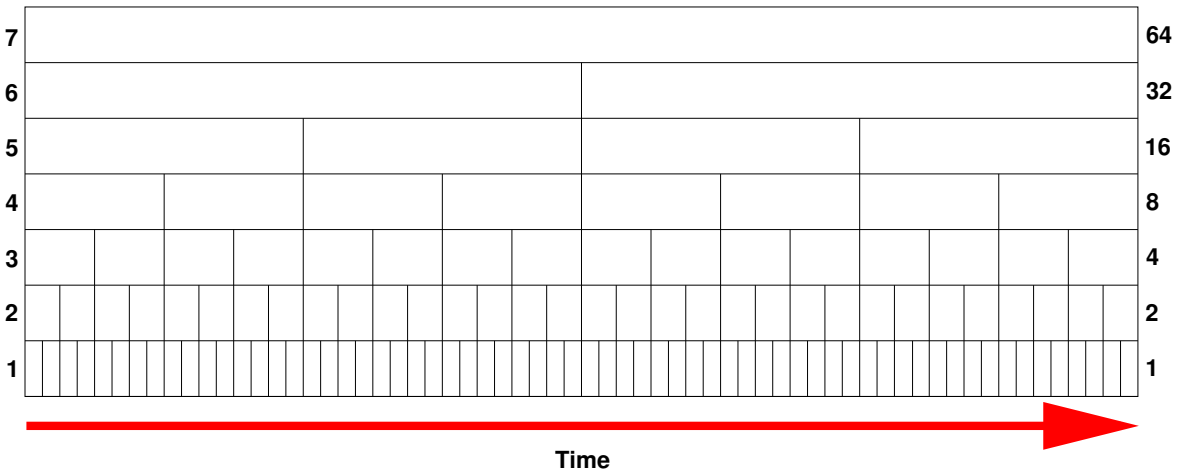


Figure 3.3 – Illustration of the block timestep hierarchy, with longer timesteps at the top. The left-hand numbers indicate the level number, and the right-hand numbers indicate the integer timestep. As time passes, particles on each level in the hierarchy have recalculations of their forces at different intervals; each box represents a single timestep for all particles on that level.

On any given simulation timestep, those particles that need their accelerations recalculated are referred to as *active* particles. All other particles are not active particles. It is common for only a tiny fraction of all particles to be active in any given simulation timestep.

3.6.3 Leapfrog integration scheme

We use the kick-drift-kick Leapfrog integration scheme. The Leapfrog scheme is symplectic and time-reversible, and consequently conserves energy. This makes it an excellent integrator for self-gravitating simulations, as orbits are modelled without secular changes in orbital energy. It is also computationally efficient, requiring only one acceleration calculation per integration step.

We define the length of the timestep (from step n to $n+1$) as $\Delta t = t(n+1) - t(n)$. At the half-step $n + 1/2$ and full-step $n + 1$, the kick-drift-kick Leapfrog scheme is defined by

$$\boldsymbol{r}_{n+1/2} = \boldsymbol{r}_n + 1/2 \boldsymbol{v}_n \Delta t \ , \tag{3.8}$$

$$\boldsymbol{v}_{n+1/2} = \boldsymbol{v}_n + 1/2 \boldsymbol{a}_n \Delta t \ , \tag{3.9}$$

$$\boldsymbol{r}_{n+1} = \boldsymbol{r}_n + \boldsymbol{v}_{n+1/2} \Delta t \ , \tag{3.10}$$

and

$$\boldsymbol{v}_{n+1} = \boldsymbol{v}_{n+1/2} + 1/2 \boldsymbol{a}_n \Delta t \ . \tag{3.11}$$

3.7 Sink particles

Prior to the introduction of *sink particles* by Bate, Bonnell & Price (1995), simulations of star formation were typically halted after the formation of the first dense object. As gas density increases, shorter timesteps are required. As gas collapses to stellar densities, computational constraints make it difficult to follow these small timesteps for more than a single object, even with modern computing facilities. Sink particles allow a large number of SPH particles in a dense object to be replaced by a single sink particle. The sink particle represents a simple model of a protostar, tracking only the centre of mass position and velocity, and angular momentum of the protostar. The sink particle feels gravitational forces but does not interact hydrodynamically with its surroundings. Sink particles accrete particles that fall within their *accretion radius*, allowing the evolution of the protostar to be followed as they accrete gas from the surrounding medium.

3.7.1 Sink particle criteria

A number of sink criteria have been devised to ensure that sinks are introduced to replace only gravitationally bound dense objects. Any SPH particle which fulfils the sink criteria is removed and replaced with a sink particle. We use the following criteria:

- the candidate particle must have a density above the *sink density*, which we set to $10^{-11}\,\mathrm{g\,cm^{-3}}$;
- the candidate particle should be at a local minimum of the gravitational potential energy; it should have the lowest gravitational potential energy of all of its SPH neighbours;
- the candidate particle and its SPH neighbours should be bound, considering only the gravitational potential energy and the kinetic energy. We disregard the thermal energy as this can appear to unbind objects on the scale of a particle smoothing kernel which are gravitationally bound on larger scales;
- the local value of the velocity divergence at the position of the candidate particle should be negative, indicating the object is currently collapsing; and
- the candidate particle must not be so close to an existing sink that, if the candidate particle was replaced with a sink, their accretion radii (see Section 3.7.2) would overlap.

Sink criteria should be used with caution if the sink density is chosen to prevent gas reaching a density higher than that resolved by SPH, as described in Section 2.4.1. If the sink criteria prevent the formation of sinks, then gas can reach densities higher than those resolved and this may lead to erroneous results. For example, if a realistic equation of

state is used which accounts for the opacity limit and resultant increase in temperature at a critical density (see Section 2.4.1), then gas will fragment into distinct objects at that density. If the sink density is lower than that density, gas will still be isothermally collapsing and fragmenting at the sink density. Sensible sink criteria may identify that the gas is not a distinct, self-gravitating object, prevent the creation of a sink, and the gas density may increase above the resolution limit. However if the sink density is much higher than the critical density, then all gas at this density can be expected to have already formed dense, distinct objects, and the sink criteria (other than the density criterion) should rarely prevent the formation of a sink.

3.7.2 Sink particle accretion

Once a new sink has been created, it immediately accretes all SPH particles within its *accretion radius*. This is chosen by calculating the *sink density smoothing length*, $h_{\rm SINK}$, which a particle would be expected to have at the sink density. The accretion radius is then (for our kernel) $2h_{\rm SINK}$. This means that all sink particles have the same accretion radius, roughly matching the size of the smoothing kernels of particles at the highest expected particle density.

Sink particles do not feel hydrodynamic forces, and have larger masses than other particles, but otherwise their motion is followed as for any other particle. They do not have a density or a smoothing length, and therefore do not have SPH neighbours. When calculating the kernel softened gravitational forces described in Section 2.4, the sink density smoothing length $h_{\rm SINK}$ is used. The centre of mass position and velocity of the sink are updated to include contributions from the accreted particles. The angular momentum of the sink is increased to include the contributions from sink particles, but plays no further part in the simulations.

Sink particles continue to accrete particles after their creation. SPH particles that fall within the accretion radius of a sink at the end of a subsequent timestep are accreted to that sink and removed. If a particle falls within the accretion radius of more than one sink, it is accreted by the closest sink to the particle. Accreted particles are removed from the simulation. The BH trees are modified to remove references to the accreted particles. This removes the need to rebuild the tree, although the tree must still be restocked as usual.

Chapter 4

Parallelization

State-of-the-art scientific computing requires access to large computing facilities. Individual computing units are often insufficient for modern research, requiring the use of multiple computing units simultaneously. The methods used to achieve this parallelization depend on the nature of the computing facilities available and on the nature of the problem to be solved.

4.1 Methods of parallelization

We define a number of terms to simplify this discussion, as illustrated in Figure 4.1:

- A *computing unit* processes a series of instructions; it is the smallest processing unit on which a computing code can be independently processed.
- A *memory unit* is a rapid-access store for the code and the data being processed, and is accessed by the *computing unit*.
- A *node* is a single unit, which may contain several *computing units*, but contains only a single *memory unit* accessible by all *computing units* within it.
- A *shared-memory machine* is a single *node* which has a large number of *computing units*.
- A *distributed-memory cluster* is a collection of independent *nodes*, which can communicate only by explicit messaging, typically over a network link.
- A *code* is an executable program.
- A *process* is a single copy of the *code*, running on one or more computing units on the same node.
- A *thread* is a part of a *process*, which runs on a single *computing unit*. Every *process* has at least one *thread*.

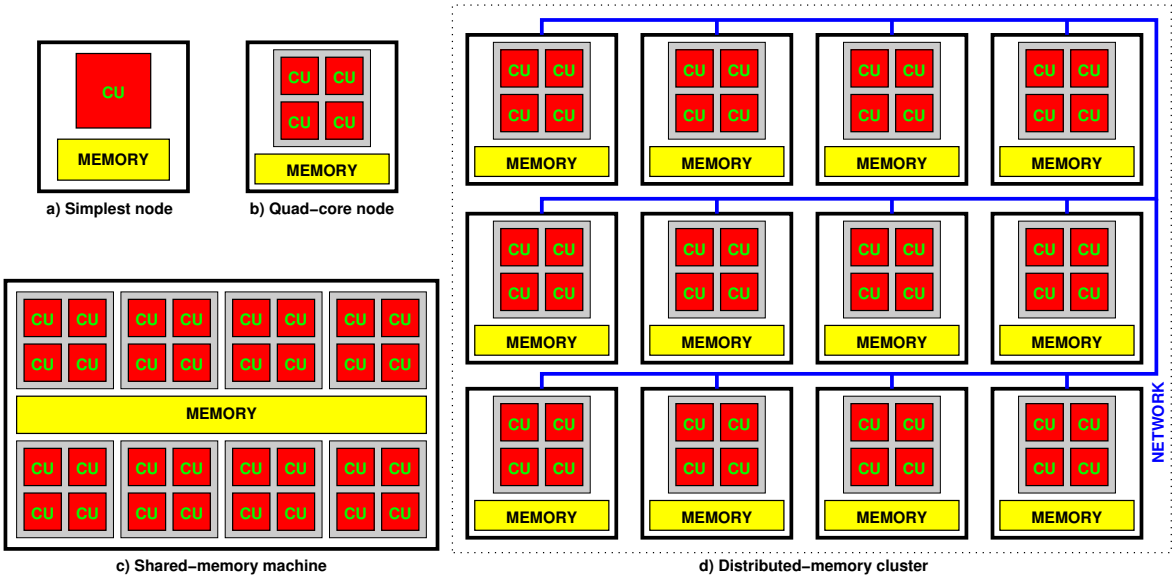


Figure 4.1 – Four possible computing facilities: a) The simplest possible *node* of one *computing unit* and *memory unit*; b) A typical *node* of a single quad-core CPU; c) A shared-memory machine with 32 *computing units* and a single *memory unit*, containing of 8 quad-core CPUs; d) A distributed-memory cluster with 48 *computing units* and 12 separate *memory units*, made up of 12 separate *nodes* each with a single quad-core CPU. Red boxes marked 'CU' are *computing units*, sometimes contained with grey boxes representing individual CPUs. Yellow boxes marked 'memory' are *memory units*. Blue lines indicate network links.

We can relate our definitions to the hardware typically used in modern computing equipment. Due to current technical limitations in the maximum speed of a single processing unit, a typical desktop computer contains a single multi-core CPU. In this case, each core is a complete *computing unit* by our definition. A typical machine used for scientific computing may have several multi-core CPUs, but *computing unit* still refers to each individual core, as each core is capable of independent processing. The *memory unit* in a desktop computer is usually referred to as the RAM.

A code running without parallelization is a *serial* code, which has a single main *thread*. A serial code has full access to the *memory unit*, but as it has a single *thread* it can use only one *computing unit*. Where it is desired to run the same code over some set of parameters, a trivial form of parallelization is to run multiple *processes* of a serial code on a single *node*; each *process* will use a single *computing unit*. To speed up computation of a single large job a more advanced form of parallelization must be used.

4.1.1 Shared-memory parallelization – OpenMP

Shared-memory parallelization is typically used when all the computing units to be used share a common memory space, so that all computing units can access the same memory. This form of parallelization can be easier to implement and may have fewer computational overheads. However, it can be difficult to implement efficiently with a large number of

computing units, and requires relatively expensive shared memory machines.

Shared-memory parallelization uses a single *process*, which typically uses sufficient *threads* to use all the *computing units* on a shared-memory machine. The simplest method of implementing shared-memory parallelization for scientific work is through the language extension OpenMP. OpenMP permits sections of code to be parallelized onto multiple threads, each of which can run on one computing unit. The code initially begins with only one *master* thread active, and the remaining *worker* threads idle. The code proceeds in serial until the first OpenMP parallelized section, indicated by OpenMP extensions, is reached. Typically this will be an indexed loop over some range of values, which will then be split up into contiguous chunks of that range. The chunks will be distributed to each thread to be processed. Once this is complete, any necessary book-keeping can take place, and the worker threads returned to idle. The code then continues in serial until the next OpenMP section.

OpenMP parallelization is used in the standard (non-MPI) version of SEREN described in Hubber *et al.* (2011). In SEREN, OpenMP is typically used to parallelize over lists of particles. All particles need to have the same basic operations performed on them: the calculation of smoothing lengths and densities; the calculation of SPH and thermal quantities; the calculation of forces; and integration using the Leapfrog scheme. The loops performing these tasks for each particle are parallelized so that each thread receives a chunk of the particles, and therefore a chunk of the work. Further details are given by Hubber *et al.* (2011).

Shared-memory parallelization is also used by other SPH codes, such as VINE (Wetzstein *et al.* 2009), and SEREN's predecessor DRAGON (Goodwin, Whitworth & Ward-Thompson 2004; Turner *et al.* 1995).

4.1.2 Distributed-memory parallelization – MPI

The alternative to shared-memory parallelization is distributed-memory parallelization. In this case, each computing unit is separate and can communicate only through explicit messages, typically over a network link. This form of parallelization is often more difficult or even impossible to implement efficiently, and may have more computational overheads. However, some problems that can be parallelized with this method scale well even when using thousands or tens of thousands of computational units. Distributed-memory computing clusters are often considerably larger, than equivalent shared-memory facilities.

As distributed-memory parallelized codes cannot access a globally shared memory, they must run multiple separate *processes*, typically one per computing unit, and communicate via *messages*. Most distributed-memory parallelized codes use an MPI (Message Passing Interface) library to pass messages between these processes. Each process, referred to as an *MPI task*, is a complete and initially identical copy of the code. The code is therefore always parallelized, as opposed to a code using OpenMP which is primarily a serial code

with parallel sections. A typical approach is to divide the problem into sections, and have each MPI task work on a section. Communication is required where sections may overlap or interact. As a result, MPI parallelization is significantly more complex than OpenMP parallelization.

MPI parallelization is used by a wide range of AMR (adaptive mesh refinement) codes, such as FLASH (Fryxell *et al.* 2000), RAMSES (Teyssier 2002), ZEUS_MP (Hayes *et al.* 2006) and ENZO (O'Shea *et al.* 2004). It is also used by a number of SPH codes, such as GADGET (Springel, Yoshida & White 2001), GADGET-2 (Springel 2005) and GASOLINE (Wadsley, Stadel & Quinn 2004).

4.1.3 Hybrid parallelization

Modern CPUs tend to include multiple cores per CPU, and individual nodes of a computing cluster may contain more than one CPU. The result is that each node may contain several computing units. Due to the overheads of MPI parallelization, it may be more efficient to use shared-memory parallelization on each individual node, and use distributed-memory parallelization to link the nodes.

This is the approach taken with our new SEREN-MPI code. Only one MPI process is run per node, minimizing the MPI overhead. Within each node, OpenMP parallelization is used to take advantage of multiple computing units per node. This hybrid approach minimizes the division of domains in SEREN-MPI, and increases the overall efficiency. The OpenMP parallelization is effectively identical to the approach used in standard SEREN.

4.2 General modifications to SEREN-MPI

SEREN-MPI is a new hybrid OpenMP/MPI parallelized version of the SEREN code. Each node runs a single MPI task, which is itself a complete version of the SEREN-MPI code. Each MPI task works on a subset of the total particles in the simulation, and a single associated region of space which we refer to as a *domain*. These domains are cuboids and may overlap. Each MPI task is allocated a single domain, and is unaware of anything outside its own domain. Any calculations in a given task may require communication with other tasks to determine if there are particles in those tasks that must be considered.

4.2.1 Block timesteps in MPI

SEREN-MPI uses multiple particle timesteps in the same way as used in standard SEREN (see Section 3.6.2). When each task calculates the smallest optimal timestep, it is necessary for the tasks to explicitly communicate through MPI to find the smallest timestep across all domains. This is shared across all tasks, and used for calculation of timesteps. This ensures that time is kept synchronised across all domains.

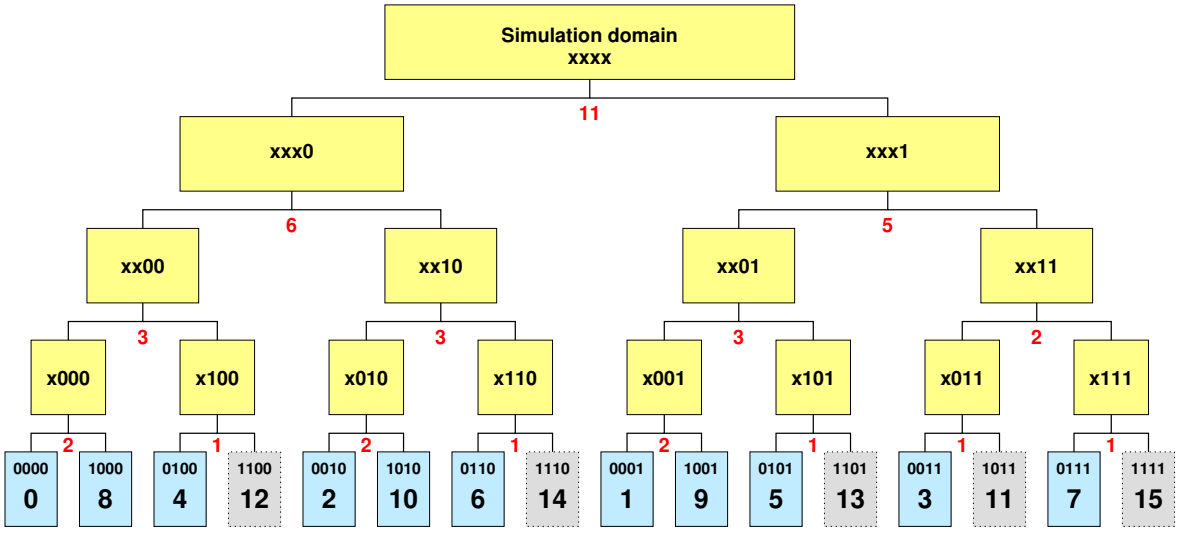


Figure 4.2 – The binary decomposition tree for 11 MPI tasks, showing the tree structure (yellow boxes), the eleven *domains* (light blue boxes), inactive tree cells not associated with domains (light grey boxes), the binary digits used to allocate tasks to domains, and the tree weighting factors (in red).

4.2.2 Domain decomposition and load-balancing

The *simulation domain* is the complete region of space being simulated. This is divided into one spatial *domain* per task. Each domain is assigned to a single MPI task. Initially, and immediately after a load-balancing step, these domains do not overlap, and all particles in an domain are therefore assigned to a single MPI task. As particles move, the domains expand with them, causing domains to overlap. Particles do not change their assignment to an MPI task between load-balancing steps. Each MPI task does not know about particles that are not assigned to it, even if (due to overlapping domains) they fall inside the MPI task's domain. The MPI task must communicate with other MPI tasks to determine if there are particles within their overlapped boundaries.

The domain decomposition is performed by repeated bisection of the *simulation domain*. Cuts are made in arbitrary dimensions in a binary tree structure in order to equalize the amount of *work* in each domain. The algorithm can deal with an arbitrary number of MPI tasks while attempting to remain balanced.

Figure 4.2 shows the division of cuts for 11 MPI tasks. Each split in the tree represents a cut. The figure also shows how tasks (numbered from 0 to $N-1$ where N is the total number of tasks) are allocated using a binary number. For each cut, the 'left' side is indicated by a 0, and the 'right' side is indicated by a 1. The digit of the first cut is set in the least-significant bit (LSB). The digit of the next cuts are set in the next least-significant bit, and so on until the bottom level of the tree. The tasks are then allocated according to the binary value as shown. Digits higher than the highest MPI task number are inactive, and not used.

In order to use the tree to evenly decompose the simulation into chunks of similar computational cost, it is necessary to attempt to quantify *work*. SEREN-MPI builds a 128^3 grid of *work* that spans the entire simulation domain, and uses a number of sources of information to build this grid. This grid, or a portion of the grid, can then be cut to arbitrarily divide the *work*. Initially, the only information available is the positions of the particles, and so the initial domain decomposition simply places approximately equal particles numbers into each domain. This suffices for simple problems, but for more advanced decomposition, a number of criteria are used, described in Section 4.2.3.

Tree balancing factors are calculated by simply adding the number of active cells in the tree at each cut, as shown by red digits in Figure 4.2. At a cut, the amount of *work* on each side is distributed according to these weighting factors; for example in the figure the first cut would attempt to place six-elevenths of the work on the 'left' side, and five-elevenths of the work on the 'right' side, as this should eventually result in equal distribution of work.

The equal division of work does not depend on the axis in which cuts are made, and we are free to choose any Cartesian axis at any point. We calculate an analogue of the moment of inertia of each cell in the domain decomposition tree along each Cartesian axis, replacing mass with the values from the grid of *work*. Cutting along the axis with the highest moment of inertia would cut across, rather than along, filaments. Cutting along a filament would maximise the number of particles along a boundary and increase the amount of communication required. It is better to cut across a filament, minimizing communication. Our work-based analogue to the moment of inertia similarly aims to minimize communication by cutting on the axis with the highest value.

4.2.3 Particle *work*

For each particle p, a *work* value Ψ_p is calculated, defined as

$$\Psi_p = N_{\mathrm{LOAD}} \left(\xi_{\mathrm{BASE}} + \frac{\tau_{\mathrm{MIN}}}{\tau_p} \right) \left(\xi_{\mathrm{HYDRO}} \times n_{\mathrm{HYDRO}} + \xi_{\mathrm{GRAV}} \times n_{\mathrm{GRAV}} \right) , \tag{4.1}$$

where N_{LOAD} is the estimated number of timesteps to the next load-balance, τ_{MIN} is the current simulation timestep, τ_p is the timestep of the particle p. n_{HYDRO} and n_{GRAV} are counts, for each particle, of the number of hydrodynamic calculations (effectively the number of neighbours) and the number of gravitational calculations (as defined below) that the particle has performed in the previous timestep. The relative estimated cost per particle, per hydrodynamic calculation and per gravity calculation can be adjusted with the parameters ξ_{BASE}, ξ_{HYDRO} and ξ_{GRAV} respectively. For sink particles we replace the hydrodynamic and gravitational parameters with a fixed parameter ξ_{SINK}. The number of

gravitational calculations is set as

$$n_{\mathrm{GRAV}} = \xi_{\mathrm{QUADRUPOLE}} \times n_{\mathrm{CELL}} + \xi_{\mathrm{MONOPOLE}} \times n_{\mathrm{PARTICLE}} \,, \tag{4.2}$$

where $\xi_{\mathrm{QUADRUPOLE}}$ and ξ_{MONOPOLE} are parameters as before, n_{CELL} is the number of tree cells for which the particle used a quadrupole expansion, and n_{PARTICLE} is the number of other particles for which the particle used direct kernel-softened gravity.

We use the following parameters:

$$\begin{aligned} \xi_{\mathrm{BASE}} &= 0.0 \,, \\ \xi_{\mathrm{HYDRO}} &= 0.01 \,, \\ \xi_{\mathrm{GRAV}} &= 0.00001 \,, \\ \xi_{\mathrm{MONOPOLE}} &= 1.0 \,, \\ \xi_{\mathrm{QUADRUPOLE}} &= 2.5 \,, \end{aligned}$$

and

$$\xi_{\mathrm{SINK}} = 1000 \,.$$

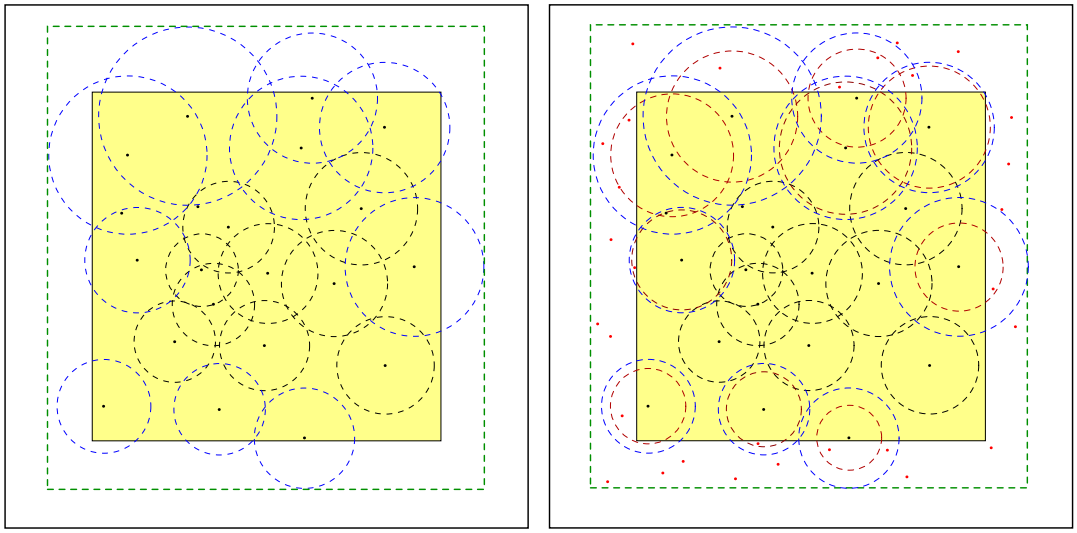
This particle *work*, Ψ_p, includes an estimate of the number of timesteps the particle will undergo before the next load-balancing step. As particle timesteps can vary, this is only an estimate. Similarly, by predicting the number of timesteps for each particle, the code attempts to load-balance at regular wall-time intervals. If the simulation is rapidly evolving, many or all of these predictions may be wrong. Despite these caveats the method works well for all problems so far tested.

4.3 Finding smoothing lengths and densities

Finding smoothing lengths and densities, as described in Section 3.3, is the first challenge for an MPI task. Only *active* particles (see Section 3.6.2) need to have their smoothing lengths recalculated in each timestep. As each MPI task is only aware of its own particles, it will over-estimate smoothing lengths near the boundaries of the domain, as particles will be missing neighbours. Each calculated smoothing length therefore represents a maximum possible smoothing length. We define a box containing the smoothing kernels of all active particles: this is a *smoothing kernel bounding box* (SKBB). Only particles inside the SKBB can affect the smoothing lengths of this domain's particles.

Figure 4.3a shows an example of a domain after the first pass. Particles whose smoothing kernels overlap another domain are indicated; these smoothing lengths are probably over-estimates.

Once each MPI task has calculated a first pass of calculating smoothing lengths, each task calculates its SKBB, and broadcasts this to all other domains. Each domain then



(a) First pass (b) Second pass

Figure 4.3 – The two-pass method of finding smoothing lengths in MPI. The local domain is shown by a yellow box; a sample of local particles and their smoothing kernels are indicated by black points and dashed-line circles. The smoothing kernel bounding box is shown as a dark green dashed line. Ghost particles are shown as red points. Black smoothing kernels are calculated correctly on the first pass, blue smoothing kernels are potential over-estimates and are replaced on the second pass with the correct dark red smoothing kernels.

identifies any of its own particles that fall within the SKBB of other domains. This includes particles that are not active, as they may still contribute to the smoothing length of an active particle. Each domain then sends the mass, positions and velocities of these particles to the relevant domains, so that every domain receives the positions of all particles within their SKBB. This information is stored as *ghost* particles. Using these ghost particles, each domain can recalculate smoothing lengths for the particles whose smoothing kernels overlapped another domain boundary.

Figure 4.3b shows the results of the second pass on the example domain shown in Figure 4.3a. Ghost particles have been imported, allowing a recalculation of those particles whose smoothing kernels overlapped another domain boundary. Their smoothing lengths have been reduced due to the additional contribution from the ghost particles.

Finally, ghost particles are used in the calculation of the local velocity divergence. This can, as described in Price 2004, be calculated without knowledge of neighbouring particle density. This is useful as ghost particles are sent before the second pass of smoothing lengths and density calculation is complete. Velocity divergence is used later to calculate timesteps, to calculate artificial viscosity and when searching for new sink particles.

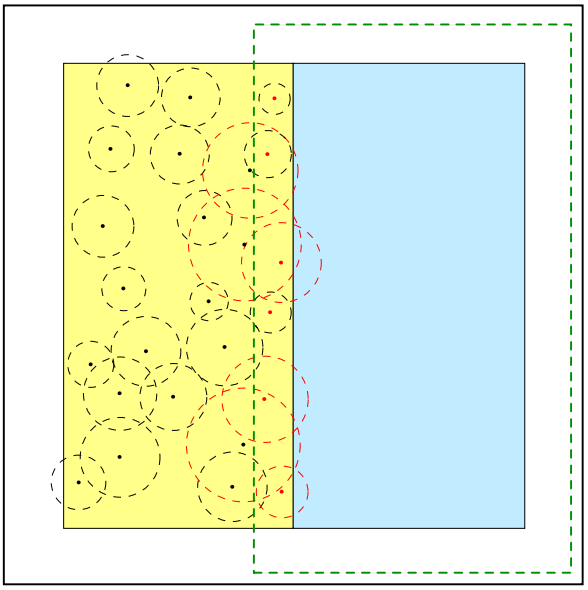


Figure 4.4 – Showing particles in one local domain (pale yellow box) which need to be exported to another domain (pale blue box). A sample of local particles and their smoothing kernels are indicated by black points and dashed-line circles. Particles to be exported because they fall within the blue domain's smoothing kernel bounding box (dashed green line) are shown in red. Particles to be exported because their smoothing kernels overlap the blue domain are indicated by red smoothing kernel circles.

4.4 Hydrodynamic forces

Hydrodynamic forces are short-ranged. We therefore need to consider only particles near the domain boundaries. We use a similar scheme to Springel (2005) to calculate hydrodynamic forces by exporting particles. First a smoothing kernel bounding box is constructed, similar to the one used in Section 4.3. However this SKBB contains the smoothing kernels of all particles in the domain, rather than just active particles.

This SKBB is broadcast to other domains as before, where it is used to determine if a particle may have neighbours in a neighbouring domain. For hydrodynamic forces, a particle b is a neighbour of particle a when particle b falls within the smoothing kernel of particle a or particle a falls within the smoothing kernel of particle b. This ensures forces are fully symmetrized. A particle therefore needs to be exported to a domain if either its smoothing kernel overlaps the domain boundary (and could include particles in another domain), or is overlapped by the SKBB of the other domain (and so could fall within the smoothing kernel of a particle in that domain).

If a particle is to be exported, only the relevant properties required for calculation of forces are sent to the other domain. A particle may be sent to several domains in this way. Hydrodynamic forces are calculated as described in Section 3.3. The resultant force is returned to the original domain of the particle, and combined with the forces due to local particles.

4.5 Gravitational forces

Gravity is a long-ranged force. The method used for hydrodynamic forces would be inefficient, as all particles would need to be exported for forces. In a similar way to Springel (2005), we use pruned gravity trees to reduce the number of particles that need to be exported for gravity calculations.

Each MPI task builds a gravity tree for its own domain (Section 3.2.1). This gravity tree is used for local tree-gravity (Section 3.5). Each MPI task makes a copy of this gravity tree, and prunes it leaving only the top few layers; we use the top four layers. The tree-building routine is modified to ensure there are no leaf cells on these levels. This pruned tree, composed entirely of link cells, is then broadcast to every other MPI task, and each task receives a copy of the tree for every domain. The pruned tree is much smaller than the complete tree, and it is not computationally expensive to send it to all MPI tasks.

Every particle then has tree-gravity calculated for the pruned tree of each other domain in turn. For distant domains, the multipole acceptance criteria (Section 3.5.1) may often allow a particle to complete a tree-gravity walk using only the cells available in the pruned tree. For efficiency, the tree-walk is performed first, and the tree-gravity force contributions are calculated only if the walk is successfully completed.

For less distant domains, a particle may find that it has to open a cell on the bottom level of the pruned tree. This information is not available in the pruned tree, and so the tree-walk is abandoned. Instead the particle must be exported to the relevant MPI task. As for the hydrodynamic forces, particles may be exported to several MPI tasks for force calculation, and the results returned and combined with local forces. Particles very near to boundaries will need to be exported, as they will need to calculate kernel-softened gravity with individual particles. However, particles far from boundaries will rarely need to be exported to any domain.

4.6 Sink particles in MPI

There will be relatively few sink particles in any simulation, compared with the number of SPH particles, and so it is not computationally expensive for all MPI tasks to be aware of all sink properties. Each sink is 'owned' by a single task, and that task calculates forces and moves the sink in a similar way to SPH particles. That MPI task then broadcasts the updated information about the sink to all other MPI tasks.

4.6.1 Sink particle tests with ghost particles

Searching for new sinks, as described in Section 3.7.1, requires information about neighbouring particles that is not available until after the initial calculations of SPH quantities described in Section 3.3. We therefore use a new set of ghost particles. First each domain identifies candidate particles by finding particles with a density greater

than the sink density. Candidate particles must not be so close to an existing sink that, if the candidate particle was replaced with a sink, their accretion radii (see Section 3.7.2) would overlap.

For each candidate particle, a *search sphere* is defined that is the larger of the particle's smoothing kernel and the sink accretion radius. This is because the particle needs to consider other particles within its smoothing kernel for sink creation criteria, and if a sink is created, it will accrete particles within the sink accretion radius.

For each domain, a *search sphere bounding box* (SSBB) is then constructed just large enough to contain the search spheres for all the candidate particles in this domain. This is broadcast to all domains. Each domain sends ghost particles that fall within the SSBB of another domain in an analogous way to that described in Section 4.3. These ghost particles contain the information relevant to the sink creation criteria: position, velocity, mass and gravitational potential energy.

Each ghost particle also contains the ID of the corresponding SPH particle on its original domain, and each MPI task receiving ghost particles tracks which task each ghost particle originated from. This allows the original particle to be identified for any ghost particle.

Each candidate particle then checks, using the ghost particles, that: the candidate particle is at a local minimum of the gravitational potential energy; that the candidate particle and its neighbours are bound, disregarding the thermal energy; and that the local velocity divergence at the particle is negative. If the candidate particle passes all sink creation criteria, it is replaced with a sink particle.

4.6.2 Sink particle accretion in MPI

Once a new sink has been created, all particles within the accretion radius are accreted and removed. Ghost particles contain sufficient information for the sink properties to be updated following the accretion as described in Section 3.7.2. Local particles can simply be removed from the simulation. The originating MPI task of each ghost particle is known, and each task that sent ghost particles is sent a list of the any accreted particles. Each MPI task is then able to remove particles that have been accreted on other tasks.

This method of accretion is used only for new sink particles. For existing sinks, each MPI task is aware of the position and accretion radius of all sinks, and can determine which of its particles have been accreted by each sink. For sinks whose properties are calculated on a different MPI task, it is necessary to send that task the details of the accreted particles. For each sink, there will be a set of accreted particles. The total mass and centre-of-mass of this set is calculated, together with the angular momentum around this centre of mass and around the sink particle. This information from each MPI task is then collected and combined for each sink, allowing the calculation of the new position and velocity of the sink particle due to accretion.

Chapter 5

Analysis techniques

Later in this work we present the results of SPH simulations. This chapter describes the techniques we use to analyse the data from these simulations. We describe the methods we use to visualize SPH simulation data as cross-section and column-density plots, and as gridded data. We use this gridded data to identify material belonging to a dense layer in a collision simulation. We describe the analysis of the growth of instabilities in those dense layers. Finally we describe the analysis of sink particles, and the construction of mass functions.

5.1 The SPLASH visualization package

In this work most plots of simulations are made using the freely-available SPLASH visualization package developed by Price (2007). This allows the discrete particle data to be interpolated into a more useful form.

We use three major functions of SPLASH in our analysis. Firstly, we can take a cross-section through the simulation, and measure the value of some quantity along the plane of the cross-section. Secondly, we can produce integrated column plots, such as a plot of column density. Finally SPLASH allows SPH data to be interpolated onto a three-dimensional grid, facilitating further analysis.

SPLASH uses the standard SPH interpolation described in Section 2.1. It uses the standard M_4 kernel without the modification of Thomas & Couchman (1992). The interpolation of a quantity A at a point $\boldsymbol{r}$ is then given by a weighted sum over all particles who fall within a smoothing kernel surrounding the point $\boldsymbol{r}$. The weights are calculated using a kernel function $W(\boldsymbol{r}, h)$, giving

$$A(\boldsymbol{r}) \approx \sum_b m_b \frac{A_b}{\rho_b} W\left(|\boldsymbol{r}_b - \boldsymbol{r}|, h\right) , \tag{5.1}$$

where b is the particle index, $\boldsymbol{r}_b$, ρ_b, m_b and A_b are the values of the position, density, mass and quantity A for the particle b, and h is the chosen smoothing length for the interpolation point. The choice of smoothing length may depend on the details of the

interpolation. For example if data is being interpolated onto a grid then the smoothing length should be larger than the half the grid spacing.

If the quantity A is set to one everywhere, then the interpolated value at any point should also be approximately one, such that

$$1 \approx \sum_b \frac{m_b}{\rho_b} W\left(|\boldsymbol{r}_b - \boldsymbol{r}|, h\right) . \tag{5.2}$$

An interpolation may be normalized by dividing by the quantity on the right-hand side of Equation 5.2. This provides greater accuracy in simulations without free surfaces as described by Price (2007).

5.1.1 Cross-section plots

SPLASH allows a two-dimensional cross-section to be taken through the output of an SPH simulation. This is illustrated in Figure 5.1. A two-dimensional grid of square pixels is set up on the plane of the cross-section. Particles whose smoothing kernels overlap pixels on the cross-section will contribute to those pixels. Any particle property, such as a temperature, pressure or a component of velocity, can be interpolated onto the cross-section. In this section we assume the desired quantity is density.

For each pixel (x, y) on the cross-section plane, SPLASH calculates the sum

$$A_{(x,y)} \approx \sum_b m_b \frac{A_b}{\rho_b} W\left(|\boldsymbol{r}_b - \boldsymbol{r}_{(x,y)}|, h\right) , \tag{5.3}$$

where $\boldsymbol{r}_{(x,y)}$ is the position of centre of the pixel. The smoothing length h is chosen such that

$$h = \max(h_b, \frac{\Delta x}{2}) , \tag{5.4}$$

where h_b is the smoothing length of particle b and Δx is the pixel size.

For reasons of efficiency SPLASH does not use this summation order directly. Instead SPLASH loops over particles, and calculates their contribution to grid pixels.

5.1.2 Column plots

SPLASH can calculate the value of a quantity integrated along columns, such as column density. This is illustrated in Figure 5.2. A two-dimensional grid of square pixels defines the columns, which for simplicity we assume are along the z axis. Integrating the interpolation for a two-dimensional cross-section along z we find

$$\int_{-\infty}^{\infty} A_{(x,y)} \, \mathrm{d}z \approx \int_{-\infty}^{\infty} \sum_b m_b \frac{A_b}{\rho_b} W\left(|\boldsymbol{r}_b - \boldsymbol{r}_{(x,y)}|, h\right) \mathrm{d}z , \tag{5.5}$$

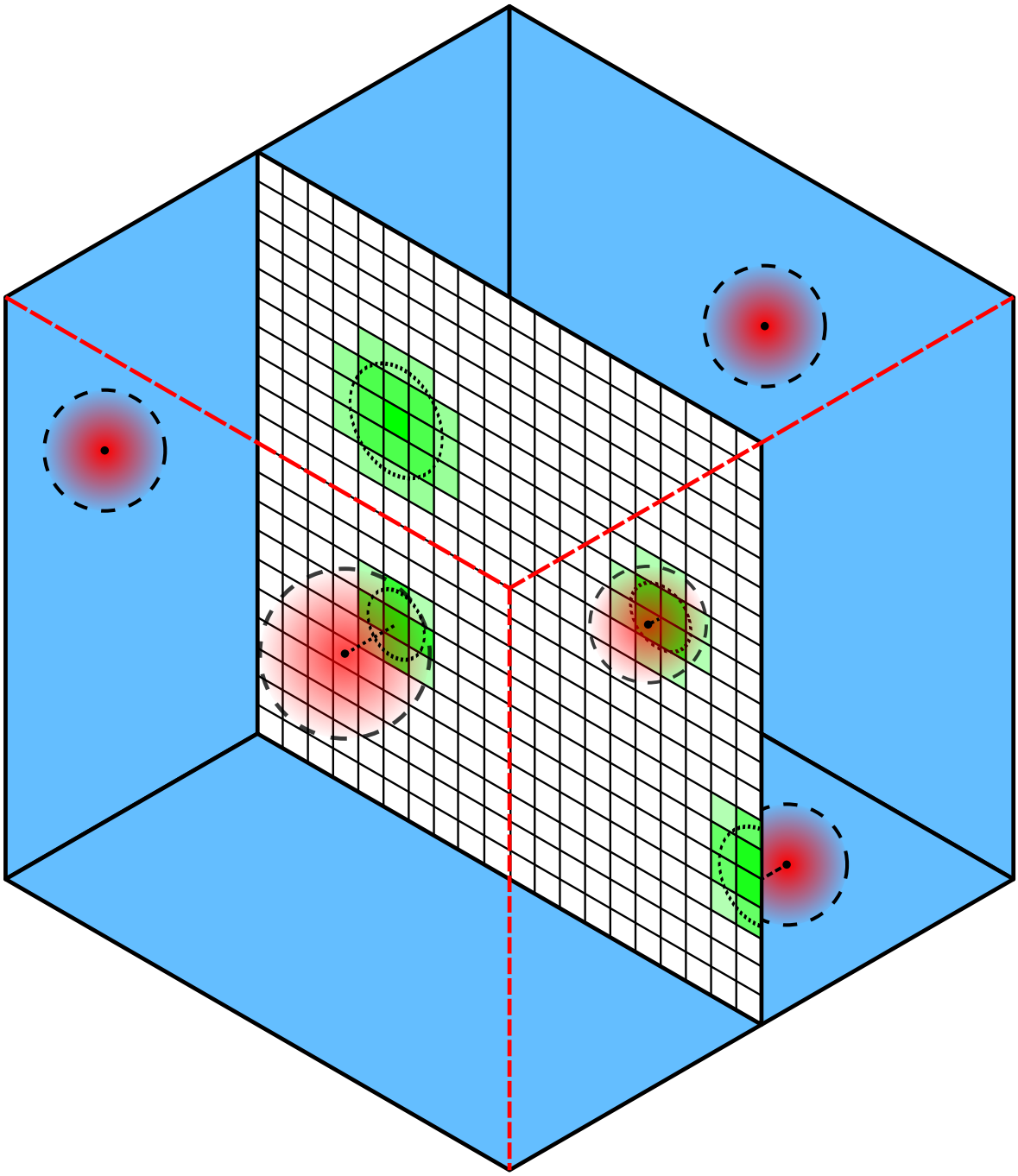


Figure 5.1 – An illustration of a cross-section plot with SPLASH. Five SPH particles are shown as black dots surrounded by their smoothing kernels; one SPH particle behind the cross-section is not plotted. The smoothing kernels are plotted with dashed lines and filled with a red gradient representing the kernel function. The cross-section is shown as a white grid of pixels. SPH particles close to the cross-section contribute to the pixels within the region overlapped by their smoothing kernels. This overlapped region is indicated with a dotted black circle, and the contribution to pixels is indicated in green. The SPH particle hidden behind the cross-section can be seen by its contribution to pixels.

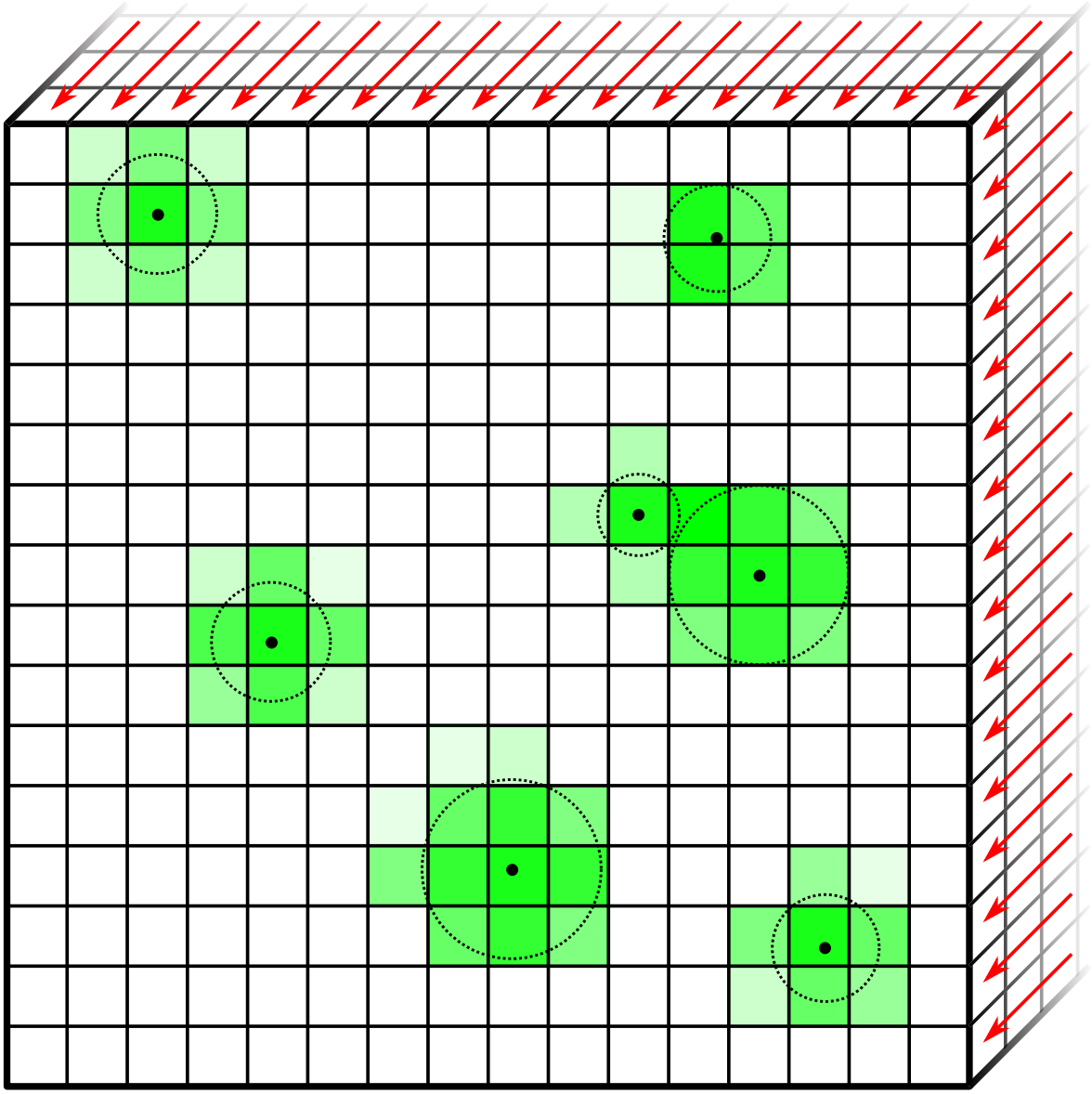


Figure 5.2 – An illustration of a column density plot with SPLASH. The three-dimensional volume of SPH particles is flattened to two dimensions along one axis, as shown by the red arrows. Seven SPH particles are shown as black dots surrounded by their smoothing kernels. The smoothing kernels are plotted with dotted lines. The column density plot is shown as a white grid of pixels. SPH particles contribute to the pixels within the region overlapped by their smoothing kernels, and the contribution to pixels is indicated in green.

where $\boldsymbol{r}_{(x,y)}$ is the perpendicular distance in the xy plane to the nearest point along the centre of the column. We use the same choice of smoothing length as for the cross-section case.

Price (2007) shows that this can be rearranged to

$$\int_{-\infty}^{\infty} A_{(x,y)} \approx \sum_b m_b \frac{A_b}{\rho_b} \int_{-\infty}^{\infty} W\left(|\boldsymbol{r}_b - \boldsymbol{r}_{(x,y)}|, h\right) \mathrm{d}z \,, \tag{5.6}$$

and the integral of the kernel can be replaced by defining the column-integrated kernel

$$Y(\boldsymbol{r}, h) = \int_{-\infty}^{\infty} W(\boldsymbol{r}, h) \, \mathrm{d}z \,, \tag{5.7}$$

which can be calculated in advance. This gives a final summation

$$\int_{-\infty}^{\infty} A_{(x,y)} \approx \sum_b m_b \frac{A_b}{\rho_b} Y\left(|\boldsymbol{r}_b - \boldsymbol{r}_{(x,y)}|, h\right) \,. \tag{5.8}$$

As before this is calculated by looping over particles, rather than grid cells, as this is more efficient.

5.1.3 Grid interpolation

SPLASH allows SPH quantities to be interpolated onto a three-dimensional grid, which is normally centred on the region of interest. We use a cubic grid where the indices x, y and z run from 1 to N, giving N^3 grid cells. We choose to use $N = 256$ for all results presented in this work. If the extent of the grid is from 0 to L along each axis, then the grid spacing is

$$s = \frac{L}{N} \tag{5.9}$$

and each grid cell (x, y, z) covers the coordinates $\big(s(x-1), s(y-1), s(z-1)\big)$ to $\big(sx, sy, sz\big)$.

SPH particles contribute to the grid cells they overlap in a similar way to that for a cross-section, illustrated in Figure 5.1, except that the particles are contributing to a three-dimensional grid of cells instead of a two-dimensional grid of pixels.

The value of some quantity at a grid cell (x, y, z) centred at a position $\boldsymbol{r}_c$ is

$$A_{(x,y,z)} = \sum_b m_b \frac{A_b}{\rho_b} W\left(|\boldsymbol{r}_b - \boldsymbol{r}_c|, h\right) \,, \tag{5.10}$$

where the sum is taken over all particles whose smoothing kernels overlap the grid cell. The quantities A_b, m_b and ρ_b are the value, for particle b, of the quantity A, the mass and the density respectively. W is the standard kernel function described in Section 2.1.1. As

before the smoothing length h is taken as the larger of the particle smoothing length h_b and half the size of a grid cell. When interpolating density to a grid this becomes

$$\rho_{(x,y,z)} = \sum_b m_b W\left(|\boldsymbol{r}_b - \boldsymbol{r}_c|, h\right) . \tag{5.11}$$

5.2 Layer identification

In this thesis we explore the growth of various instabilities in simulations of head-on collisions of approximately symmetric gas flows. Such collisions create a shock-compressed layer that can become unstable to both bending and breathing-mode instabilities. We first define a *collision plane* perpendicular to the collision axis and positioned at the initial interface between the two flows. The layer initially forms along this plane, although the layer can be displaced from the plane by instabilities.

Before an analysis of such instabilities can be started, it is necessary to identify the extent of the layer. We describe how we identify which regions we consider part of the layer, and how we reject some regions of high density which are probably not associated with the layer. Finally we use this information to determine two measures of the thickness of the layer.

Normally instabilities would occur in two dimensions across the layer. If however we are considering an instability which we have excited only along one dimension, we first reduce the three-dimensional grid to two dimensions. If the instability is acting along the y axis, then we average densities across the z axis, such that

$$\rho_{(x,y)} = \frac{\sum_{z=1}^{N} O(z)\rho_{(x,y,z)}}{\sum_{z=1}^{N} O(z)} , \tag{5.12}$$

where $O(z)$ is an occupation factor that determines how much of the cell is part of the simulation. If the entire grid from $z = 1$ to $z = N$ represents part of a simulation, then $O(z) = 1$ for all cells.

In some cases our simulations do not fit across the entire grid. This is the case if one dimension is much smaller in extent than one or both of the other dimensions. The limits of the simulation over the grid cells is defined as z_1 to z_2 in grid units, where z_1 and z_2

are generally not integers. The occupation factor is then given by

$$O(z) = \begin{cases} 0 & \text{if} \quad z < \lceil z_1 \rceil \\ z - z_1 & \text{if} \quad z = \lceil z_1 \rceil \\ 1 & \text{if} \quad \lceil z_1 \rceil < z < \lceil z_2 \rceil \\ z_2 - (z-1) & \text{if} \quad z = \lceil z_2 \rceil \\ 0 & \text{if} \quad z > \lceil z_2 \rceil \end{cases}, \tag{5.13}$$

where $\lceil x \rceil$ indicates the ceiling operator, which returns the smallest integer not less than x.

To simplify notation, we hereafter assume that we have an instability acting in two dimensions. Unless otherwise stated, the equivalent case for an instability acting in one dimension can be found by simply dropping references to z, so that $G_{(x,y,z)}$ becomes $G_{(x,y)}$ for some grid quantity G.

5.2.1 Density threshold

In order to apply our analysis to only the layer, it is necessary to identify the layer. We first apply a simple density criterion to each grid cell. If the density in the grid cell is above a density threshold then the cell is considered a *potential layer cell*. In this work we use a density threshold of five times the background density. If the density is below the critical density then it is not considered further in our analysis. We construct a grid of *mask* cells $M_{(x,y,z)}$, where the value of each mask cell is one if the matching density grid cell is above the density threshold and is otherwise zero. Each cell for which the mask is one is a potential layer cell.

In simulations with velocity fields of supersonic turbulence, regions of gas which are distant from the layer can reach high densities. To filter out some of these regions, we estimate the position and thickness of the layer, and reject dense regions which are distant from the layer.

We consider columns of grid cells through our mask grid along the collision axis. Within each column we find the median position along the x axis of the potential layer cells $X_{(y,z)}$. This gives an first approximation to the position of the centre of the layer at each grid column. We use the median as it is more robust against distant outlying cells which are not part of the layer. If there are no cells above the density threshold in a column, then we set $X_{(y,z)} = 0$ for a collision symmetric about $x = 0$.

We then consider, for each potential layer cell in that column, the absolute distance from the cell to the estimated centre of the layer. We then take the median value of these absolute distances to get an estimated layer half-thickness $D_{y,z}$ for each column. Again, we set this to zero for any columns with no potential layer cells. Finally we take the

median of these half-thicknesses, giving an averaged half-thickness over the whole layer of $\langle D_{\text{layer}} \rangle$. We ensure that this value is greater than the size of two grid cells.

5.2.2 Identifying regions

We now identify each dense *region* of the layer. Cells are considered touching if they have adjacent sides, but not if they touch only through their corners. Each dense region of the layer is defined by a *group* of potential layer cells connected by their sides to themselves and no other potential layer cells. We find such groups by a simple algorithm. We describe this algorithm in two dimensions for simplicity; it is easily extended to three dimensions.

Before beginning the algorithm we initialise a grid, known as the *group grid*, to be of equal size to the mask grid. We set each cell in the group grid equal to zero. The *group number* of a cell is the value of the cell in this group grid. We keep a counter for a *new group number*, which we begin at one. We also keep a list of each cell in each group we create, and update this list as we add cells to a group.

An example of the execution of this algorithm is shown in the subfigures (*a*) to (*f*) of Figure 5.3. We begin at the top left corner of the mask grid, as in subfigure (*a*). We move through the grid by taking single steps to the right. When we reach the end of each row, we descend one row, move to the left edge and continue as before.

At each step we test whether we are in a potential layer cell. If not, then we continue. Otherwise, we find G_{u} and G_{l}, which are the group number of the cell above and the group number of the cell to the left respectively. If these cells are outside the grid boundaries then we use a group number of zero.

- If neither G_{u} nor G_{l} is greater than zero, we begin a new group as this implies that neither the cell to the left nor the cell above is a potential layer cell. We set the group number of our current cell to the *new group number* and increment the new group number by one.
  This is as shown in subfigure (*a*) by the creation of group 1, and then in subfigure (*b*) by the creation of a new group 2.

- If only G_{l} is greater than zero, we set the group number of the current cell to G_{l}. Similarly if only G_{u} is greater than zero, we set the group number of the current cell to G_{l}.
  In subfigure (*b*) after creating the first group 2 cell the next cell also becomes group 2. We then create a new group 1 cell where the potential layer cell falls immediately below the original group 1 cell. In subfigure (*c*) we first create a group 3 cell, noting that touching only at corners does not qualify as connected. We then continue and create a group 4 and a group 5 cell, before beginning to create a line of group 6 cells.

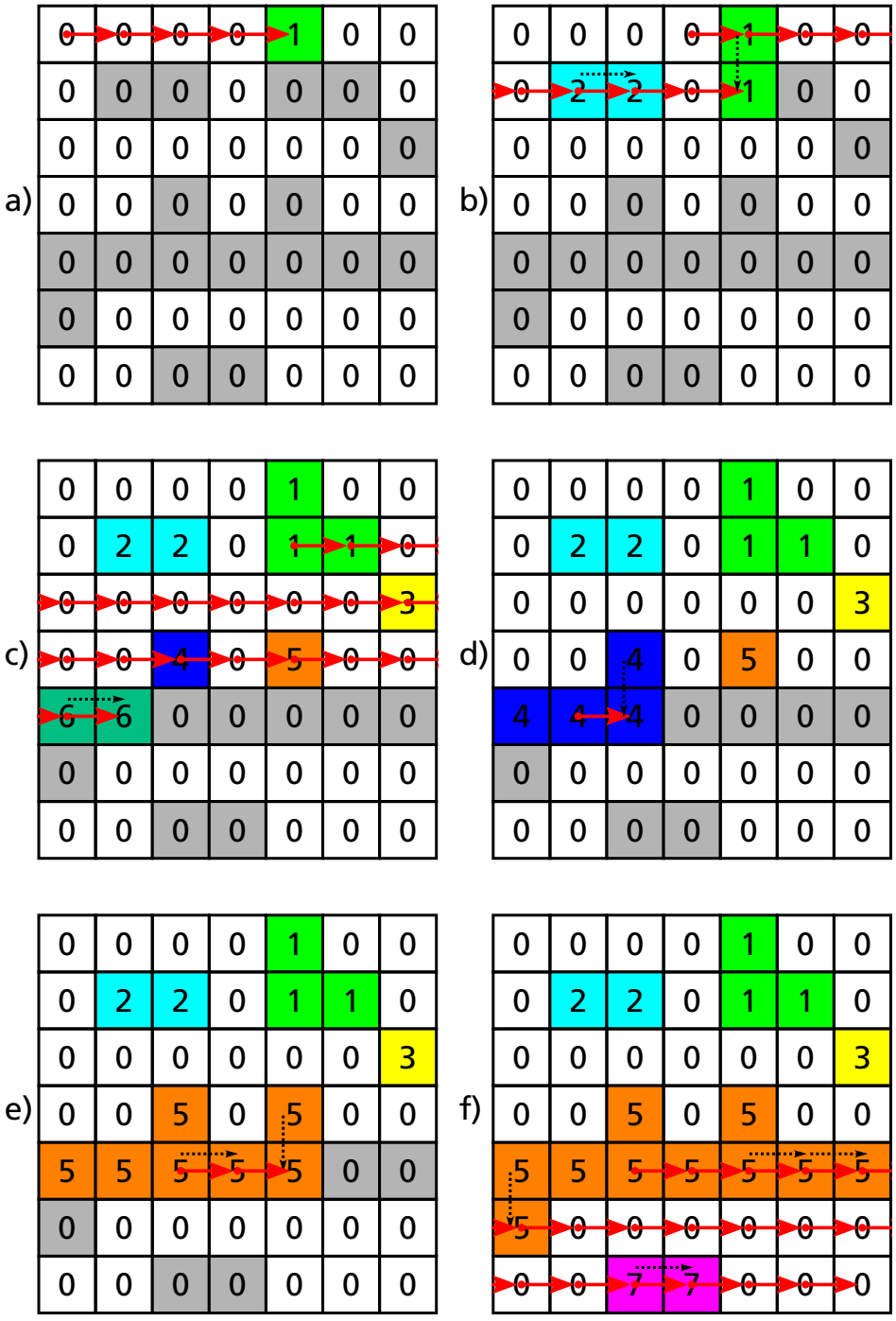


Figure 5.3 – Algorithm for finding connected regions. White grid cells indicate a non-layer grid cell. Grey grid cells indicate a potential layer cell that has not been allocated a group number. The number in each cell is the corresponding group number for that cell. Red arrows indicate the path of the algorithm through the grid. Black dashed arrows indicate where a group number has been taken from another cell. The different grids show different points in the execution of the algorithm.

- If both G_u and G_l are greater than one, then it we set the group number of the current cell to G_u. We move all cells in group G_l to the group G_u. We do this by looping over all cells in group G_l and updating their group numbers on the group grid to G_u. We also add these cells to our list of cells for group G_u.
  In subfigure (*d*) we see a cell which has different groups above and below it. We therefore give it the group number of the cell above, and replace all group 6 cells with group 4 cells. In subfigure (*e*) the process is repeated, except that we replace group 4 cells with group 5 cells. Finally in subfigure (*f*) we complete our walk through the grid, creating one more group in the process.

After walking the entire grid we can repeat the first row and the first column if these are periodic, considering G_l and G_u as wrapped around the periodic boundary. For the first row we only need to consider G_u, corresponding to cells on the bottom row. Similarly for the first column we only need to consider G_l, corresponding to cells on the last column.

Once this is completed we have a list of cells in each group, and can discard the group number grid. Each group has a unique identifier, and does not touch (except at corners) any other group.

5.2.3 Filtering regions

We consider each group in turn, and either accept it as a region of the layer or reject it as unrelated. To do this we find the size of the group along the x axis, and its approximate distance from the layer.

To find an estimate of the distance of the group from the layer, we first find the average position of all the cells in the group, rounded to the nearest grid cell. We then take the $X_{(y,z)}$, the x position of the layer at that position, as estimated in Section 5.2.1. The distance between the group and the layer is then the distance between the averaged position of the group along the x axis and the estimated position of the layer, such that

$$d_x = |G_x - X_{(y,z)}| \,, \tag{5.14}$$

where G_x is the averaged x position of the group, and y and z are the averaged y and z grid positions of the group. The size of the group along the x axis, which we refer to as L_x, is simply the maximum x extent of any group cell minus the minimum x extent of any group cell, and is given in grid units by

$$L_x = x_{\max} - x_{\min} + 1 \,, \tag{5.15}$$

where $x_{\min}$ and $x_{\max}$ are the minimum and maximum x indices of cells in the group.

We then accept the group if

$$\frac{d_x - \langle D_{\text{layer}} \rangle}{L_x} \leq C \ , \tag{5.16}$$

where $\langle D_{\text{layer}} \rangle$ is the estimated half-thickness of the layer defined in Section 5.2.1 and C is a constant. We note that $d_x - \langle D_{\text{layer}} \rangle$ is simply how far the average position of the group is outside the estimated layer thickness, becoming negative for groups with averaged positions inside the layer.

This inequality is automatically true for any group whose distance to the layer is smaller than the estimated half-thickness. For more distant groups, the group is accepted if it is C times larger than the distance between the group and the estimated edge of the layer. In our work we use $C = 1$. Any group which does not fulfil the inequality is not considered part of the layer, and is disregarded in all further analysis. Only cells that are part of groups that are accepted are considered layer cells.

5.3 Instability analysis

In this section we describe techniques used to determine the growth rates of instabilities in dense layers formed by approximately symmetric collisions of gas. Such layers are susceptible to both bending-mode and breathing-mode instabilities. Bending-mode instabilities occur when the regions of the layer become displaced away from the collision plane along the collision axis. Breathing-mode instabilities occur when gas in the layer moves parallel to the collision plane, causing the surface density of the layer to vary from point to point.

To analyse these instabilities, and their growth rates, it is necessary to define tracers for the instability that measure the amount of layer displacement or surface density change. In both cases we begin with a three-dimensional grid of gas density, created as described in Section 5.1.3. For an instability in one dimension we reduce our three-dimensional grid to a two-dimensional grid using equation 5.12. We analyse only regions considered to be part of the layer, as described in Section 5.2. We then use our tracers to measure, on a one-dimensional or two-dimensional grid, the strength of the instability at each point on the grid. Finally we can use Fourier analysis to examine the amplitude of different modes, and the rate of growth of each mode as a function of wavenumber.

5.3.1 Bending-mode instabilities

In a bending-mode instability we wish to measure the displacement of the centre of mass of a layer from the collision plane as a function of position, as illustrated in Figure 5.4. We assume that gas is colliding along the x axis, and that the colliding gas is approximately symmetric such that the initial centre of mass along the x axis at any position is at $x = 0$. The collision plane is therefore along the y–z plane at $x = 0$.

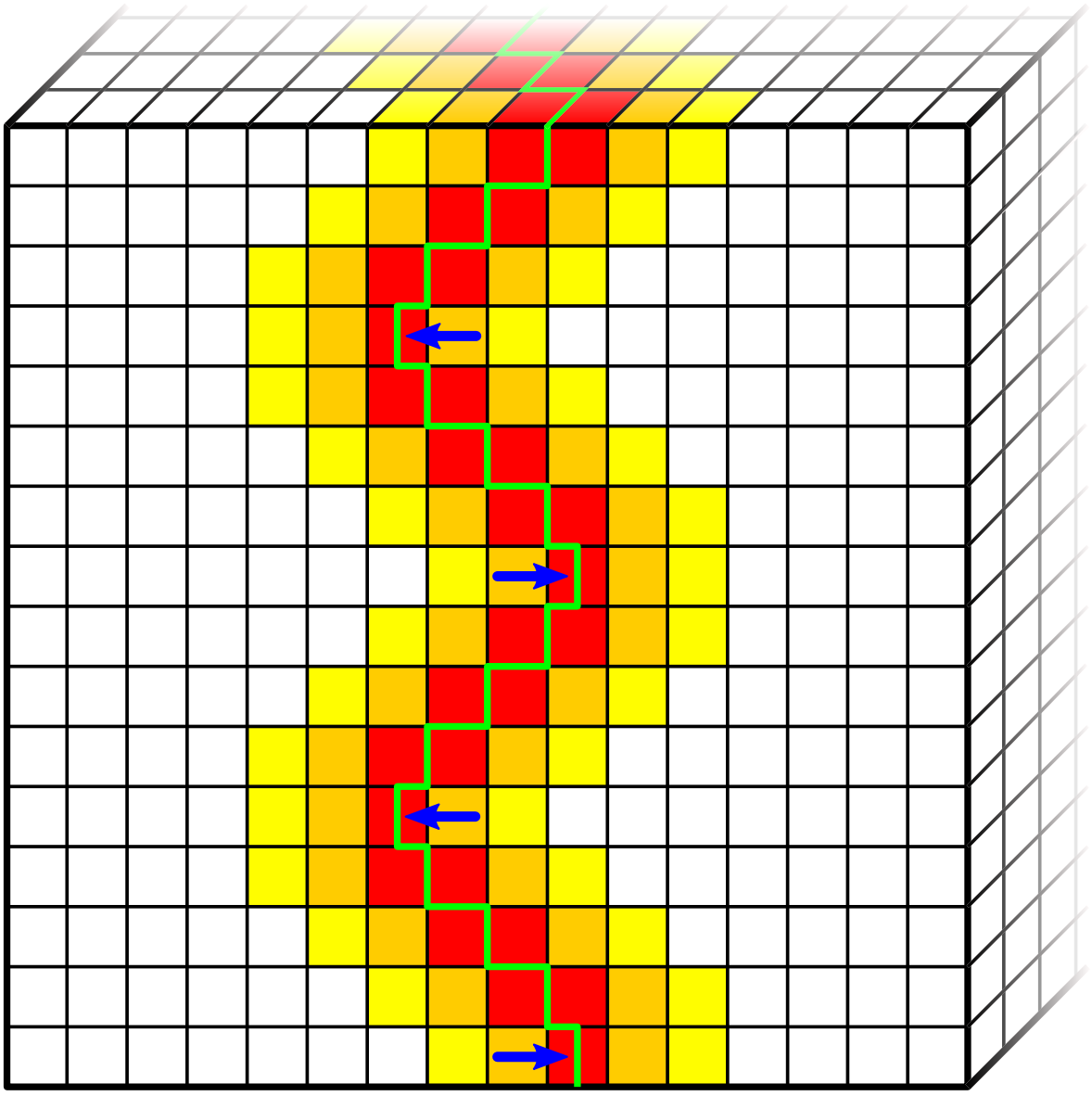


Figure 5.4 – Analysis of a bending-mode instability. One slice through the three-dimensional grid of density is shown. Yellow cells are the least dense while red cells are the most dense grid cells. White cells are cells not considered to be part of the layer and are ignored. Gas is inflowing from the left and right, creating a dense layer. The centre of mass along each horizontal column is calculated, and shown as the position of the green line. This line traces the displacement of the centre of mass of the layer from the original midpoint, indicated using blue arrows. Using every slice a two-dimensional displacement surface is defined, of which the green line is just a cross-section. This displacement surface is the two-dimensional tracer of the bending-mode instability.

We define columns along the collision axis corresponding to grid cells on the collision plane. We include only cells that have been identified as part of the layer, as defined in Section 5.2, in our analysis. As the bending-mode instability is excited, the centre of mass for each column will change. We use the position of the centre of mass along columns to find the approximate position of the layer. When analysing an instability in two dimensions we define a grid of layer displacements

$$D_{(y,z)} = \frac{\sum_{x=1}^{N} r_x \rho_{(x,y,z)}}{\sum_{x=1}^{N} \rho_{(x,y,z)}} , \tag{5.17}$$

where r_x is the x position at the centre of the grid cell, or for an instability acting along only the y axis

$$D_{(y)} = \frac{\sum_{x=1}^{N} r_x \rho_{(x,y)}}{\sum_{x=1}^{N} \rho_{(x,y)}} . \tag{5.18}$$

For any column where no cells are part of the layer, we assume the centre of mass is zero. This grid is our tracer of the bending-mode instability, which we refer to as the centre-of-mass position along the collision axis.

5.3.2 Breathing-mode instabilities

In a breathing-mode instability we wish to measure the displacement of gas within a layer parallel to the collision plane, as illustrated in Figure 5.5. We therefore need to measure the surface density across the layer, as a breathing-mode instability creates denser and less dense regions in the layer. We again assume the gas is colliding along the x axis, and so the collision plane is parallel to the y–z plane.

For an instability in two-dimensions, we find

$$C_{(y,z)} = \sum_{x=1}^{N} \rho_{(x,y,z)} , \tag{5.19}$$

or for an instability acting along only the y axis,

$$C_{(y)} = \sum_{x=1}^{N} \rho_{(x,y)} . \tag{5.20}$$

This grid of column densities represents our tracer of surface density variations, caused by a combination of breathing-mode instabilities and a constant layer accretion term due

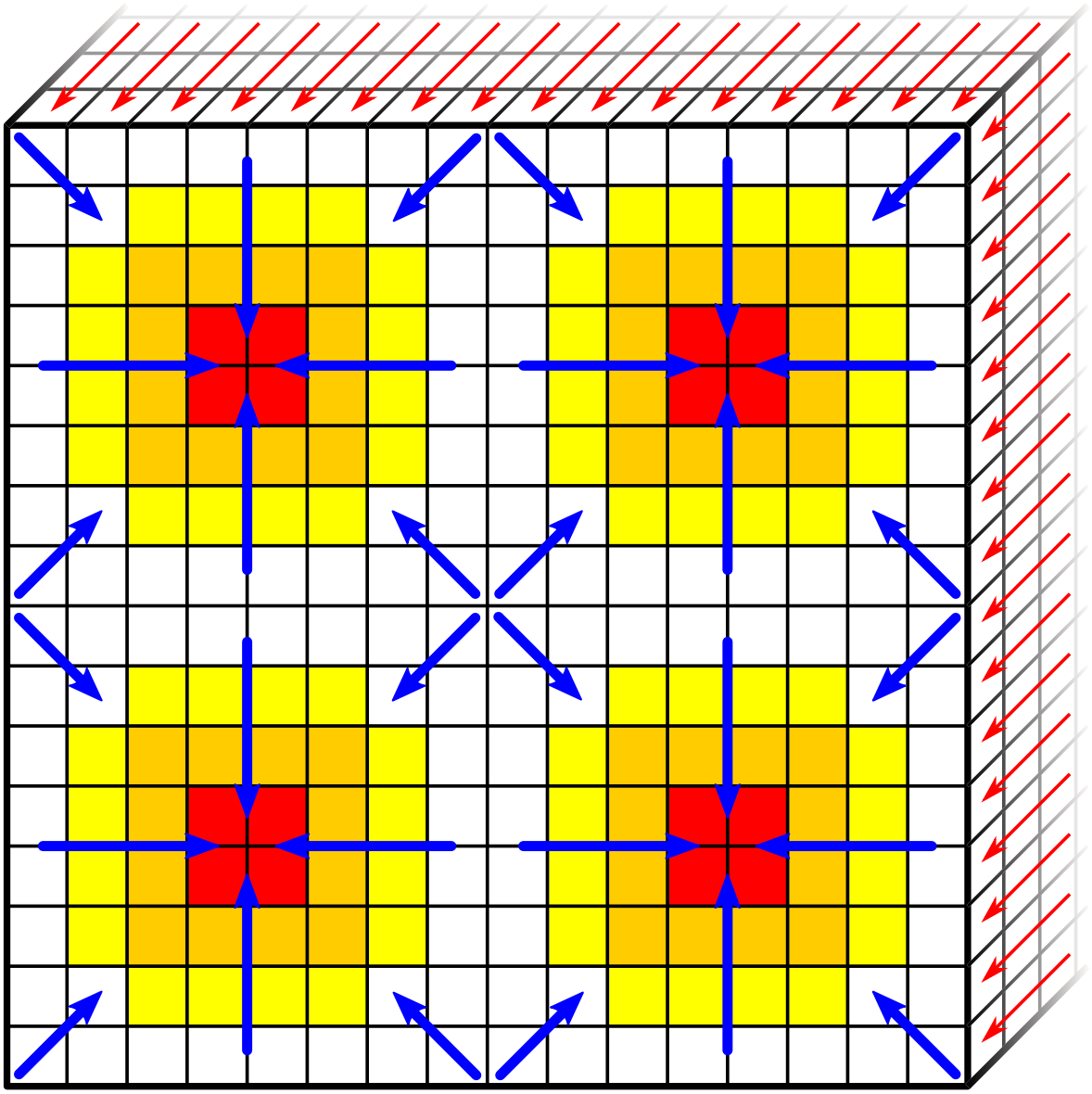


Figure 5.5 – Analysis of a breathing-mode instability. The three-dimensional volume of grid cells of density, viewed along the collision axis, is flattened to two dimensions along one axis by summation along the columns. This summation is shown by the red arrows. Yellow cells have the lowest surface density while red cells have the highest surface density. White cells are cells that are not considered to be part of the layer and are ignored. Gas in the layer is displaced parallel to the layer by the breathing-mode instability, as indicated using blue arrows, creating variations in the surface density. The surface density is therefore the tracer of the breathing-mode instability.

to the inflowing gas.

5.3.3 Fourier transform

A Fourier transform can now be performed on our grid of either displacements D or column densities C. The two-dimensional Fourier transform is given by

$$\tilde{F}_{(k_y,k_z)} = \int\int D_{(y,z)} \mathrm{e}^{-2\pi i(k_y y + k_z z)} \,\mathrm{d}x\,\mathrm{d}y \;, \tag{5.21}$$

where $\tilde{F}$ is the complex Fourier field and $\boldsymbol{k} = (k_y, k_z)$ is the wave vector. This grid has the same size as the density grid. We define a half-grid limit N_{half} as

$$N_{\mathrm{half}} = \left\lfloor \frac{N}{2} \right\rfloor \;, \tag{5.22}$$

where the floor operator $\lfloor x \rfloor$ returns the largest integer not greater than x. When N is even, k_y and k_z run from $-N_{\mathrm{half}} + 1$ to N_{half}. When N is odd, k_y and k_z run from $-N_{\mathrm{half}}$ to N_{half}. The $k = 0$ grid cell in each direction is then found at the grid cell index

$$c = \left\lfloor \frac{N+1}{2} \right\rfloor \;. \tag{5.23}$$

The Fourier transform can be equivalently taken for the surface density grid C, and can also be taken in one dimension for a one-dimensional grid of displacements or column densities. For efficiency we do not use equation 5.21 directly. Instead we use a Fast Fourier Transform provided by the Python module Numpy. When calculating the Fourier transform in this way the wavelength to wavenumber relation is defined so that

$$\lambda = \frac{1}{|\boldsymbol{k}|} \;, \tag{5.24}$$

where $\boldsymbol{k}$ is the wavenumber vector returned by the Fourier transform and λ is the associated wavelength.

We are not interested in the phase information in the Fourier field, and so we use only the Fourier amplitude A, given by

$$A = \sqrt{\tilde{F}\;\tilde{F}^*} \;, \tag{5.25}$$

for each value of the Fourier field.

If we have excited a perturbation at a particular wavenumber, and we are only interested in the amplitude at this wavenumber, then this can be obtained directly from either a one-dimensional or two-dimensional Fourier field. If we are interested in the Fourier amplitude as a function of the magnitude of the wavenumber k, then for a one-

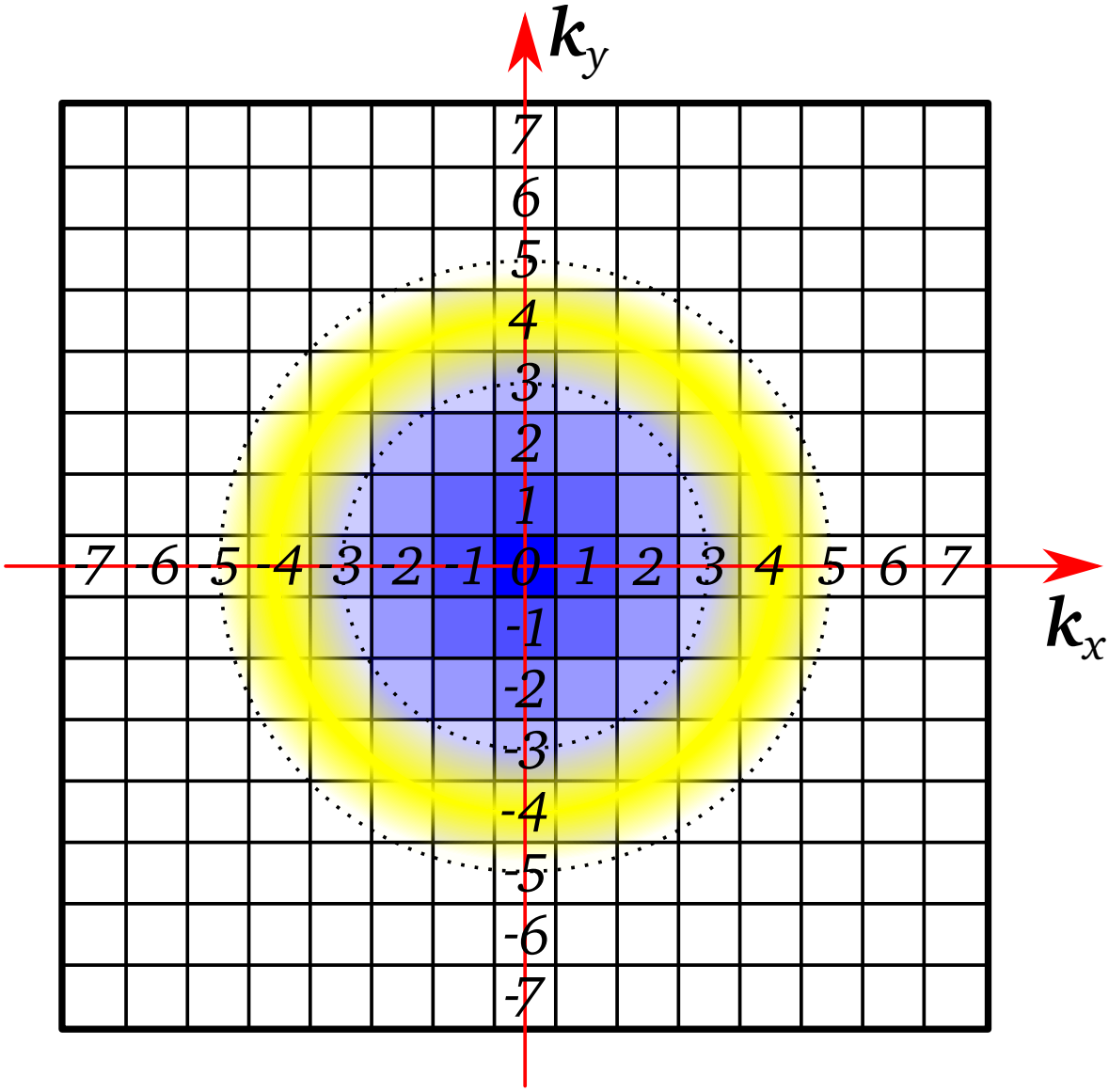


Figure 5.6 – Circular averaging of a two-dimensional Fourier spectrum. The value of the Fourier amplitude at each vector $\boldsymbol{k}$ is indicated in blue, with darker being a larger value and lighter being a lower value. The circular wavelet function used for circular averaging is shown in yellow and centred on $k = 4$. The wavelet function is moved from $k = 0$ to the maximum value of k to generate the continuously averaged Fourier spectrum $A(\boldsymbol{k})$.

dimensional Fourier transform we already have $A(k)$. This is defined at integer values of the wavenumber k. If however we have taken a two-dimensional Fourier transform then we need to take a circular average.

We convert the integer grid of wavevectors $\boldsymbol{k}$ into a continuous one-dimensional function $A(|\boldsymbol{k}|)$ by calculating a weighted sum at a large number of fractional values of $|\boldsymbol{k}|$. The weighted sum is defined such that

$$A(|\boldsymbol{k}|) = \frac{\sum W(\boldsymbol{k}_{ij} - \boldsymbol{k})\, A_{ij}}{\sum W(\boldsymbol{k}_{ij} - \boldsymbol{k})} \tag{5.26}$$

where the sums are over all wavevectors $\boldsymbol{k}_{ij}$ with amplitude A_{ij}. The weighting function $W(k)$ is defined such that

$$W(\boldsymbol{k}) = \begin{cases} 1 - |\boldsymbol{k}| & \text{if } |\boldsymbol{k}| \leq 1 \\ 0 & \text{if } |\boldsymbol{k}| > 1 \end{cases} . \tag{5.27}$$

This method samples the two-dimensional grid by moving a circular wavelet function of width one and a triangular profile outwards from $|\boldsymbol{k}| = 0$. Figure 5.6 shows this wavelet function at $k = 4$. The resultant Fourier spectrum $A(|\boldsymbol{k}|)$ indicates the Fourier amplitude as a function of wavenumber. We disregard the Fourier spectrum at values of $k < 1$, as this includes a contribution from the $k = 0$ component. This component is not usually interesting when analysing perturbations.

In this section we have defined our FFT in terms of dimensionless wavenumbers. These are wavenumbers in terms of the box size over which we perform the FFT. When analysing data we present wavenumbers as 'physical' wavenumbers, with units of inverse length, to remove this dependence on box size. These are found by dividing the dimensionless wavenumbers by the box size. When using physical wavenumbers, amplitudes are still defined only at discrete wavenumbers. However, those wavenumbers are no longer defined at all integers, but could be defined at regular spacing.

5.3.4 Rates of growth

To convert Fourier spectra into rates of growth, we take the Fourier spectrum $A(k)$ at a set of times i. We then use the central differential method to find the rates of growth $\omega(k)$ at time i as a function of wavenumber k. This is

$$\omega(k)_i = \frac{A(k)_{i+1} - A(k)_{i-1}}{2\Delta t} , \tag{5.28}$$

where Δt is the time interval between timesteps.

The rate of growth ω for the Fourier spectrum derived from the centre-of-mass position of the layer is the rate at which perturbations grow in size, and is a velocity. For the Fourier spectrum derived from the surface density of the layer, ω is the rate at which surface density perturbations grow, and has units of surface density over time. These are not 'growth rates' in the usual sense of being the reciprocal of a timescale. However they can be converted to such a growth rate for a particular perturbation of amplitude η by

$$\tau^{-1} = \frac{\omega}{\eta} , \tag{5.29}$$

where τ^{-1} is the growth rate.

5.3.5 Layer width and thickness

It is useful to have a measure of the size of the layer. We consider two measures which we call the *thickness* and the *width*. The thickness does not include voids within the layer, whereas the width is a measure of the maximum extent of the layer.

To find the layer thickness, we begin at one end of a column along the collision axis. We walk through our density grid, noting whether a cell is inside the layer. A cell is inside the layer if it passes the criterion described in Section 5.2.3. This also implies it will have a density above the density threshold.

As we transition from a cell that is not in the layer to a cell that is in the layer, we interpolate between densities to find the estimated edge of the layer. In such a case we will always be moving from a cell which is below the density threshold to a cell that is above it or vice versa. This is only true because we only accept or reject complete groups which are completely surrounded by cells below the density threshold. The position of the estimated edge between a cell (x,y,z) and a cell (x+1,y,z) is given by

$$x_{\text{edge}} = x + \frac{\rho_t - \rho_{(x,y,z)}}{\rho_{(x+1,y,z)} - \rho_{(x,y,z)}} \,. \tag{5.30}$$

As we move through the column, we keep track of the total accumulated length in each section of the layer, interpolating in and out of the layer. The total accumulated length is the *thickness*.

Our alternative measurement is the *width*, where we consider, in each column, only the minimum and maximum extent of the layer along the collision axis. The width is therefore at least as great as the thickness, and if the layer includes voids or folds the width can be considerably greater. If outlying regions have not been effectively suppressed, the width can be dramatically overestimated. However the width can be more useful if we simply wish to know how bloated a layer has become.

At large bending angles, these measurements are overestimates as they measure the length through the layer parallel to the collision axis. The true thickness, perpendicular to the layer, may be much smaller. We use the layer position $D_{(y,z)}$ or $D_{(y)}$ determined in Section 5.3.1 to estimate at each point across the layer the true width and thickness.

At each point we calculate the spatial derivative of the layer position. In one dimension we use the one-dimensional five-point stencil

$$\frac{\mathrm{d}D_{(y)}}{\mathrm{d}y} = \frac{D_{(y-2)} - 8\,D_{(y-1)} + 8\,D_{(y+1)} - D_{(y+2)}}{12\,\Delta y} \,, \tag{5.31}$$

where Δ is the grid spacing in the y direction, and we wrap appropriately around periodic boundaries. In two dimensions we calculate the vector gradient of the layer using the

same one-dimensional five point in two dimensions, such that

$$\nabla D_{(y,z)} = \frac{D_{(y-2,z)} - 8\,D_{(y-1,z)} + 8\,D_{(y+1,z)} - D_{(y+2,z)}}{12\,\Delta y}\,\boldsymbol{i} + \frac{D_{(y,z-2)} - 8\,D_{(y,z-1)} + 8\,D_{(y,z+1)} - D_{(y,z+2)}}{12\,\Delta z}\,\boldsymbol{j}\,, \tag{5.32}$$

where Δz is the grid spacing in z direction. For our simulations the grid spacing is the same across all dimensions. We then take the magnitude of the vector gradient to give the gradient in the direction of steepest ascent.

The bending angle ϕ is then simply $\tan^{-1}(m)$ where m is the gradient, and for a thickness or width l we can estimate a corrected length across the layer

$$l_{\text{corrected}} = l\cos\left(\phi\right)\,. \tag{5.33}$$

We use this corrected length wherever we quote a width or thickness. Once the thickness and width have been calculated for each column, we calculate the mean values for the entire layer. Normally we collide gas flows which extend across the entire domain, and these averages are calculated across all cells. However, if we are colliding two spheres only a fraction of the grid will be involved in the collision. We assume the spheres are of radius r_{s} and are initially at $x = \pm r_{\text{s}}$, so that they are just touching, and are centred on and move along the $y = z = 0$ line. We define this line as the collision line, and for each cell find r, the perpendicular distance to this line. We then only include cells in any average if r is less than a critical value r_{max}. This critical value is given by

$$r_{\text{max}} = \sqrt{V_{\text{E}}t\left(2r_{\text{s}} - V_{\text{E}}t\right)}\,, \tag{5.34}$$

where V_{E} is the collision velocity and t is the elapsed time since the start of the collision. This is simply the radius of the circle formed by the intersection of the collision plane and the sphere as it moves into the collision. Outside of this circle, no layer will have formed, and the width and thickness should be ignored and not included in the average.

5.4 Sink properties

In our simulations we use sink particles, which are described in Section 3.7. Each sink represents a protostar which accretes mass from its surroundings through its gravitational attraction. We use the position, velocity and mass of each sink to identify some statistical properties of sinks for each star-forming simulations. These quantities are output from our simulations at regular time intervals. Certain properties, such as the total mass in sinks with respect to time, are easily calculated and do not require detailed explanation. The only analysis we perform which does requires explanation is the construction of a

sink mass function, which we describe below.

5.4.1 Mass function

We begin with a large number of sink masses. We construct a histogram of these sink masses, where we attempt to place an equal number of sinks into each bin. If the total number of sinks is not a multiple of the number of sinks per bin, we only partially fill the first bin. By placing equal numbers of sinks into each bin, the binning error is the same for all bins, and is the square root of the number of sinks per bin. We construct our mass function in logarithmic space.

We begin by calculating the number of bins. If we have $N_{\rm sinks}$ sinks in total, we use a number of bins given by

$$B = \left\lceil \frac{N_{\rm sinks}}{20} \right\rceil \, , \tag{5.35}$$

where the brackets indicate the ceiling function, which returns the lowest integer greater than its argument. Each bin except the first will contain 20 sinks. The first bin may contain fewer sinks, given by

$$p_{\rm first} = N_{\rm sinks} - p(B-1) \, , \tag{5.36}$$

where p is the number of sinks per completely-occupied bin. The list of sink masses is then sorted in order of increasing mass. Sinks are placed into each bin in order, with $p_{\rm first}$ sinks in the first bin and p sinks in each subsequent bin. We then need to find the mass ranges of each bin.

We define the logarithmic boundaries of a bin i as x_{i-1} and x_i. The first bin therefore has boundaries x_0 and x_1, the second bin has boundaries x_1 and x_2, and so on. The boundary between bins i and $i+1$ is given by

$$x_i = \frac{\log_{10}\left(m_{\rm i,last}\right) + \log_{10}\left(m_{\rm i+1,first}\right)}{2} \, , \tag{5.37}$$

where $m_{\rm i,last}$ is the mass of the most massive sink in bin i, and $m_{\rm i+1,first}$ is the mass of the least massive sink in bin $i+1$. We deal with the first and last boundary separately. The first boundary is given by

$$x_0 = \min\left[\log_{10}\left(m_{\rm min}\right), 2\log_{10}\left(\overline{m}_0\right) - x_1\right] \, , \tag{5.38}$$

where $\min(a,b)$ returns its smallest argument, $m_{\rm min}$ is the mass of the least massive sink and $\overline{m}_0$ is the mean mass of sinks in the first bin. The last boundary is given by

$$x_B = \max\left[\log_{10}\left(m_{\rm max}\right), 2\log_{10}\left(\overline{m}_B\right) - x_{B-1}\right] \, , \tag{5.39}$$

where $\max(a,b)$ returns its largest argument, $m_{\rm max}$ is the mass of the most massive sink

and $\overline{m}_B$ is the mean mass of sinks in the last bin. The value of bin i is then given by

$$H_i = \frac{p_i}{N_{\text{sinks}}(x_i - x_{i-1})} \tag{5.40}$$

where p_i is the number of sinks in bin i. The logarithmic bin boundaries can be converted to boundaries in linear mass space by

$$M_i = 10^{x_i} \,, \tag{5.41}$$

for each boundary x_i. The histogram is then constructed by plotting H_i for each bin i, which covers the mass interval M_{i-1} to M_i.

We also construct a smooth, continuous version of a histogram using the superposition of Gaussian functions. This smoothing is entirely arbitrary, and is performed only to make noisy histograms clearer. We again sort sinks into order of increasing mass. For each sink i of mass m_i and logarithmic mass $\log_{10}(m_i)$, we determine the local spacing of logarithmic sink masses around its logarithmic sink mass. We then construct a Gaussian for each sink of the form

$$f_i(x) = \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left[\frac{-(x-m_i)^2}{2\sigma_i^2}\right] \,, \tag{5.42}$$

where σ_i is the width of the Gaussian for sink i. This is given by the sum in quadrature of a global width and a sink width. The global width σ_g is given by the average mass spacing between sinks over a range from sink $\left\lfloor\sqrt{N}\right\rfloor$ to $\left\lceil N - \sqrt{N}\right\rceil$. This removes the effect of outliers to give a global spacing. The global spacing is therefore

$$\sigma_g = \frac{m_{\left\lceil N-\sqrt{N}\right\rceil} - m_{\left\lfloor\sqrt{N}\right\rfloor}}{\left\lceil N - \sqrt{N}\right\rceil - \left\lfloor\sqrt{N}\right\rfloor} \,. \tag{5.43}$$

For each sink i, the local spacing is calculated over the 20 sinks from sink $i-10$ to $i+10$. Where the sink i is one of the ten least massive or ten most massive sinks, and therefore some of these neighbours do not exist, then some of the sinks that do exist are included twice, but with their masses reflected to the other side of the mass m_i. The local spacing is given by

$$\sigma_{i,\text{local}} = \sqrt{\frac{1}{20}\sum_{j=-10}^{10}\left(\frac{m_{i+j} - m_i}{|j|}\right)^2} \,, \tag{5.44}$$

where the sum does not include $i = j$. The combined spacing is then given by

$$\sigma_i = \varepsilon\sqrt{\sigma_{i,\text{local}}^2 + \sigma_g^2} \,, \tag{5.45}$$

where ε is a numerical factor that determines the amount of smoothing. We set ε equal to

20 for our mass functions.

Using equation 5.45 with equation 5.42, we obtain our smooth curve from the normalized superposition of the smoothed curve for each sink, given by

$$f(x) = \frac{1}{N_{\text{sinks}}} \sum_{i=1}^{N_{\text{sinks}}} f_i(x) \,. \tag{5.46}$$

Chapter 6

Gravitational instability

Gravitational instability is simply the growth of structure under self-gravity. If gravitational forces exceed the combination of any opposing thermal, rotational, turbulent and magnetic support, then gas will tend to contract towards the centre of gravity. If this contraction is relatively unimpeded then it can be described as gravitational collapse. Gas undergoing gravitational collapse will increase in density until the pressure in the gas increases, slowing the collapse. Gravitational collapse is an important process in star formation as all star formation begins with the collapse of low-density gas.

Masunaga & Inutsuka (2000) studied the gravitational collapse of a $1\,\mathrm{M}_\odot$ cloud. At the start of the collapse, the cloud cools to reach a minimum temperature of 5 K. At low densities of less than $10^{-14}\,\mathrm{g\,cm^{-3}}$ the cloud is still optically thin, is able to radiate away heat efficiently, and so remains cold and approximately isothermal at around 5 K to 20 K. As a result the cloud is in approximately freefall collapse unimpeded by gas pressure.

As the density increases to the opacity limit at approximately $10^{-12}\,\mathrm{g\,cm^{-3}}$ the gas becomes optically thick, and the temperature in the cloud begins to rise. The equation of state of the gas becomes approximately adiabatic, with an exponent of ~ 1.4. This slows the collapse of the cloud, as radiation is no longer able to escape efficiently, and the *first core* is formed. This first core slowly undergoes Kelvin–Helmholtz contraction until it reaches a density of approximately $10^{-7}\,\mathrm{g\,cm^{-3}}$ and at a temperature of approximately 2000 K. This causes molecular hydrogen to dissociate, lowering the adiabatic exponent from ~ 1.4 to ~ 1.1. The first core rapidly collapses to a new core density of $10^{-2}\,\mathrm{g\,cm^{-3}}$ and a temperature of 10^4 K. While the first core survives for approximately 10^5 years, its eventual collapse takes only months. This collapse forms the *second core*, which is the hydrostatically supported protostar. This will slowly cool and contract into a main sequence star.

6.1 Jeans instability

The Jeans instability was first investigated by Jeans (1902). This instability is the tendency of isothermal, uniform density gas to gravitationally fragment if slightly perturbed. The

critical mass scale in this instability is the Jeans mass. We can estimate the Jeans mass by considering a spherically symmetric volume of gas. If this volume has radius R, mass M and sound speed c_s, then the approximate crossing time of sound waves is simply

$$t_{\rm cross} \sim \frac{R}{c_s} \,. \tag{6.1}$$

Sound waves will act to smooth out any density perturbations across the volume provided that the perturbations grow slower than this crossing time. This implies that the volume will not collapse gravitationally unless the freefall timescale is shorter than the crossing time. The freefall timescale of a uniform density cloud is

$$t_{\rm freefall} \sim \frac{1}{\sqrt{G\rho}} \,. \tag{6.2}$$

The condition for collapse is therefore

$$t_{\rm freefall} \lesssim t_{\rm cross} \,, \tag{6.3}$$

or

$$\frac{1}{\sqrt{G\rho}} \lesssim \frac{R}{c_s} \,. \tag{6.4}$$

When these timescales are equal, the sphere is marginally unstable. Larger objects will collapse gravitationally, while smaller objects are stabilised by sound waves and do not collapse. Rearranging gives the *Jeans length*

$$R_{\rm J} \sim \frac{c_s}{\sqrt{G\rho}} \,. \tag{6.5}$$

For a medium of approximately uniform density, this defines the *Jeans mass*

$$M_{\rm J} \sim \frac{c_s^3}{\sqrt{G^3\rho}} \,. \tag{6.6}$$

Appendix A describes other methods to derive a Jeans length, and thus a Jeans mass, which give more accurate results. We choose to use a Jeans radius

$$R_{\rm J} = \sqrt{\frac{15}{4\pi\rho G}} c_s \,, \tag{6.7}$$

and a Jeans mass

$$M_{\rm J} = \sqrt{\frac{375}{4\pi G^3\rho}} c_s^3 \,. \tag{6.8}$$

6.2 Gravitational instability in a layer

In star formation the most commonly encountered form of the gravitational instability in isothermal gas is the three-dimensional spherically symmetric Jeans instability. However the gravitational instability can take different forms in different geometries. In a two-dimensional layer of gas with a uniform surface density there is a similar two-dimensional gravitational instability.

The three-dimensional Jeans instability has no preferred length scale that collapses faster than all other scales. Although the Jeans length gives the smallest scale that is gravitationally unstable, all larger scales are also unstable. These larger scales collapse faster than the smaller scales. The result is that any initially static and spherically symmetric distribution of gas larger than the Jeans radius will collapse monolithically towards the centre of mass.

However the gravitational instability in a layer does have a preferred length scale. Scales larger or smaller than this preferred scale collapse more slowly, creating a pattern of fragmentation at the preferred scale. This pattern may leave an imprint in any later star formation.

6.2.1 Layer formed by colliding flows

In most cases the layer is formed by the collision of supersonic flows and confined by ram pressure. We assume that the ram pressure dominates over the thermal pressure in the inflowing gas. We define a Mach number,

$$\mathscr{M} = \frac{v_{\rm ext}}{c_{\rm s}} \,, \tag{6.9}$$

where $v_{\rm ext}$ is the velocity of the inflowing gas. For pressure balance we have

$$\rho_{\rm layer} c_{\rm s}^2 = \rho_{\rm ext} v_{\rm ext}^2 \,, \tag{6.10}$$

where $\rho_{\rm layer}$ is the gas density in the layer, $c_{\rm s}$ is the sound speed in the layer, and we use the ideal gas law $P = \rho c_{\rm s}^2$. Combining this with equation 6.9 gives

$$\rho_{\rm layer} = \rho_{\rm ext} \mathscr{M}^2 \,. \tag{6.11}$$

The surface density of the layer grows with time as it accretes inflowing gas on both sides, giving

$$\Sigma = 2\rho_{\rm ext} v_{\rm ext} t \,, \tag{6.12}$$

where t is the age of the layer. Using equations 6.9 and 6.11 this becomes

$$\Sigma = 2\rho_{\rm ext} c_{\rm s} \mathscr{M} t \,. \tag{6.13}$$

6.2.2 Jeans length in a layer

As in section 6.1, we can calculate the Jeans length by equating the crossing time with the gravitational timescale. We consider a region of the layer of linear size L. The sound crossing time for the layer is

$$t_{\rm cross} \sim \frac{L}{c_{\rm s}} \,. \tag{6.14}$$

The freefall timescale for gravitational collapse of this region parallel to the layer is then

$$t_{\rm ff} \sim \sqrt{\frac{L^3}{GM}} \,. \tag{6.15}$$

The mass of a region of size L is approximately $M \sim L^2\Sigma$, giving

$$t_{\rm ff} \sim \sqrt{\frac{L}{G\Sigma}} \,. \tag{6.16}$$

Equating these two timescales gives

$$\frac{L}{c_{\rm s}} \sim \sqrt{\frac{L}{G\Sigma}} \,, \tag{6.17}$$

and rearranging yields the Jeans length

$$L_{\rm J} \sim \frac{c_{\rm s}^2}{G\Sigma} \,, \tag{6.18}$$

and the Jeans mass

$$M_{\rm J} \sim \frac{c_{\rm s}^4}{G^2\Sigma} \,. \tag{6.19}$$

Vishniac (1994) uses a detailed linear model of a shock-confined slab, and derives an instantaneous growth rate of the gravitational instability

$$\omega^2 \sim 4\pi^2 c_{\rm s}^2 k^2 - 4\pi^2 G k \Sigma \,, \tag{6.20}$$

where k is the wavenumber of the perturbation. The numerical coefficients in equation 6.20 differ slightly from those of Vishniac (1994), as we define

$$k = \frac{1}{\lambda} \,, \tag{6.21}$$

where λ is the wavelength, instead of the angular wavenumber

$$k_{\rm angular} = \frac{2\pi}{\lambda} \,. \tag{6.22}$$

We similarly adjust other equations involving wavenumbers throughout this chapter.

This relation is similar to the growth rate for the three-dimensional Jeans instability, but with a different Jeans length. When ω^2 is negative this corresponds to expanding or collapsing solutions. The critical wavenumber is when $\omega^2 = 0$, and so

$$k_{\rm J} \sim \frac{G\Sigma}{c_{\rm s}^2} \,. \tag{6.23}$$

This gives a critical wavelength, which is the Jeans wavelength, of

$$\lambda_{\rm J} \sim \frac{c_{\rm s}^2}{G\Sigma} \,, \tag{6.24}$$

which is identical to the Jeans length $L_{\rm J}$ we have just derived by considering the sound crossing time across the layer. We can also substitute equation 6.13 to obtain the critical wavenumber in terms of time and Mach number, which is

$$k_{\rm J} \sim \frac{2G\rho_{\rm ext}\mathscr{M}\,t}{c_{\rm s}} \,. \tag{6.25}$$

6.2.3 Fastest growing mode by considering forces

In order to collapse a perturbation needs to be both Jeans-unstable and to have had sufficient time to grow. At each instant, only wavelengths longer than the Jeans length are unstable against gravitational collapse. As the surface density of the layer increases with time the Jeans length decreases, allowing smaller modes to become unstable.

Equation 6.20 shows that although larger modes become unstable earlier, smaller modes have larger growth rates, and so collapse faster than the larger modes. The eventual pattern of fragmentation will be set by those modes that both become unstable and successfully grow and collapse fastest.

Whitworth *et al.* (1994) considered the time evolution of the layer to find the mode which first becomes unstable as the layer evolves. Each element of size L feels an inward acceleration at its edge

$$g \sim \frac{GM}{L^2} - \frac{\nabla P}{\rho} \,. \tag{6.26}$$

The first term represents the gravitational attraction from the mass of the region, while the second term represents the hydrostatic support from pressure gradients. Remembering that $M \sim L^2\Sigma$, using $\nabla P = c_{\rm s}^2\,\nabla\rho$ and approximating $\nabla\rho$ as ρL^{-1}, we obtain

$$g \sim G\Sigma - \frac{c_{\rm s}^2}{L} \,. \tag{6.27}$$

The edge of each region must travel a distance L for collapse. We can use the equation of

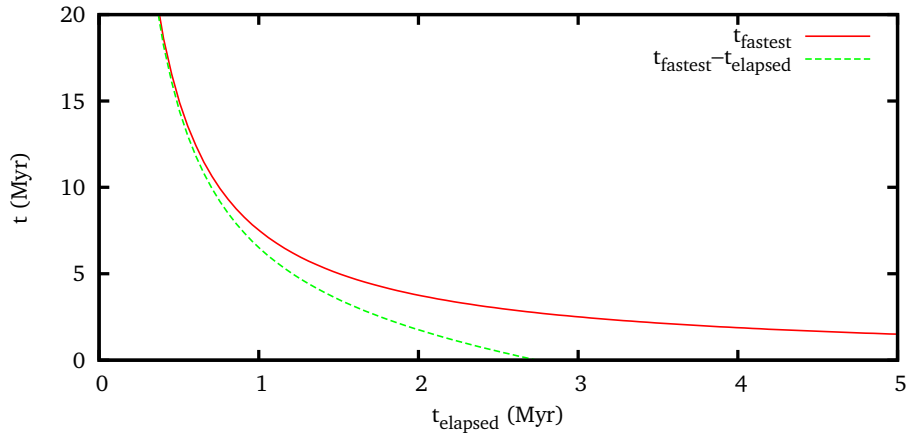


Figure 6.1 – Fastest timescale of gravitational fragmentation t_{fastest} as a function of the time elapsed since the start of the collision t_{elapsed}. The sound speed is $0.2\,\text{km}\,\text{s}^{-1}$ and the speed of the collision is $4\,\text{km}\,\text{s}^{-1}$. The solid red line shows t_{fastest}, while the dashed green line shows $t_{\text{fastest}} - t_{\text{elapsed}}$. When the green line reaches $t = 0$, the fastest and elapsed times are equal.

motion at constant velocity,

$$s = \frac{1}{2} a t^2 \,, \tag{6.28}$$

where s is distance, a is acceleration and t is time, and there is no initial velocity. This gives us the growth time

$$t_{\text{growth}} \sim \sqrt{\frac{L}{g}} \sim \sqrt{\frac{2L^2}{G\Sigma L - c_s^2}} \,. \tag{6.29}$$

Each mode needs at least this long to grow. The function reaches a minimum, corresponding to the fastest growing mode, for a wavelength

$$L_{\text{fastest}} \sim \frac{2c_s^2}{G\Sigma} \,, \tag{6.30}$$

which is simply twice the Jeans length given by equation 6.18. The growth time at this minimum is then

$$t_{\text{fastest}} \sim \frac{2c_s}{G\Sigma} \,. \tag{6.31}$$

Non-linear fragmentation will occur at approximately this time after the start of the collision. Figure 6.1 shows this timescale as a function of time for a colliding layer. When the fastest and elapsed times are equal, the fastest growing mode has just had enough time to grow.

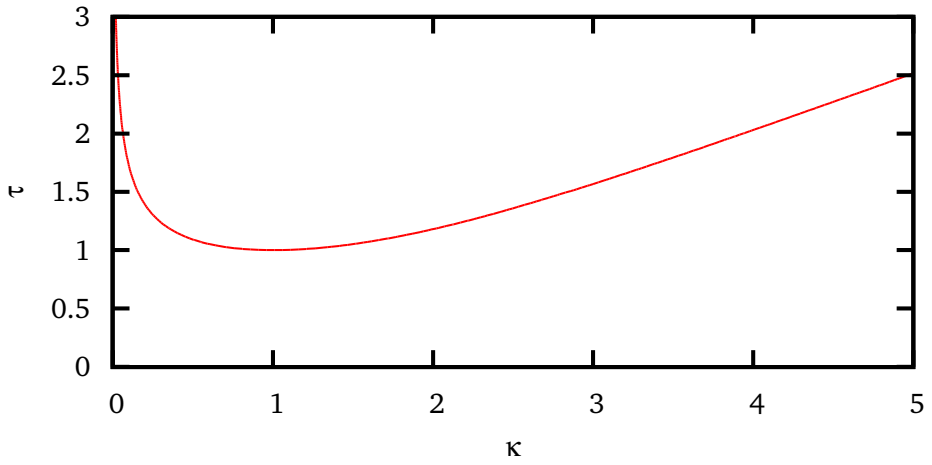


Figure 6.2 – Transition time $\tau = \frac{t}{t_c}$ as a function of wavenumber $\kappa = \frac{k}{k_c}$.

6.2.4 Fastest growing mode by considering growth rates

The same result can be shown by considering the relative growth rates of the gravitational instability and accretion onto the layer. Iwasaki & Tsuribe (2008) use the growth rates calculated by Vishniac to calculate the timescale at which the gravitational instability dominates. The approximate growth rate of the non-linear part of the gravitational instability is given by

$$t_{\mathrm{grav}}^{-1} = \sqrt{-\omega^2}\,. \tag{6.32}$$

The timescale t_{grav} is equivalent to the growth time in equation 6.29. The gravitational instability will become significant when the growth rate of the non-linear part of the instability becomes larger than the growth rate of the layer through accretion. Equation 6.13 shows that the surface density of the layer increases linearly with time through accretion. We define the accretion timescale as the time required to reach the current surface density, which is by definition simply the elapsed time t. The growth rate of the layer through accretion is therefore equal to t^{-1}. At the transition time these growth rates are equal, giving

$$\frac{1}{t} = 2\pi\sqrt{kG\Sigma - c_s^2 k^2}\,. \tag{6.33}$$

This is equivalent to the requirement that the fastest growing mode must have had sufficient time to grow before the evolution becomes non-linear. We combine this with

equation 6.13 to obtain

$$\frac{1}{t} = 2\pi\sqrt{2G\rho_{\text{ext}}\mathscr{M}c_{\text{s}}kt - c_{\text{s}}^2k^2}\,. \tag{6.34}$$

Equation 6.34 depends on both time and Mach number. It is advantageous to define a growth rate that does not depend on Mach number, and to define time and wavenumber in terms of dimensionless variables. Following Iwasaki & Tsuribe (2008) we define k and t in terms of a dimensionless wavenumber $\kappa = k/k_{\text{c}}^{-1}$ and time $\tau = t/t_{\text{c}}^{-1}$, where k_{c} is defined as

$$k_{\text{c}} = \frac{\sqrt{G\rho_{\text{ext}}\mathscr{M}}}{c_{\text{s}}\sqrt{2\pi}}\,, \tag{6.35}$$

and t_{c} is defined as

$$t_{\text{c}} = \frac{1}{\sqrt{2\pi G\rho_{\text{ext}}\mathscr{M}}}\,. \tag{6.36}$$

Equation 6.34 can then be written in a far simpler form as

$$\frac{1}{\tau} = \sqrt{2\kappa\tau - \kappa^2}\,. \tag{6.37}$$

This definition is independent of Mach number. Figure 6.2 shows the relationship between τ and κ. The transition time occurs earliest, at $\tau = 1$, for the mode $\kappa = 1$. Equivalently, the transition time occurs earliest after the critical time t_{c} for the critical wavenumber k_{c}. We therefore describe t_{c} and k_{c} as the critical time and the critical wavenumber respectively.

Substituting $k = \kappa k_{\text{c}}$ and $t = \tau t_{\text{c}}$ into equation 6.25 yields the dimensionless Jeans wavenumber

$$\kappa_{\text{J}} = 2\tau\,, \tag{6.38}$$

which shows that at the critical time t_{c}, the Jeans wavenumber is twice the critical wavenumber. All modes longer than this Jeans wavenumber, i.e. between $\kappa = 0$ and $\kappa = 2$, are unstable. However the critical wavenumber $\kappa = 1$ is the first mode to both become unstable and have had sufficient time to grow. This matches the result of Whitworth *et al.* (1994) given in equation 6.30 which finds that the fastest growing mode at the transition time is twice the size of the Jeans length.

6.3 Previous computational work

Although the analysis of the gravitational instability described so far finds the fastest growing mode at the earliest time that a mode is both unstable and has had time to grow, the later evolution is not predicted. Larger modes become unstable first, but smaller modes, which are not at first unstable, may later grow sufficiently quickly to overtake the larger modes. Such future behaviour is difficult to predict analytically.

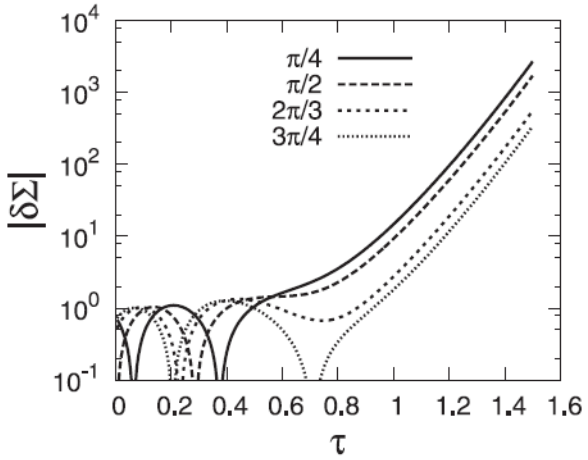


Figure 6.3 – Evolution of perturbations in column density as a function of time (Iwasaki & Tsuribe 2008). The initial perturbation is set up with one of four initial phases. The gravitational instability causes the perturbation to initially oscillate and later collapse.

Iwasaki & Tsuribe (2008) study the time evolution of the gravitational instability using a numerical model. They use the upwind finite-difference method to directly integrate linearized perturbation equations, with Rankine–Hugoniot boundary conditions. They simulate the supersonic collision of two clouds of equal size, uniform density, and which extend infinitely in the xy plane. The results are consistent with results from a one-dimensional non-linear hydrodynamics code.

They study the growth of an initial perturbation as a function of both Mach number and of the amplitude and wavenumber of the initial perturbation. Figure 6.3 shows the time evolution of a perturbation in column density. At early times, the layer is thin and the time evolution of surface density is dominated by accretion. The initial perturbation is oscillatory at these times. At later times, the gravitational instability dominates over the accretion of gas, and the perturbation begins non-linear growth.

The transition time is the time when the gravitational instability begins to dominate, and perturbations switch from oscillating modes to growing modes. The numerical results closely match the analytic prediction given in equation 6.37, and are proportional to $\mathscr{M}^{-1/2}$. As predicted, the $\kappa = 1$ mode is the first to become unstable. However, the mode which first becomes non-linear, and therefore is the fastest-growing, is $\kappa = 2$. Interestingly this is actually the Jeans length, calculated at the critical time when the $\kappa = 1$ mode has had sufficient time to grow.

Solving equation 6.37 for $\kappa = 2$ gives the transition time for the fastest growing mode

at $\tau = 1.18$. Since $t = \tau t_c$,

$$t_{\text{transition}} \approx \frac{1.2}{\sqrt{2\pi G \rho_{\text{ext}} \mathscr{M}}} \,. \tag{6.39}$$

The non-linear time is defined as the time where surface density perturbations become larger in amplitude than the unperturbed surface density, and depends on the wavenumber and initial amplitude of the initial perturbation. The non-linear time is well-fitted by

$$\tau_{\text{nl}} \approx 2.4\, \delta_0^{-0.1} \mathscr{M}^{-1/2} \,, \tag{6.40}$$

where δ_0 is the amplitude of the initial perturbation relative to the unperturbed surface density. This is equivalent to

$$t_{\text{nl}} \approx \frac{2.4\, \delta_0^{-0.1}}{\sqrt{2\pi G \rho_{\text{ext}} \mathscr{M}}} \,. \tag{6.41}$$

Chapter 7

Non-linear thin shell instability – theory

A number of instabilities are important in the star formation process. Instabilities can collect and compress gas, aiding star formation, or can disperse gas, slowing or stopping star formation.

The non-linear thin shell instability (NTSI) was first described by Vishniac (1994). This instability occurs in shock-compressed layers, such as those confined between converging flows. The instability grows from initial perturbations in the layer, and is caused by the non-alignment of ram pressure and thermal pressure. It should not be confused with the similar *Vishniac instability* (Vishniac 1983), which occurs in expanding shells driven into an ambient medium.

As illustrated in Fig. 7.1, ram pressure acts along the collision axis, while thermal pressure is always normal to the surface of the layer. In the NTSI this causes shear in the outer parts of the layer, driving material towards the extremities of the perturbation.

The NTSI is a non-linear instability that can only grow from existing perturbations with an amplitude η larger than or comparable to the thickness of the layer. If the layer is comparatively thick it is less likely that pre-existing perturbations will be of sufficient amplitude. Increasing the collision velocity should lead to more rapid growth of the NTSI, as faster collisions produce a larger density contrast and therefore a thinner layer.

Vishniac (1994) derives the growth rate and the range of unstable wavelengths of the NTSI. The growth rate of the NTSI is given by

$$\tau^{-1}(k) \sim c_{\mathrm{s}} k^{3/2} \eta^{1/2} \,, \tag{7.1}$$

where k is the wavenumber and c_{s} is the sound speed. The NTSI therefore grows faster at smaller wavelengths, but it can only grow within a limited range of wavenumbers. At very large wavelengths, secular damping terms will suppress the NTSI. We can define R as the propagation distance over which zeroth order shock properties evolve. For a stationary slab this is equal to $V_{\mathrm{E}} t$, where V_{E} is the velocity of the inflowing gas and t is the time elapsed since the layer started to form. This sets a maximum wavelength condition that

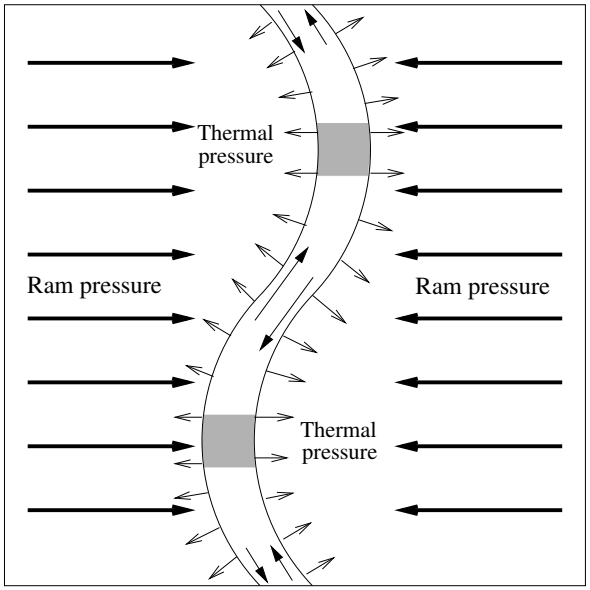


Figure 7.1 – The non-linear thin shell instability. Gas flows in from the left and right, confining the layer with ram pressure (large arrows). The dense gas in the layer resists the compression with thermal pressure (small arrows), but this is always normal to the surface of the layer. The misalignment of ram and thermal pressure causes shear in the layer (shown with arrows within the layer), leading to density enhancements (shown in dark grey).

the NTSI is only present when

$$k \gg \left(\frac{V_S}{c_s}\right) R^{-1} , \tag{7.2}$$

where V_S is the velocity of the inflowing gas relative to the shock at the edge of the layer, and c_s is the sound speed in the layer. The minimum wavelength, and therefore the fastest growing wavenumber, is set by the thickness of the layer. As a bending instability, the NTSI cannot excite any mode on wavelengths shorter than twice the layer thickness L, giving the condition that

$$k < \frac{1}{2L} . \tag{7.3}$$

This corresponds to a bending angle in the layer of 90° and is more stringent than the criteria $kL < 1$ suggested by Vishniac.

At higher collision velocities the initial layer will be thinner, and the NTSI can excite shorter wavelength modes that grow more rapidly. It is also easier to create a perturbation of amplitude greater than the layer thickness. At low collision velocities the range of unstable wavelengths can disappear, suppressing the NTSI.

Finally Vishniac shows that once excited by an initial perturbation the NTSI will grow until saturated at a rate

$$c_s k^{3/2} L^{1/2} < \tau^{-1} < c_s k . \tag{7.4}$$

The lower limit represents the growth rate of the NTSI with a perturbation of the same

size of the layer; this is approximately the smallest perturbation that should be able to excite the NTSI. The upper limit prevents the perturbation from growing to more than its wavelength while remaining subsonic. There is therefore some uncertainty as to the effective growth rate of the NTSI with time, especially as the approach of Vishniac 1994 is time-independent and does not account for the secular evolution of the layer.

7.1 Previous computational work

Blondin & Marks (1996) study the evolution of colliding flows that form a dense, shock-compressed layer using the grid code VH-1[3] (Colella & Woodward 1984). The computational volume is periodic in the y and z directions. Gas enters the box at the 'left' (minimum x) and 'right' (maximum x) boundaries with a velocity of V_{E} and $-V_{\mathrm{E}}$ respectively. Initially the box contains a slab of uniform density and width L across the box at $x = 0$, with a sinusoidal displacement of the slab in the x direction, as shown in Figure 7.2. An isothermal equation of state is used, implying rapid cooling.

They examine the effect of the NTSI on the evolution of a range of initial perturbations at a variety of inflow velocities. They are not able to show conclusively the predicted growth rate $\tau^{-1}(k) \propto k^{3/2}$, since they do not calculate growth rates explicitly as a function of wavenumber. However they show qualitatively that smaller wavelengths show faster growth, as predicted.

As the NTSI progresses the thickness of the layer increases. The bending of the slab becomes increasingly extreme until the inflowing gas can no longer excite the NTSI and the instability is saturated, as shown in Figure 7.3. At this point, the slab becomes bloated and supported by supersonic turbulence.

They also show that the layer is unstable to the Kelvin–Helmholtz instability, with the fastest growing mode being on the order of the thickness of the layer. Later, when the internal motions of the slab become supersonic, the Kelvin–Helmholtz instability is suppressed.

Klein & Woods (1998) performed axisymmetric 2D simulations of colliding atomic clouds using an adaptive mesh refinement grid code. This included a realistic cooling function that included cosmic ray heating, background UV photoionization and photo-electric heating, recombination cooling, collisional ionization cooling and grain cooling. Each cloud is of mass $1.15\,\mathrm{M}_{\odot}$ and radius $0.74\,\mathrm{pc}$, giving initial density of $20\,\mathrm{cm}^{-3}$.

Smooth clouds with no initial perturbation are not unstable to the NTSI, as shown in Figure 7.4. This is because the NTSI is a non-linear instability and requires an initial perturbation. When the clouds have an initial surface perturbation the shock-compressed layer between the colliding cloud becomes unstable to the NTSI and develops the characteristic bending modes. This is shown in Figure 7.5, where the upper and lower clouds have an $m = 5$ and an $m = 7$ perturbation respectively. Klein & Woods do not

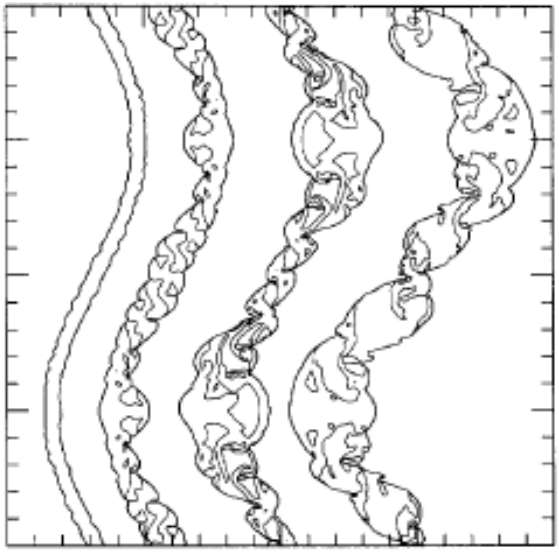


Figure 7.2 – Growth of the NTSI in a shock-compressed slab with an initial sinusoidal perturbation (Blondin & Marks 1996).

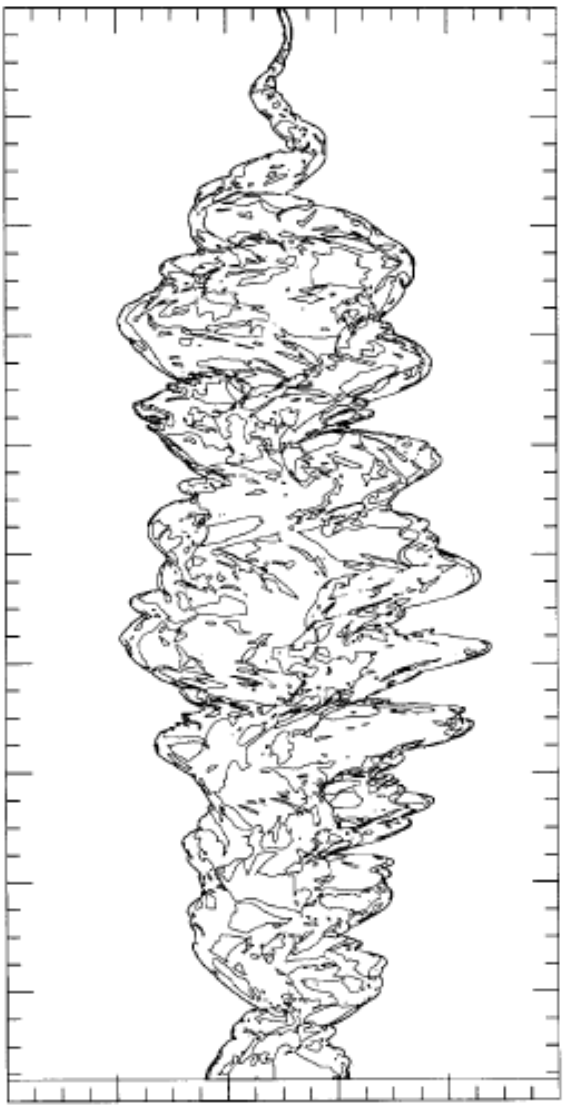


Figure 7.3 – Space-time diagram that shows the growth and saturation of the NTSI. The contours show gas density. The x axis represents distance perpendicular to the slab. The y axis represents both distance along the slab and time, showing the gas density at increasing time down the y axis. This represents a progression of time from top to bottom (Blondin & Marks 1996).

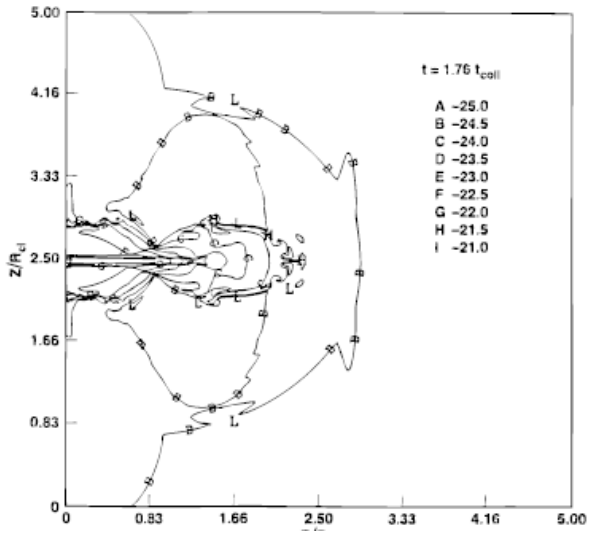


Figure 7.4 – Gas density of colliding smooth clouds after 1.78 collision times (Klein & Woods 1998). The clouds collide along the z axis, shown in units of the cloud radius.

Reprinted from the Astrophysical Journal, Vol. 498, Apr 1998, Richard I. Klein & D. Tod Woods, Bending mode instabilities and fragmentation in interstellar cloud collisions: a mechanism for complex structure, pp 777–799, Copyright (1998), reproduced by permission of the AAS

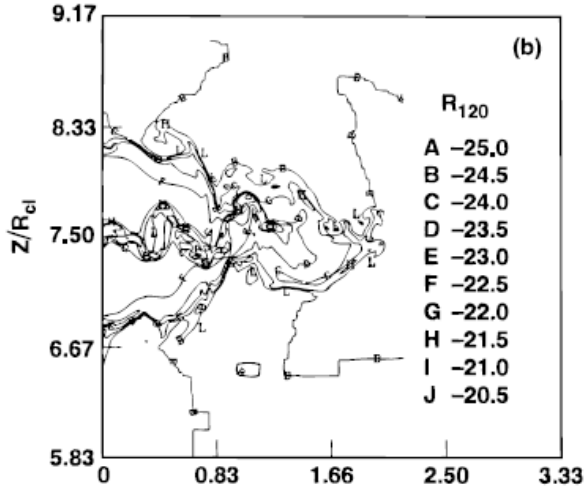
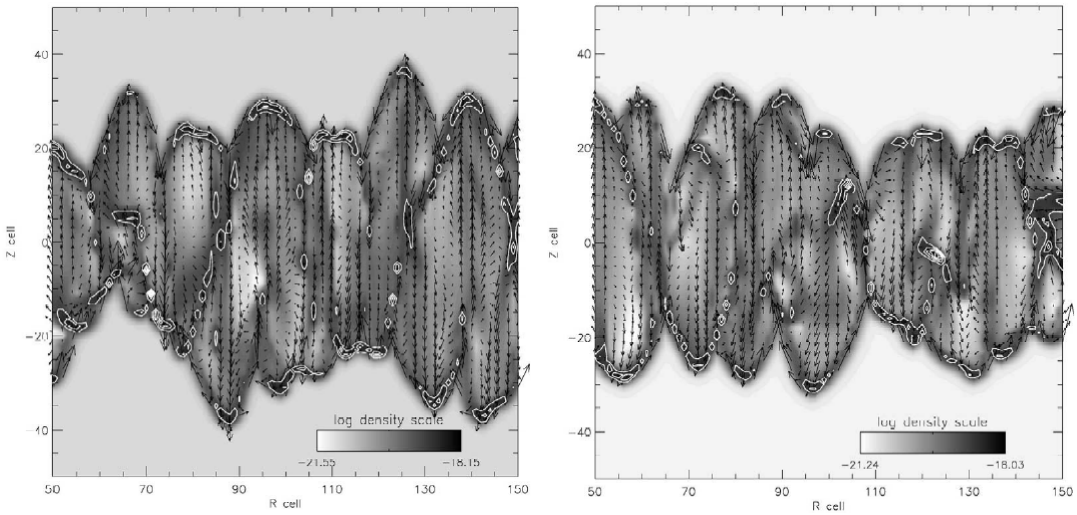


Figure 7.5 – Gas density of colliding clouds with surface perturbations after 1.38 collision times (Klein & Woods 1998). The clouds collide along the z axis, shown in units of the cloud radius.

Reprinted from the Astrophysical Journal, Vol. 498, Apr 1998, Richard I. Klein & D. Tod Woods, Bending mode instabilities and fragmentation in interstellar cloud collisions: a mechanism for complex structure, pp 777–799, Copyright (1998), reproduced by permission of the AAS



(a) Non self-gravitating (b) Self-gravitating

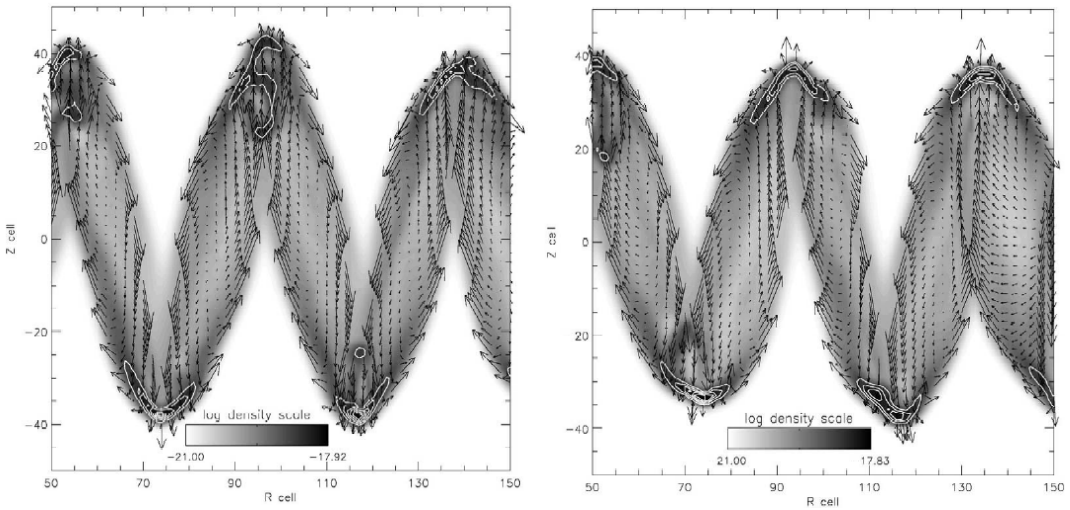
Figure 7.6 – (Hueckstaedt 2003). Gas density of a shock-compressed slab with full cooling.

address the dependence of the growth rate of the NTSI on wavelength, and therefore are not able to test the analytic prediction of Vishniac.

Klein & Woods repeat the collision of perturbed clouds using both an isothermal and an adiabatic equation of state. As expected, the isothermal case is strongly susceptible to the NTSI. In the adiabatic case the heating of the gas reduces the amount of compression, and therefore reduces the maximum density. It also reduces the effect of initial perturbations; there is little difference between the collisions of smooth and initially perturbed clouds. A combination of these effects prevents the NTSI from appearing in their adiabatic collisions.

Hueckstaedt (2003) simulated the NTSI in a cylindrical 2D calculation using the CDFLIB grid code (Kashiwa *et al.* 1994). They used the cooling method of Neufeld, Lepp & Melnick (1995), applying either full (100%) or reduced (30%) cooling. They also considered the effects of self-gravity in some simulations.

Gas flows into the computational volume from the top and bottom with a velocity of $\pm 5\,\mathrm{km\,s^{-1}}$. This gas has a density of $10 \times 10^{-21}\,\mathrm{g\,cm^{-3}}$ and a temperature of 100 K, and forms a dense slab at the centre of the box. The slab is initially unperturbed, and as such is not susceptible to the NTSI. Self-gravity is not used in this first stage, as the intention is simply to build up the mass of the slab. After 0.16 Myr, the slab has reached a density approximately one thousand times greater than initial density, and a perturbation is imposed on the slab. This perturbation is added to the velocity of the slab along the collision, causing displacements of the slab. If the simulation includes self-gravity, it is 'switched on' at this point.



(a) Non self-gravitating (b) Self-gravitating

Figure 7.7 – (Hueckstaedt 2003). Gas density of a shock-compressed slab with cooling reduced to 30% of the expected value.

Gas is heated when it collides with the shocked layer, but rapidly cools to below the initial temperature in both full and reduced cooling simulations. In both cases, the layer becomes unstable to the NTSI once the perturbation as been added. The inclusion of self-gravity does not appear to significantly affect the evolution of the NTSI, while the effect of differing levels of cooling is much more pronounced.

If full cooling is applied, short wavelength modes are rapidly excited. This erases the initial perturbation, creating a chaotic superposition of different modes, as shown in Figure 7.6. If cooling is reduced to 30% of the full value, then the NTSI is not able to develop at these smaller wavelength modes. Instead, the initial perturbation is amplified, as shown in Figure 7.7. Hueckstaedt does not address the dependence of growth rate on wavenumber.

Chapter 8

Generating initial conditions

The first stage in conducting a numerical simulation of star formation, and arguably the most important, is to carefully decide what the initial conditions of the simulations should be. In our case this means specifying the density, temperature and velocity of the system as a function of position.

Generating initial conditions for SPH simulations is the process of creating a set of SPH particles and their properties to represent these initial conditions. In this work we always use equal mass particles, and so the first task is always to carefully place particles to reflect the required density distribution. This cannot be done for a generalized density distribution, but can be done for restricted cases such as a uniform density distribution as described in Section 8.1.

Once the particles have been placed, additional properties of the initial conditions, such as temperature and velocity, can be interpolated onto the particles. Setting particle temperatures is not required if the simulation is isothermal and so specified by the simulation, but they are usually set to a constant or a simple function of density. In our case all particles begin at 10 K.

In our simulations we impose an initial velocity field onto our particles. The velocity field is constructed by a superposition of different components. In all our simulations we have a component that creates a collision at the $x = 0$ plane, with uniform speed either side of the plane. We add further velocity field components depending on the simulation. We use monochromatic and white noise sinusoidal perturbations and both subsonic and supersonic turbulence.

In this section we also consider the resolution requirement of SPH simulations that include both hydrodynamics and self-gravity. We examine how solving the energy equation affects these resolution requirements.

8.1 Generating a glass-like, uniform density box of particles

When generating initial conditions for use in SPH simulations it is important to use a relaxed distribution of particles, as discussed by Whitworth *et al.* (1995a). We start by

placing particles at random into a box. The density of particles in the box will not be uniform, as Poisson noise will cause over-densities and under-densities. This clustering causes SPH to overestimate the average density, as particles naturally oversample the over-densities and undersample the under-densities.

Such a distribution of particles is a higher energy state in SPH; the clustering of particles represents an extra source of potential energy. If a simulation is begun with a randomized particle distribution, the particles will rapidly evolve to a lower energy state. This converts the potential energy into additional kinetic energy, which is not usually desired. Simulations using SPH should therefore take care to relax their initial conditions to remove this source of non-physical energy.

We convert the initially random distribution of particles to a glass-like distribution by allowing them to relax. We give the box periodic boundaries and evolve it isothermally using SEREN. We use hydrodynamics with artificial viscosity, and do not include self gravity. We do not use time-dependent viscosity, as viscosity is important for the relaxation process. Although the particles are initially at rest, they will move away from the over-densities and towards the under-densities. The resultant kinetic energy is then slowly dissipated until the system evolves to the lowest energy state. Given enough time the particles will form a glass-like distribution of particles.

This process can be computationally expensive for large numbers of particles. It is easier to relax a small box, with fewer particles than required, and then use repeated images of this box to build a larger box. Since the small box is created in a simulation with periodic boundaries, the edges between repeated images of the small box will be seamless.

Once this relaxation process is completed, SPH particles will tend to maintain their highly anti-clustered distribution, where particles are evenly spaced. Any other distribution represents a higher energy state, and particles will rapidly adjust themselves to revert to a relaxed distribution.

8.2 Imposing velocity perturbations

In our simulations it is often necessary to impose velocity perturbations onto the initial conditions. In our simulations we begin with a velocity field that causes a collision between two flows of gas. We can then add either sinusoidal perturbations to the interface of the collision, or a turbulent velocity field over all the gas in a simulation. Once a velocity field has been generated, we then have to impose it on the SPH particles themselves.

A simple velocity field for colliding flows is

$$\boldsymbol{v}(x,y,z) = -V\frac{x}{|x|}\,, \tag{8.1}$$

where V is the collision velocity. If the initial conditions are symmetric about $x = 0$, then this creates a symmetric collision along the x axis and the subsequent formation of a layer in the yz plane. We can add a turbulent velocity field or a sinusoidal perturbation to this velocity field before imposing it on the SPH particles of the initial conditions.

8.2.1 Wavenumber convention

In this chapter, we use the symbol k to represent a wavenumber. In all of the equations in this chapter, this represents a integer wavenumber in terms of the length of the box. Therefore for perturbations across a box of length L, the integer wavenumber is simply

$$k = \frac{m}{L}\,, \tag{8.2}$$

where m is the number of perturbations across the box. Our wavenumbers do not include a factor of 2π at any point in this work. The integer wavenumber k corresponds to a 'physical' wavenumber $k_{\rm physical}$ by

$$k_{\rm physical} = \frac{k}{L}\,. \tag{8.3}$$

In this work we provide values of k in physical units, such as pc^{-1}. This prevents our results and our initial conditions from depending on the size of the box.

8.2.2 Monochromatic sinusoidal perturbations

For a monochromatic sinusoidal velocity perturbation, we add a sinusoidal perturbation of a single wavenumber to the interface of colliding gas flows. These perturbations can either be one-dimensional or two-dimensional. We assume the collision occurs along the x axis, and one-dimensional perturbations vary across the y axis. The additional velocity field due to the perturbation is then

$$\delta\boldsymbol{v}(x,y,z) = VA\,\sin\left(2\pi k\frac{y}{L_y}\right) f(x)\,, \tag{8.4}$$

for a one-dimensional perturbation or

$$\delta\boldsymbol{v}(x,y,z) = VA\,\sin\left(2\pi k_{\rm 2D}\frac{y}{L_y}\right)\sin\left(2\pi k_{\rm 2D}\frac{z}{L_z}\right) f(x)\,, \tag{8.5}$$

for a two-dimensional perturbation, where A is the maximum amplitude of the perturbation in terms of the collision velocity V, L_y and L_z are the lengths of the box along the y and z axis, k is the integer one-dimensional wavenumber and $k_{\rm 2D}$ is a two-dimensional integer wavenumber of the perturbation. The scaling coefficient $f(x)$ smoothly reduces the strength of the perturbation to zero outside the central perturbed

zone extending a distance R_{pert} either side of the original collision plane. It is given by

$$f(x) = \begin{cases} \frac{1}{2}\left[\cos\left(\frac{x\pi}{R_{\text{pert}}}\right) + 1\right] & \text{if } |x| < R_{\text{pert}} \\ 0 & \text{if } |x| \geq R_{\text{pert}} \,. \end{cases} \tag{8.6}$$

For a two-dimensional perturbation, we use the integer wavenumber $k_{2\text{D}}$ in equation 8.5, which is the relevant wavenumber along the y and z axes. Constraining the wavenumbers in this way does not give a good representation of the true wavenumber, and so we use the one-dimensional wavenumber in our analysis. The wavenumber $k_{2\text{D}}$ is related to the one-dimensional wavenumber k by

$$k = \sqrt{2}\, k_{2\text{D}} \,. \tag{8.7}$$

This means that it is not possible to excite the same wavenumbers of one-dimensional and two-dimensional perturbations, as we are constrained to use integer wavenumbers of k for one-dimensional perturbations and integer wavenumbers of $k_{2\text{D}}$ for two-dimensional perturbations.

This, and every other form of velocity perturbation we use, does not affect the position or magnitude of the velocity discontinuity at the collision plane. Instead the perturbation adds or subtracts to the velocity on both sides of the discontinuity. This adds extra positive or negative momentum to the gas close to the collision plane. The effect is to push the shock front to one side and so create bending in the shock-compressed layer, while not affecting the Mach number of the collision.

8.2.3 White noise sinusoidal perturbations

For one-dimensional white noise perturbations, we use a velocity perturbation made up of perturbations of equal amplitude over a range of wavenumbers k_{min} to k_{max}. Each perturbation has a randomly generated phase ϕ_k selected from a uniform distribution from 0 to 2π. The one-dimensional white noise perturbation is then

$$\delta \boldsymbol{v}(x,y,z) = V A f(x) \sum_{k=k_{\text{min}}}^{k_{\text{max}}} \sin\left(2\pi k \frac{y}{L_y} + \phi_k\right) \,. \tag{8.8}$$

For two-dimensional white noise perturbations, the velocity perturbation is made up of perturbations of equal amplitude over a range of the two-dimensional wavevector $\boldsymbol{k} = (k_y, k_z)$. Perturbations are included where $k_{\text{min}} \geq k_y \geq k_{\text{max}}$ and $k_{\text{min}} \geq k_z \geq k_{\text{max}}$. Each wavevector $\boldsymbol{k}$ has a randomly generated y and z phase component $\phi_{\boldsymbol{k},y}$ and $\phi_{\boldsymbol{k},z}$ selected from a uniform distribution between 0 and 2π as for the one-dimensional case.

The two-dimensional white noise perturbation is then

$$\delta \boldsymbol{v}(x,y,z) = V A f(x) \sum_{\boldsymbol{k}} \sin\left(2\pi k_y \frac{y}{L_y} + \phi_{\boldsymbol{k},y}\right) \sin\left(2\pi k_z \frac{z}{L_z} + \phi_{\boldsymbol{k},z}\right) , \tag{8.9}$$

where the summation over the wavevector $\boldsymbol{k}$ is over all wavevectors where $k_{\min} \leq k_y \leq k_{\max}$ and $k_{\min} \leq k_z \leq k_{\max}$ as previously described.

8.2.4 Generating a turbulent velocity field

In our work we use only divergence-free turbulent Fourier spectra with a power law form $A(k) \propto k^{\alpha}$. We generate the velocity field in a similar way to that described by Mac Low *et al.* (1998). We create a three-dimensional grid where each grid cell represents a wave vector $\boldsymbol{k}$. We use a grid of size $N_{\text{grid}} = 128$ along each side, giving 128^3 grid cells. The wave vectors are then limited between -63 and 64 along each dimension. We set a limiting $k_{\max} = 63$.

For each wave vector we generate a random value $\mathscr{R}_{\text{u}}$ chosen from a uniform distribution between 0 and 1, and a random value $\mathscr{R}_{\text{g}}$ selected from a Gaussian distribution centred on zero and with a standard deviation of 1. We use these to set a vector amplitude $\boldsymbol{A}(\boldsymbol{k})$ and a phase $\boldsymbol{\phi}(\boldsymbol{k})$ for each wave vector so that

$$\boldsymbol{A}(\boldsymbol{k}) \quad = \quad \begin{cases} \mathscr{R}_{\text{g}} |\boldsymbol{k}|^{\alpha} & \text{if} \quad |\boldsymbol{k}| < k_{\max} \\ 0 & \text{otherwise} \end{cases} , \tag{8.10}$$

and

$$\boldsymbol{\phi}(\boldsymbol{k}) \quad = \quad 2\pi \mathscr{R}_{\text{u}} . \tag{8.11}$$

Each vector component of the amplitude and phase vectors represents a Cartesian component of the final velocity field in real space. We now perform Helmholtz decomposition. This allows any Fourier field to be decomposed into two parts such that

$$\tilde{F} = \tilde{F}_{\text{com}} + \tilde{F}_{\text{sol}} , \tag{8.12}$$

where $\tilde{F}_{\text{com}}$ is the compressive part and $\tilde{F}_{\text{sol}}$ is the solenoidal part. The amplitude of the compressive component is given by

$$\boldsymbol{A}_{\text{com}} = \boldsymbol{A} \cdot \boldsymbol{k} . \tag{8.13}$$

Taking the Fourier transform of the resulting amplitude vectors would result in a curl-free field of compressive turbulence. We want divergence-free solenoidal turbulence, so we

find the solenoidal component of the amplitude

$$\boldsymbol{A}_{\rm sol} = \boldsymbol{A} - \boldsymbol{A} \cdot \boldsymbol{k} \ . \tag{8.14}$$

We separate the three-dimensional grid of amplitude vectors $\boldsymbol{A}_{\rm sol}(\boldsymbol{k})$ into the three component grids $A_{{\rm sol},x}(\boldsymbol{k})$, $A_{{\rm sol},y}(\boldsymbol{k})$ and $A_{{\rm sol},z}(\boldsymbol{k})$. We separate the three-dimensional grid of phase vectors in the same way. We then take a Fourier transform of each matching pair of amplitude and phase grids. For efficiency we use the FFTW (Fastest Fourier Transform in the West) library (Frigo & Johnson 2005). The three Fourier transforms give the x, y and z components of the velocity field. The velocity field is then normalised to the desired average speed. This finally yields our divergence-free velocity field of the desired Fourier spectrum.

8.2.5 Interpolating the velocity field onto SPH particles

We use trilinear interpolation to impose the velocity field onto SPH particles. We consider our grid of velocities to represent a grid of points at the centres of cells. We stretch this grid over the particle distribution as desired, and loop over each particle within the grid. We first scale the position $\boldsymbol{r}_p$ of the particle p into units of the grid size

$$\boldsymbol{S}_p = \frac{\boldsymbol{r}_p N_{\rm grid}}{D} \ . \tag{8.15}$$

We convert this into an integer grid index

$$\boldsymbol{G}_p = \left\lfloor \boldsymbol{S}_p \right\rfloor \ , \tag{8.16}$$

where the floor operator $\lfloor x \rfloor$ returns the largest integer not greater than x, and a fractional position vector

$$\boldsymbol{g}_p = \boldsymbol{S}_p - \boldsymbol{G}_p \ . \tag{8.17}$$

Numbering the velocity grid points from 0 to $N_{\rm grid} - 1$ along each axis, each particle falls within the points $\boldsymbol{G}_p$ and $\boldsymbol{G}_p + \boldsymbol{1}$, where $\boldsymbol{1}$ represents the vector $(1,1,1)$. The velocity of the particle is then given by

$$\boldsymbol{v}_{\rm p,turb} = \boldsymbol{V}\left(\boldsymbol{G}_p\right) + \boldsymbol{g}_p \left[\boldsymbol{V}\left(\boldsymbol{G}_p + \boldsymbol{1}\right) - \boldsymbol{V}\left(\boldsymbol{G}_p\right)\right] \tag{8.18}$$

where $\boldsymbol{V}(\boldsymbol{G})$ is the value of the velocity grid at point $\boldsymbol{G}$, and $\boldsymbol{v}_{\rm p,turb}$ is the velocity of the particle p due to the imposed turbulent velocity field.

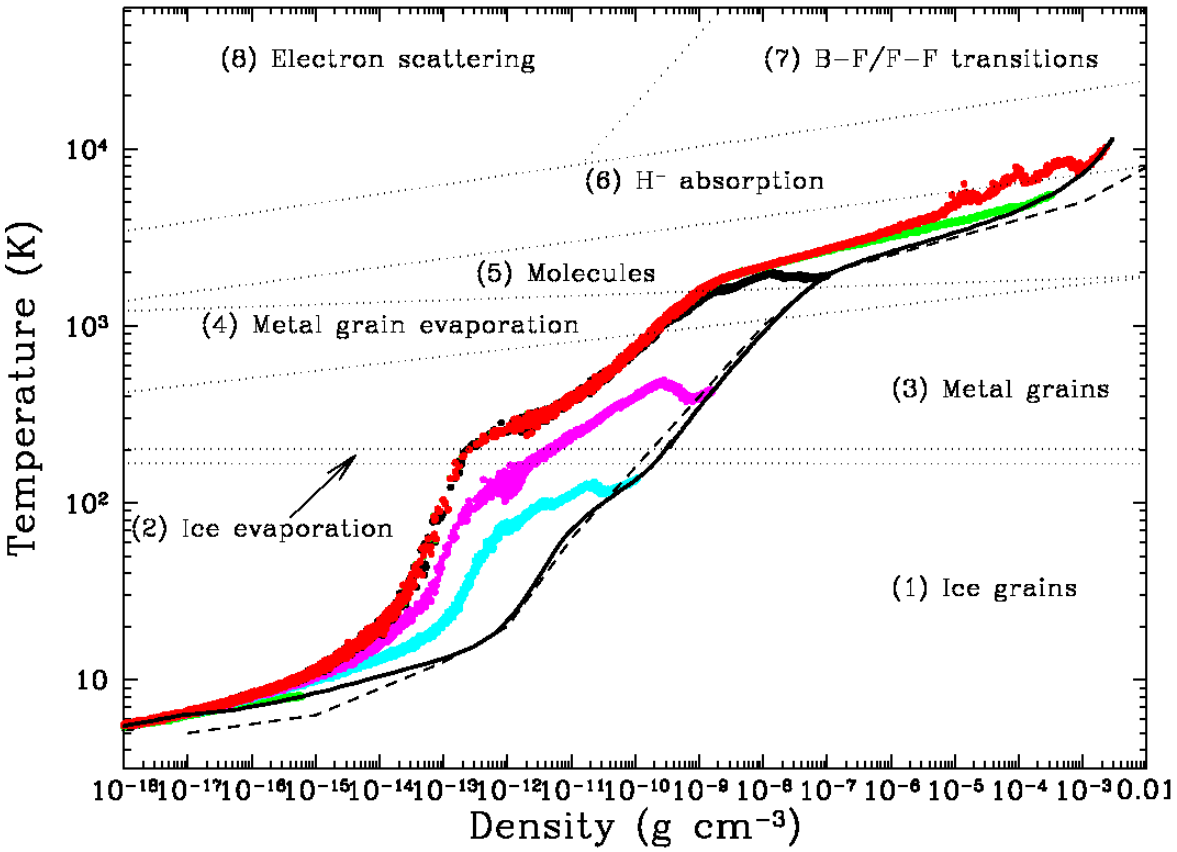


Figure 8.1 – The evolution of temperature and density in a collapsing $1\,M_\odot$ cloud (Stamatellos *et al.* 2007). A time sequence of snapshots are shown, indicated in cyan, magenta, black, green and red. At each snapshot the temperature and density of every SPH particle is plotted using an appropriately coloured dot. The solid black line is the track of the temperature and density at the centre of the cloud at each timestep. The dashed line is the results of (Masunaga & Inutsuka 2000).

8.3 Resolution requirements

Section 2.4.1 describes why it is important to resolve the Jeans mass everywhere and at all times in order to properly resolve gravitational fragmentation. As regions collapse gravitationally, their density will increase, and so the Jeans mass will decrease. For an isothermal simulation, the Jeans mass decreases monotonically with density, leading to a critical density above which simulations will not be resolved.

For more realistic simulations, the Jeans mass reaches a minimum at the opacity limit at approximately $10 \times 10^{-12}\,\mathrm{g\,cm^{-3}}$ as above this density gas become optically thick and heats adiabatically. This raises the sound speed and therefore raises the Jeans mass. If we resolve the Jeans mass at this minimum, we will always resolve the Jeans mass and will therefore obey the Jeans criterion.

Bate, Bonnell & Bromm (2003) show that the smallest mass SPH is able to adequately

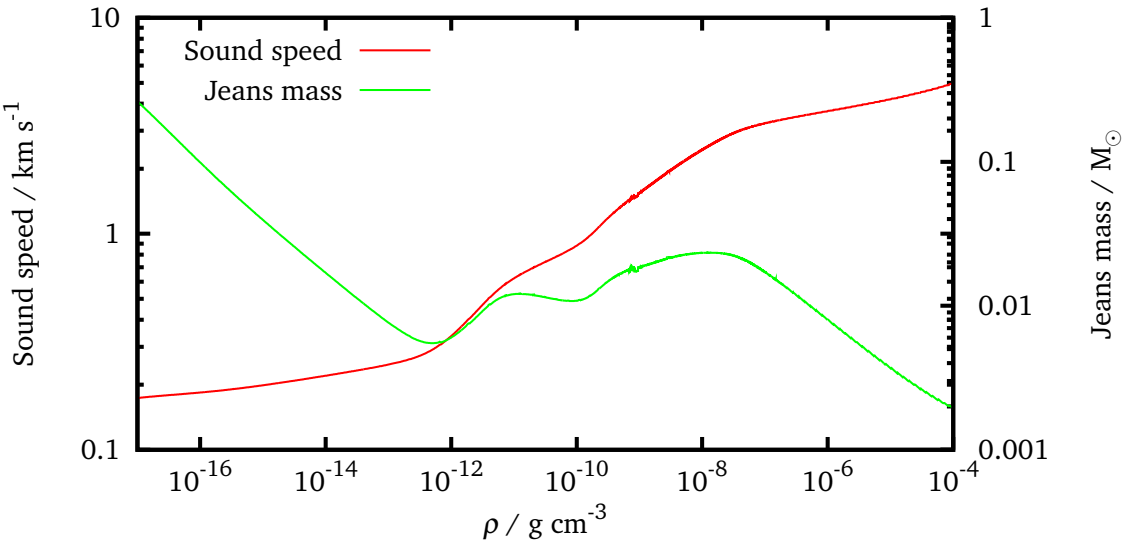


Figure 8.2 – The evolution of sound speed and the Jeans mass at the core of a collapsing $1\,M_\odot$ cloud. At each snapshot the temperature and density of the densest SPH particle is recorded. The equivalent sound speed and Jeans mass is plotted against the particle density.

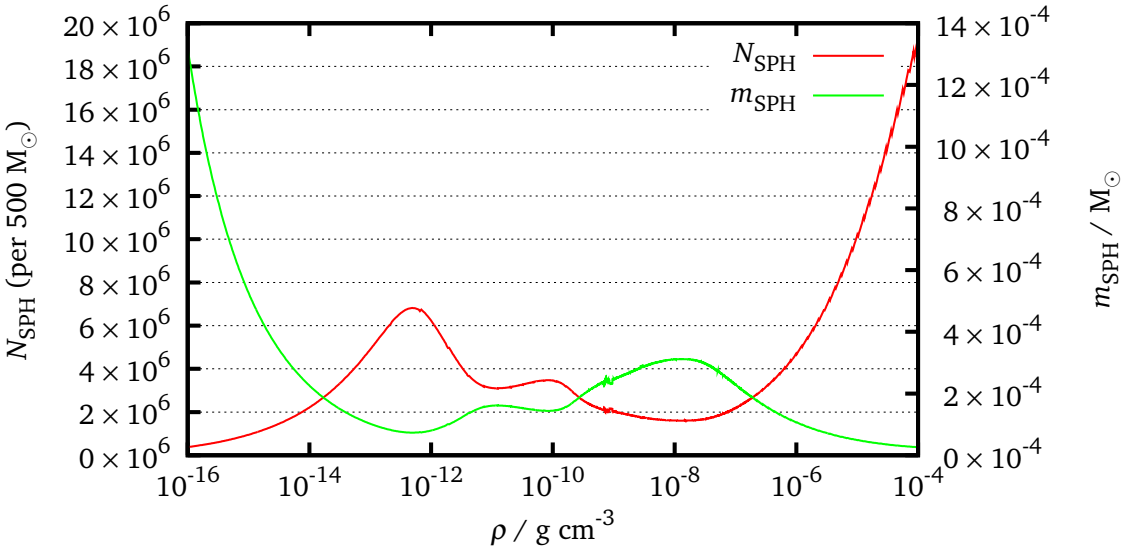


Figure 8.3 – The particle mass $m_{\rm SPH}$ required as a function of density to fulfil the Jeans criterion, together with the total number of particles $N_{\rm SPH}$ required in a $500\,M_\odot$ simulation.

resolve is approximately $1.5\mathcal{N}_{\rm NEIB}$ SPH particles, where $\mathcal{N}_{\rm NEIB}$ is the number of neighbours in a smoothing kernel. If we assume there are typically 50 neighbours in each particle's smoothing kernel, this gives a minimum resolvable mass of approximately $75m_{\rm SPH}$ where $m_{\rm SPH}$ is the mass of an SPH particle.

8.3.1 Critical density for isothermal equation of state

As described in Section 6.1, the Jeans mass for a sphere of uniform density ρ is

$$M_{\rm J} = \sqrt{\frac{375}{4\pi G^3 \rho}} c_{\rm s}^3 \, . \tag{8.19}$$

For an isothermal gas, $c_{\rm s}$ is not a function of density, and so we can equate this to the minimum resolvable mass

$$M_{\rm min} \approx 1.5 \, \mathcal{N}_{\rm NEIB} \, m_{\rm SPH} \, , \tag{8.20}$$

to obtain the minimum resolvable density

$$\rho_{\rm min} \approx \frac{375 c_{\rm s}^6}{9\pi \mathcal{N}_{\rm NEIB}^2 m_{\rm SPH}^2 G^3} \, . \tag{8.21}$$

This can be restated as

$$\rho_{\rm min} \approx 1.2 \times 10^{-13} \, {\rm g \, cm^{-3}} \left(\frac{c_{\rm s}}{0.2 \, {\rm km \, s^{-1}}} \right)^6 \left(\frac{\mathcal{N}_{\rm NEIB}}{50} \right)^{-2} \left(\frac{m_{\rm SPH}}{5 \times 10^{-5} \, {\rm M_\odot}} \right)^{-2} . \tag{8.22}$$

8.3.2 Minimum particle mass for barotropic equation of state

If we are using a barotropic equation of state which is isothermal up to some critical density $\rho_{\rm B}$ and then becomes adiabatic, then this can be used to determine the particle mass required to resolve the Jeans mass at that critical density. As the temperature increases above this density, the Jeans mass increases, and so the Jeans mass is always resolved. This is

$$m_{\rm SPH} \approx \sqrt{\frac{375}{9\pi \rho_{\rm min} G^3}} \frac{c_{\rm s}^3}{\mathcal{N}_{\rm NEIB}} \, , \tag{8.23}$$

and can be restated as

$$m_{\rm SPH} \approx 1.7 \times 10^{-5} \, {\rm M_\odot} \left(\frac{\rho_{\rm min}}{10^{-12} \, {\rm g \, cm^{-3}}} \right)^{-1/2} \left(\frac{c_{\rm s}}{0.2 \, {\rm km \, s^{-1}}} \right)^3 \left(\frac{\mathcal{N}_{\rm NEIB}}{50} \right)^{-1} , \tag{8.24}$$

8.3.3 Minimum particle mass when solving the energy equation

The method we are using to solve the energy equation, described in Section 3.4, is not barotropic, and does not have a fixed relationship between density and temperature. As the Jeans mass depends on sound speed and thus temperature, there is not a fixed

relationship between density and the Jeans mass.

We therefore repeat a test of Stamatellos *et al.* (2007), and set up a simulation following the collapse of a $1\,\mathrm{M}_\odot$ sphere. We track the density and temperature of the densest particle at each timestep. This gives an approximate empirical relation between density and the minimum expected temperature.

We can consider a thought experiment where we consider the collapse of a spherical cloud. We can assume that one particle remains at the core of the cloud, and call this the core particle. As the cloud collapses, the density and temperature of the core increase, giving a core particle temperature/density relation. The core particle is always in a minimally shielded state for its density – it is not surrounded by any higher-density material.

We can now consider a time t in this simulation where the core particle is at some density ρ. The core particle will have a temperature T_{core}. If we now consider a later time $t + \Delta t$, the core particle will now have a higher density. We will now find particles at the density ρ not at the centre but in a shell around the centre. These particles are not in a minimally shielded state. They are partially shielded by higher-density material, and will therefore have a higher equilibrium temperature T than the core particle did when it was at density ρ at time t.

This matches the results shown in Figure 8.1 (Stamatellos *et al.* 2007). When the centre of the cloud has reached a certain density, the temperature of the densest particle is lower than that of any other particle that subsequently reaches that density.

The core particle temperature/density relation therefore gives a minimum temperature/density relation for all particles. All particles at a given density ρ be at a higher temperature than the core particle did when it was at density ρ. Since the Jeans mass is higher at higher temperatures, this gives a relationship between density and minimum Jeans mass.

Using this relationship between minimum temperature and density, we can find a minimum sound speed

$$c_{\mathrm{s}} = \sqrt{\frac{\gamma k_{\mathrm{B}} T}{\bar{m}}}\,, \tag{8.25}$$

where γ is the adiabatic index, k_{B} is the Boltzmann constant, T is the temperature and $\bar{m}$ is the average particle mass. For our non-isothermal simulations the adiabatic index is 1.4, and $\bar{m} = 2.35$ for gas at less than $10^3\,\mathrm{K}$. As described in Section 6.1, the Jeans mass for a sphere of uniform density ρ is

$$M_{\mathrm{J}} = \sqrt{\frac{375}{4\pi G^3 \rho}} c_{\mathrm{s}}^3\,, \tag{8.26}$$

where G is the gravitational constant.

Our results are shown in Figure 8.2. Below $10^{-12}\,\mathrm{g\,cm^{-3}}$ the sound speed increases very slowly as the core density increases. The Jeans mass therefore decreases. Above $10^{-12}\,\mathrm{g\,cm^{-3}}$, the sound speed begins to rise more rapidly. This is the opacity limit, and causes the Jeans mass to increase. We create sink particles at $10^{-11}\,\mathrm{g\,cm^{-3}}$ as described in Section 3.7. The later evolution of the Jeans mass is therefore not important; however, we note the Jeans mass does not reach a smaller minimum until $10^{-5}\,\mathrm{g\,cm^{-3}}$.

Using the requirement that we resolve the Jeans mass with at least $1.5\mathcal{N}_{\mathrm{NEIB}}$ particles, and assuming $\mathcal{N}_{\mathrm{NEIB}} \approx 50$, we can calculate the required mass per particle

$$m_{\mathrm{SPH}} = \frac{M_{\mathrm{J}}}{75}\,. \tag{8.27}$$

This is shown as a function of density in Figure 8.3, together with the total number of particles required for a $500\,\mathrm{M}_{\odot}$ simulation.

Chapter 9

NTSI – simulations

In this chapter, we describe results from simulations of colliding gas flows. The gas flows are anti-symmetric and of uniform density. We impose an asymmetric velocity perturbation on the gas. As the gas collides, a dense layer is formed. The initial velocity perturbation causes bending in this layer and this triggers the NTSI, as described in Chapter 7.

Vishniac (1994) predicts that the NTSI excited by a perturbation in one dimension should have a growth rate

$$\tau^{-1}(k) \propto k^{3/2}\,, \tag{9.1}$$

where k is the wavenumber of the perturbation. We refer to this as the one-dimensional NTSI. To test this relation, we perform simulations including one-dimensional, monochromatic perturbations at a range of wavenumbers. Since real collision-induced dense layers are likely to be subject to random perturbations in two dimensions, we also run tests with a two-dimensional monochromatic perturbation. We compare the growth rates between the one-dimensional and two-dimensional cases.

Generally, simulations of cloud–cloud collision will not have monochromatic perturbations. Instead they may have a range of initial perturbations. We explore two sets of mixed-mode initial perturbations: white noise perturbations over a range of wavenumbers, and a turbulent velocity field applied to the inflowing gas. By comparing our results with the one-dimensional monochromatic case, we can show whether these alternative perturbations also trigger the NTSI, and empirically determine growth rates for the NTSI when triggered by these forms of perturbations.

9.1 Initial conditions

Each simulation is run in a box which is periodic along the y and z axes, and open along the x axis. This is shown in Figure 9.1. Uniform density gas fills the box between the limits $-R_x < x < R_x$, and across the periodic limits $-R_y < y < R_y$ and $-R_z < z < R_z$. The gas flows collide along the x axis at the $x = 0$ collision plane. For non-turbulent simulations, the gas has a uniform speed V_E directed towards the collision plane outside of a central perturbed zone of size $\pm R_\mathrm{pert}$. We refer to any velocity perturbation which we

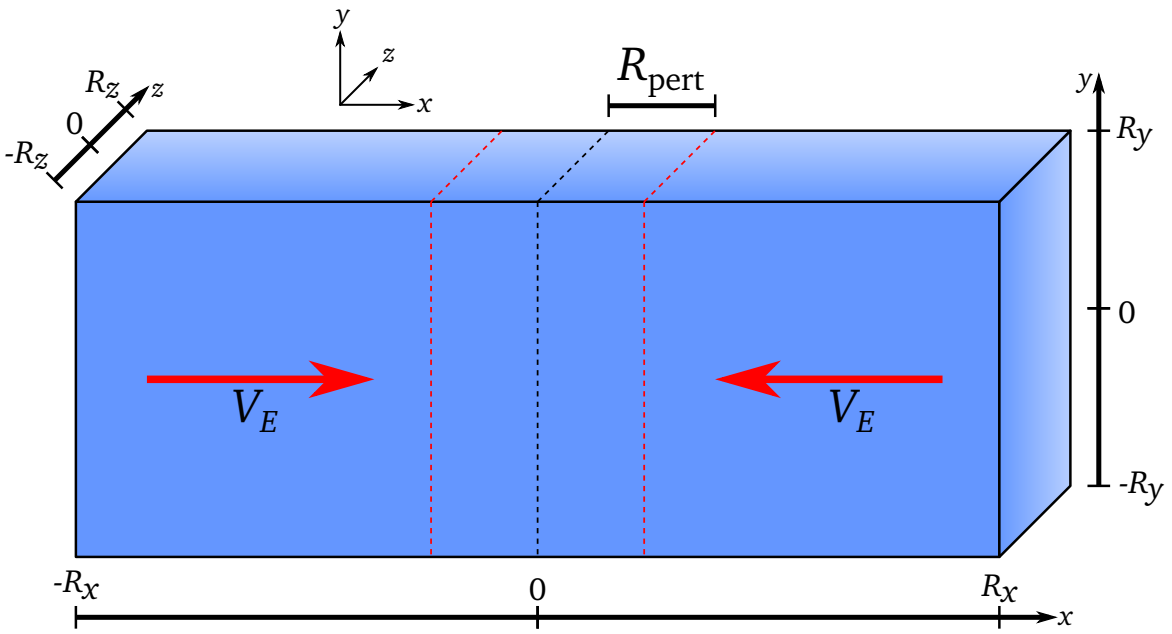


Figure 9.1 – Illustration of the initial conditions of a simulation of the NTSI. The box is periodic in the y and z axes. The black dotted line indicates the collision plane. Red dotted lines indicate the extent of the zone of velocity perturbations.

impose on the initial conditions as the *initial perturbation*.

9.1.1 Physical parameters

We use an initial density of $10^{-21}\,\mathrm{g\,cm^{-3}}$. The gas is isothermal with a temperature of 10 K and a sound speed of $0.188\,\mathrm{km\,s^{-1}}$. The collision velocity V_E is $3.76\,\mathrm{km\,s^{-1}}$, giving a Mach number $\mathscr{M}$ of 20. We set the amplitude of velocity perturbations A to 0.5 for monochromatic perturbations and 0.2 for white noise perturbations over a maximum perturbation distance R_pert of 0.05 pc. For simulations with turbulence, we use a Fourier spectrum of the form $A(k) \propto k^{-2}$ as described in Section 8.2.4. The average speed of the turbulent velocity field is $0.2V_\mathrm{E}$ (Mach 4) for supersonic turbulence and $0.025V_\mathrm{E}$ (Mach 0.5) for subsonic turbulence.

As the box is not infinitely long, the simulations cannot be continued indefinitely. As the ends of the gas flows are not confined, a rarefaction wave will move through the gas flow at the sound speed. Since the gas is moving with the collision speed V_E, the rarefaction wave moves, relative to the collision plane, at a speed $V_\mathrm{E} + c_\mathrm{s}$. The simulation will be a good approximation of an infinitely long gas flow until the rarefaction wave reaches the collision plane, which will happen at a time

$$t_\mathrm{r} = \frac{R_x}{V_\mathrm{E} + c_\mathrm{s}}\,. \tag{9.2}$$

To avoid re-expansion of the layer when the confining ram pressure is removed, our

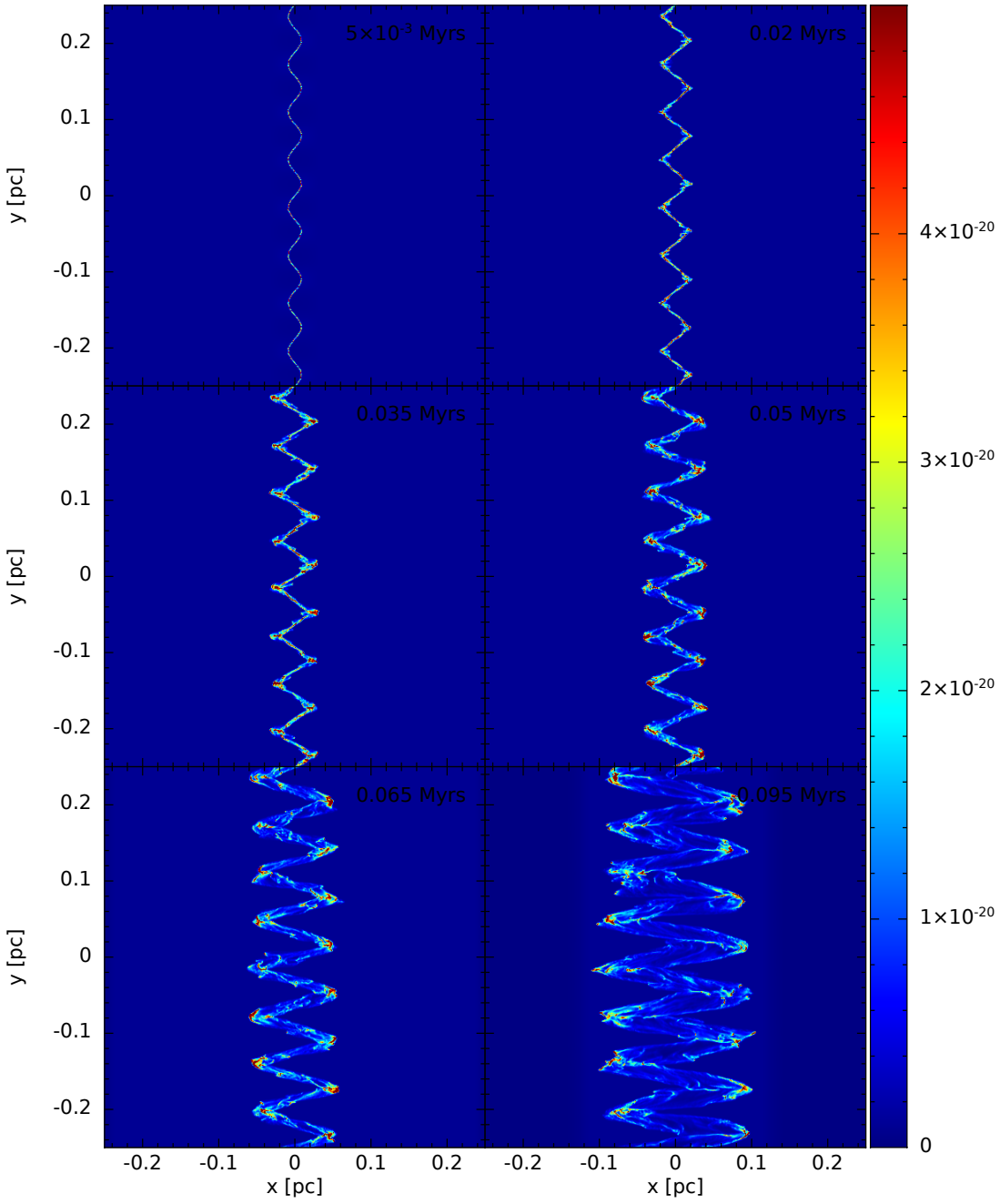


Figure 9.2 – Cross-sections of density ($\mathrm{g\,cm^{-3}}$) in the xy plane for 1DHR simulation of monochromatic perturbations of wavenumber $k = 16\,\mathrm{pc^{-1}}$. Similar plots for other simulations can be found in Appendix B.

simulations do not run for longer than this time.

Three suites of simulations have been simulated: a low-resolution suite with one-dimensional perturbations, a high-resolution suite with one-dimensional perturbations and a low-resolution suite with two-dimensional perturbations. We refer to these as the 1DLR, 1DHR and 2D simulation suites respectively. These are summarized in Table 9.1.

Each suite contains simulations of monochromatic perturbations for each of the wavenumbers 4, 8, 12, 16, 20, 24, 28 and $32\,\mathrm{pc}^{-1}$. Note that these are 'physical' wavenumbers, as described in Section 8.2.1, and so do not depend on the size of the simulation box. Each wavenumber is tested with five different realizations to improve the overall results, giving a total of 40 simulations per suite. An example of a collision with a monochromatic perturbation is shown in Figure 9.2.

The 1DHR and 2D suites also contain simulations with white noise perturbation and simulations with initial turbulence. The white noise perturbations cover a range of wavenumbers up to the limiting wavenumber $k_{\mathrm{max}} = 20$. Each of the five realizations is combined with two randomly generated sets of white noise perturbations giving ten simulations per suite.

We use both subsonic turbulence and supersonic turbulence. We produce ten sets of turbulent velocity fields with the Fourier spectrum described in Section 8.2.4. By scaling each velocity field to the appropriate average speed, we produce ten subsonic velocity fields and ten supersonic velocity fields. Each of the five realizations is combined with two subsonic and two supersonic velocity fields giving ten subsonic simulations and ten supersonic simulations per suite.

Simulating a two-dimensional perturbation can be more computationally expensive than simulating a one-dimensional perturbation, which is why we only perform a low-resolution study of two-dimensional perturbations. The 1DLR suite is performed at the same resolution as the 2D suite to allow comparison between the 1DHR suite and the 2D suite. Table 9.2 and Table 9.3 give details of the sizes and resolutions of the simulations in the three suites. Table 9.4 shows the expected timescale for the rarefaction wave to reach the collision plane; this table shows that this will not occur until after the end time of simulations in each suite.

9.1.2 Creating the collision boxes

We use the relaxation process described in Section 8.1 to create settled particle distributions. This can be computationally expensive, and so we therefore first create a box with the same y and z dimensions as desired, but one twentieth the length in the x dimension. Particles are placed randomly in this box and relaxed as normal. The full-size box is then created by lining up twenty copies of the relaxed box. This leaves no regularly repeating patterns across the y or z axes which could help or hinder the development of instabilities in the layer. Particles will repeat along the x axis at regular intervals, but

	One-dimensional		Two-dimensional
	High resolution	Low resolution	Low resolution
	1DHR	1DLR	2D
Monochromatic	5 per wavenumber[1]	5 per wavenumber[1]	5 per wavenumber[1]
White noise	10	-	10
Supersonic turb.	10	-	10
Subsonic turb.	10	-	10

[1] Wavenumbers 4, 8, 12, 16, 20, 24, 28 and 32 pc^{-1}

Table 9.1 – Number of realizations for each set of simulations in each suite

Suite	N_{SPH}[1]	Box lengths (pc)			Particles across box[2]		
		$2R_x$	$2R_y$	$2R_z$	x	y	z
1DHR	8 000 000	1	0.5	0.01	1170	585	11.7
1DLR	1 000 000	2	0.5	0.1	431	108	21.5
2D	5 000 000	2	0.5	0.5	431	108	108

[1] Total number of particles
[2] Based on average inter-particle spacing

Table 9.2 – Sizes of gas flows and simulation box for NTSI simulations

Suite	m_{SPH} ($\mathrm{M}_\odot$)[1]	Particle density (pc^{-3})	Particle spacing (pc^{-1})
High resolution 1D	9.23×10^{-9}	1.6×10^{9}	1170
Low resolution 1D	1.48×10^{-6}	10^{7}	215
2D	1.48×10^{-6}	10^{7}	215

[1] Mass per particle

Table 9.3 – Resolutions of NTSI simulations

Suite	R_x (pc)	v_{r} ($\mathrm{km\,s}^{-1}$)[1]	t_{r} (Myr)[2]	End time (Myr)
High resolution 1D	0.5	3.95	0.124	0.1
Low resolution 1D	1	3.95	0.248	0.15
2D	1	3.95	0.248	0.15

[1] Rarefaction wave velocity
[2] Time that rarefaction wave reaches collision plane

Table 9.4 – Timescales for rarefaction wave to reach collision plane in NTSI simulations

this is not expected to significantly affect the development of instabilities along the $x = 0$ plane.

This process is repeated to produce five initial realizations for each suite which differ only by the original random placement of particles. For the monochromatic simulations, each of the five realizations is used to create eight simulations, one for each wavenumber, giving a total of forty simulations.

Each realization is also used to create two simulations with white noise, two simulations with subsonic turbulence and two simulations with supersonic turbulence. The monochromatic velocity perturbations are as described in Section 8.2.2, while white noise perturbations are as described in Section 8.2.3. The particle velocities are then set as described in Section 8.2.5.

For simulations with an initial turbulent velocity field, the field is generated as described in Section 8.2.4. The velocity field is cubic and periodic, and does not fit neatly onto our simulations. To provide consistency between our simulations, we therefore stretch the grid to a $0.5\,\mathrm{pc} \times 0.5\,\mathrm{pc} \times 0.5\,\mathrm{pc}$ cube centred on the origin. This grid is then repeated over the x axis as required to cover the entire simulation, and velocities are interpolated onto particles as described in Section 8.2.5.

No other initial perturbations are included in the simulations with turbulent velocity fields. These simulations differ from other simulations in that the initial perturbations are three-dimensional and extend throughout the entirety of the simulation. However, since the simulations with one-dimensional perturbations are much taller in the y than in the z direction, the largest perturbations will only exist approximately along the y axis. Since we use a Fourier spectrum where modes with higher wavenumbers have smaller amplitudes, these will be the most significant wavenumbers. This gives us a reasonable approximation for one-dimensional perturbations.

For the 2D simulations, where the y and z axes are of equal length to the periodic velocity grid, the velocity field will be periodic across the boundary. However, for the 1DLR and 1DHR simulations, where the z axis is smaller than the y axis, the imposed velocity field will not be periodic and there will be a velocity discontinuity at the edges. This is not expected to be significant as long as the z axis is small enough that the velocity change across the z axis is negligible compared to the velocity changes along the x and y axes. The velocity field will repeat along the x axis, but this should not affect the perturbation in the y and z axes significantly.

9.2 Theory

Although it is difficult to build a simple model describing the behaviour of one of our simulations with a complex velocity field, such as a white noise perturbation or turbulence, it is possible to predict the early evolution of the layer for a simple monochromatic

perturbation. To do this we consider only the peak of the perturbation and ignore the effects of the NTSI.

Based on the work of Vishniac (1994), we can also predict a number of properties of the NTSI. We can estimate the range of wavenumbers susceptible to the NTSI. We can make an approximate estimation, as a function of wavenumber, of the time at which the NTSI becomes dominant and the corresponding initial growth rate, and we can estimate the rate at which the NTSI grows to saturation.

In this work we differentiate between the *growth rate* and the *rate of growth*. The growth rate is the reciprocal of a growth timescale which in some way reflects how long something takes to grow. The rate of growth of a quantity is simply the instantaneous rate at which that quantity is growing in time. The growth rate therefore has units of inverse time, such as Myr^{-1}, while the units of rate of growth depend on the quantity which is growing. For a bending-mode perturbation in a layer, the amplitude of the perturbation is given by the maximum displacement of the peak of the perturbation. The rate of growth is then the velocity at which the peak is moving in units of velocity, such as $\mathrm{pc\,Myr}^{-1}$. For breathing-mode perturbations, the units of the rate of growth would be units of surface density per unit time, such as $\mathrm{g\,cm}^{-2}\,\mathrm{Myr}^{-1}$.

9.2.1 The decay of monochromatic sinusoidal perturbations

We are using relatively strong perturbations in our monochromatic simulations. We can predict the behaviour of the layer, in the absence of the NTSI, by considering the mass and momentum of the layer. Our initial velocity field is of the form

$$\boldsymbol{v}(x,y,z) = V_{\mathrm{E}} \left[\frac{-x}{|x|} + A \sin\left(k\pi \frac{y}{R_y} \right) f(x) \right] , \tag{9.3}$$

where V_{E} is the collision velocity, x is the position along the collision axis, with the collision plane at $x = 0$, A is the amplitude of the initial perturbation, k is the wavenumber of the perturbation, y is the position across the collision plane, R_y is the half the length of the box along the collision plane as shown in Figure 9.1 and $f(x)$ is the smoothing function defined in equation 8.6. This velocity field includes a collision component and a monochromatic perturbation of the form described in Section 8.2.2. We consider now only the velocity perturbation $\delta\boldsymbol{v}$ at a point where the perturbation is greatest, which is then

$$\delta\boldsymbol{v} = V_{\mathrm{E}} A f(x) . \tag{9.4}$$

We can consider, at the peak of a perturbation, the extra momentum δp per unit area a due to the perturbation between symmetric limits $-d$ and d. This extra momentum per

unit area is

$$\frac{\delta p}{a} = \frac{1}{a}\int_{-d}^{d} \boldsymbol{v}\,\mathrm{d}m\,. \tag{9.5}$$

Substituting the mass element $\mathrm{d}m = \rho\, a\,\mathrm{d}x$ we obtain

$$\frac{\delta p}{a} = \frac{1}{a}\int_{-d}^{d} a\,\rho_0 \boldsymbol{v}\,\mathrm{d}x\,, \tag{9.6}$$

where ρ_0 is the original background density of the layer. Substituting in $\boldsymbol{v}$, this becomes

$$\frac{\delta p}{a} = \frac{\rho_0 V_\mathrm{E} A}{2}\int_{-d}^{d}\left[\cos\left(\frac{x\pi}{R_\mathrm{pert}}\right)+1\right]\mathrm{d}x\,, \tag{9.7}$$

assuming that $d \leq R_\mathrm{pert}$. This gives us

$$\frac{\delta p}{a} = \rho_0 V_\mathrm{E} A\left[\frac{R_\mathrm{pert}}{\pi}\sin\left(\frac{d\pi}{R_\mathrm{pert}}\right)+d\right]\,, \tag{9.8}$$

which is the additional momentum per unit area a at the peak of the perturbations due to the perturbation between $-d$ and d where $d \leq R_\mathrm{pert}$. For $d = R_\mathrm{pert}$, we obtain

$$\frac{p}{a} = \rho_0 V_\mathrm{E} A\, R_\mathrm{pert}\,, \tag{9.9}$$

which is the total additional momentum per unit area a at the peak of the perturbations.

The surface density of the layer increases with time according to

$$\frac{m_\mathrm{layer}}{a} = 2V_\mathrm{E}\rho_0 t\,. \tag{9.10}$$

By considering the total additional momentum of the perturbation and the mass of the layer, we can predict the velocity of the layer due to the perturbation alone as a function of time. There is a critical time τ where all the momentum of the perturbation has entered the layer, defined as

$$\tau = \frac{R_\mathrm{pert}}{V_\mathrm{E}}\,, \tag{9.11}$$

which for our simulations is 0.013 Myr. Before this critical time, only a portion $d = Vt$ of the perturbation momentum has entered the layer. The velocity of the layer is then

$$v_\mathrm{layer}(t) = \frac{A}{2t}\left[\frac{R_\mathrm{pert}}{\pi}\sin\left(\frac{\pi V_\mathrm{E} t}{R_\mathrm{pert}}\right)+V_\mathrm{E} t\right]\,. \tag{9.12}$$

At the critical time τ where $V_E t = R_{pert}$ this is simply

$$v_{layer}(\tau) = \frac{A V_E}{2} . \tag{9.13}$$

Integrating with respect to time gives the x position of the layer

$$x_{layer}(t) = \frac{A R_{pert}}{2\pi} \int_0^t \left[\frac{1}{t} \sin\left(\frac{V_E t \pi}{R_{pert}} \right) \right] + \frac{A V_E t}{2} , \tag{9.14}$$

giving the position of the layer at the critical time τ as

$$x_{layer}(\tau) = \frac{A R_{pert}}{2} . \tag{9.15}$$

This is only a first estimate of the early behaviour of the layer. As the layer moves into the inflowing gas it will tend to collect more mass on its leading edge than on its trailing edge. This leads to an additional source of momentum as the layer is no longer receiving equal momentum per unit time from the left and right flows. We can estimate, at the critical time, how much momentum we have failed to account for.

The momentum of the layer per unit area at the critical time is

$$\frac{p}{a} = \rho_0 V_E A R_{pert} . \tag{9.16}$$

If the layer has moved a distance x in this time, it has accumulated an extra momentum per unit area $x\rho_0 V_E$ on its leading edge, acting to slow the layer, and it has not accumulated the same amount on its trailing edge which would have accelerated the layer. The extra momentum per unit area that we are not accounting for is therefore

$$\frac{p_{missed}}{a} = 2x\rho_0 V_E , \tag{9.17}$$

and inserting our estimate of x at the critical time gives

$$\frac{p_{missed}}{a} = \rho_0 V_E A R_{pert} . \tag{9.18}$$

Since this is equal to the total momentum due to the perturbation, we have failed to account for half of the total momentum of the layer. Our estimate of the rate at which the layer slows is therefore likely to be inaccurate. Due to the piecewise velocity field, it is difficult to solve analytically for the position of the layer as a function of time, and so we integrate numerically.

At all times, we can calculate the mass of the layer per unit area $2V_E\rho_0 t$, and the velocity of the layer is always given by the total momentum per unit area absorbed by the

layer divided by the mass per area. The momentum per unit area is given by the integral of momentum elements through the section of the original velocity field that has become part of the layer. This is

$$\frac{p(x,t)}{a} = \int_{x_\mathrm{L}}^{x_\mathrm{R}} \rho_0 \, v_x(x) \, \mathrm{d}l \, , \tag{9.19}$$

where x_L and x_R are the left and right extent and $v_x(x)$ is the original velocity field given in equation 9.3. For a layer at time t and central position x, ignoring the thickness of the layer, the left and right extents are given by $x - V_\mathrm{E}t$ and $x + V_\mathrm{E}t$ respectively.

We integrate this numerically in advance between limits of 0 and l and store this as a table at a large number of values of l. For any choice of r_L and r_R, we can quickly look up from a table the total momentum per unit area at any layer position since

$$\int_{r_\mathrm{L}}^{r_\mathrm{R}} \rho_0 v_x(x) \, \mathrm{d}l = \int_0^{r_\mathrm{R}} \rho_0 v_x(x) \, \mathrm{d}l - \int_0^{r_\mathrm{L}} \rho_0 \, v_x(x) \, \mathrm{d}l \, . \tag{9.20}$$

We can now integrate the variable x with respect to time, for which we use the simple Euler method with a large number of steps. We set our initial values x_0 and v_0 to 0 and $A\,V_\mathrm{E}$ respectively, and evolve x so that

$$x_{t+\Delta} = x_t + \frac{p(x,t)}{a} \frac{a}{m(t)} * \Delta t \, , \tag{9.21}$$

where $p(x,t)/a$ is the momentum per unit area calculated from equation 9.19 , $m(t)/a$ is the mass per unit area calculated in equation 9.10, and Δt is the time interval between steps.

Figure 9.3 shows the rate of growth of the amplitude of the perturbation as predicted by our numerical solution. This rate of growth is simply the velocity of the layer at the peak of the sinusoidal perturbation. The velocity decreases very gradually until approximately the critical time 0.013 Myr, after which the velocity decreases proportional to t^{-2}.

In reality the layer is not infinitesimally thin. As the layer becomes wider, it captures more momentum and mass from the inflowing gas than we assume. The extra mass causes the layer to slow more quickly than predicted. The layer will capture more mass on the leading edge. This extra momentum will further slow the layer, although it will also move the centre of mass of the layer in the direction of motion. We have not explored these effects in detail, as they should be only small correction to our thin layer model.

9.2.2 Growth rates of the NTSI

As described in Chapter 7, Vishniac (1994) predicts the growth rate of the NTSI to be

$$\tau^{-1} \sim c_\mathrm{s} \, k^{3/2} \eta^{1/2} \, , \tag{9.22}$$

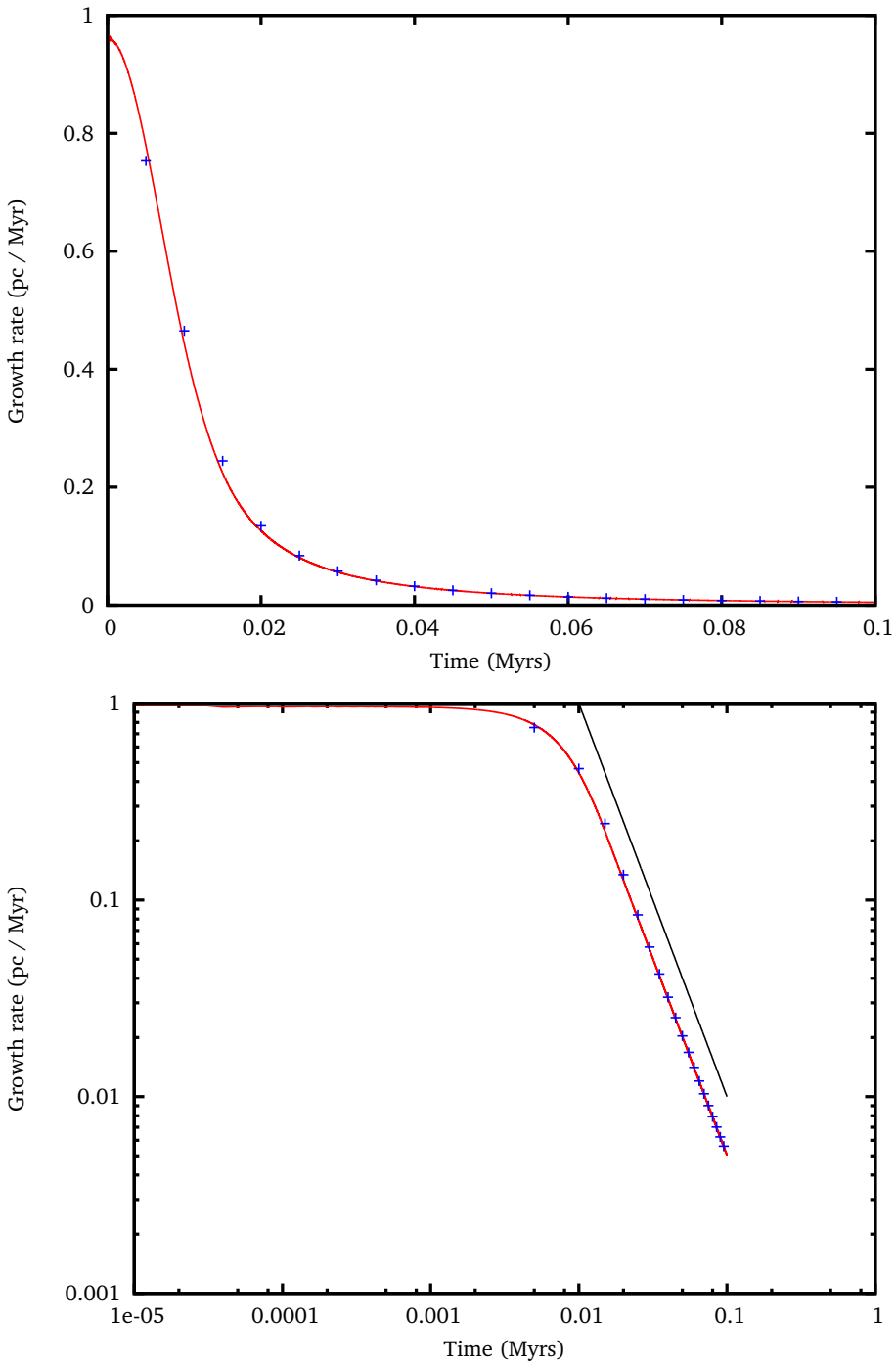


Figure 9.3 – Expected rate of growth in perturbation amplitude of a one-dimensional sinusoidal velocity perturbation. The red line shows the results of the semi-analytic solution. The blue points indicate the growth calculated by numerical differentiation of the layer position over the same intervals and using the same method as in the equivalent SPH simulations. The black line indicates shows a line $y \propto t^{-2}$.

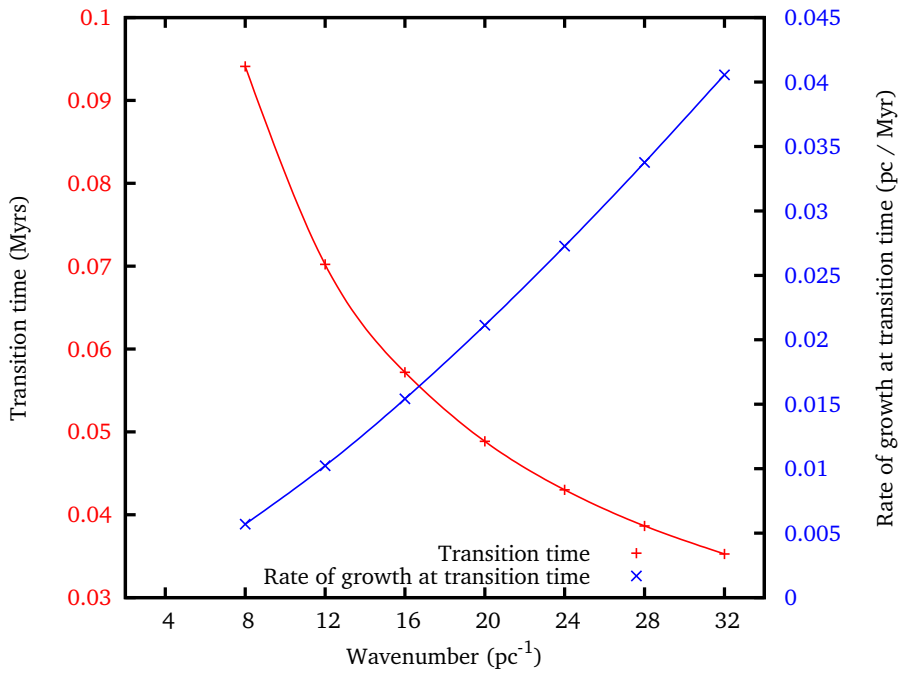


Figure 9.4 – The transition time: the time at which the rate of growth of the initial perturbation is first exceeded by the rate of growth of the NTSI triggered by that initial perturbation. The red points show this transition time as a function of wavenumber. The blue points show the rate of growth of the NTSI at this transition time. No points are plotted for the wavenumber $4\,\mathrm{pc}^{-1}$ as the transition time does not occur within the simulation time of 0.15 Myr.

for one-dimensional monochromatic perturbations in a dense shock-confined layer, where η is the amplitude of the perturbation. It is not clear what the relevant value of η is for our simulations, as it will depend on the amplitude of the decaying initial perturbations at the time when the NTSI becomes dominant.

For a single monochromatic initial perturbation, we have calculated the growth of the perturbation numerically in Section 9.2.1. The rate of growth of the initial perturbation decreases as more mass is accreted to the layer, but its amplitude continues to increase and so the growth rate of the NTSI increases. We use the results of our numerical simulation to calculate the transition time at which the rate of growth of the NTSI first exceeds the decaying rate of growth of the initial perturbation.

We need to convert the growth rate of the NTSI, which is the reciprocal of the time it would take to build the perturbation of amplitude η, into a rate of growth. We assume we have some approximately sinusoidal perturbation; the amplitude of that perturbation is then half the peak-to-peak distance, which is the distance of the peaks from the original collision plane. The speed of the peaks of the perturbation is therefore the rate at which the perturbation amplitude η is growing; this is the rate of growth of the perturbation. The speed is simply the size of the perturbation divided by the time taken for it to grow,

Wavenumber (pc^{-1})	4	8	12	16	20	24	28	32
Growth rate (Myr^{-1})	0.77	1.54	2.31	3.08	3.85	4.62	5.39	6.16
Growth time (Myr)	1.30	0.65	0.43	0.32	0.26	0.22	0.19	0.16

Table 9.5 – Maximum predicted growth rate of the NTSI to saturation

giving

$$\omega = \frac{\eta}{\tau}\,, \tag{9.23}$$

where ω is the rate of growth and τ is the growth timescale. Since the growth rate, as given in Equation 9.22, is simply the reciprocal of the growth timescale, we obtain the rate of growth of the NTSI

$$\omega \sim c_s\, k^{3/2} \eta^{3/2}\,. \tag{9.24}$$

Figure 9.4 shows, for each of our wavenumbers, the time at which the growth rate of the initial perturbation is equal to the growth rate of the NTSI due to that perturbation. It also shows the growth rate at that time for each wavenumber, which should be a minimum growth rate for simulations with that wavenumber.

Our numerical model assumes the NTSI has no effect before the transition time. In reality the NTSI will affect the amplitude of the perturbation before the transition time. While this is only an estimate, the prediction is that higher wavenumbers will grow earlier and faster, and will reach a higher minimum wavenumber.

Vishniac (1994) states that, once excited, the NTSI will grow to saturation at a rate

$$c_s k^{3/2} L^{1/2} < \tau^{-1} < c_s k\,, \tag{9.25}$$

where L is the layer thickness. As described in Chapter 9, these limits are results of the minimum and maximum size of perturbations for which the analysis of the NTSI holds. The lower limit is the growth rate of a perturbation with wavelength equal to the layer thickness, and is the smallest perturbation which should be able to excite the NTSI.

The analysis of the NTSI is only defined for small bending angles, and once the amplitude of the perturbation is comparable to the wavelength the instability will become saturated. The upper limit is then the maximum growth rate allowing the perturbation amplitude to grow to this size without becoming supersonic.

For our simulations, the upper limit of the growth rate is given in Table 9.5. From this estimate we can see that the growth time of the NTSI should be much greater than the duration of our simulations. However, we show that the layer is strongly affected by the NTSI well before the growth times predicted.

9.2.3 Wavenumbers susceptible to the NTSI

Vishniac (1994) also derives upper and lower limits on the range of wavenumbers susceptible to the NTSI. The maximum wavenumber is a result of the width of the layer; a bending-mode instability cannot be excited if the wavelength is shorter than half the thickness of the layer. This gives

$$k_{\max} = \frac{1}{2L} \,, \tag{9.26}$$

where k is the wavenumber and L is the width of the layer. As described later in Section 9.3.1, we measure the layer width for each simulation timestep and fit a single layer width growth rate $\dot{L}$ for each suite of simulations. We use this empirically measured layer width growth rate to calculate the expected maximum wavenumber at each timestep.

We can convert this to a last resolvable time for each wavenumber, given by

$$t_{\max} = \frac{1}{2k\dot{L}} \,, \tag{9.27}$$

after which the wavenumber k can no longer be resolved.

In Section 9.2.2, we explained that Vishniac (1994) predicts a maximum growth rate of the NTSI of $\sim c_s k$. If the growth rate of the layer due to the secular evolution of the layer is larger than this, then the NTSI will be suppressed. He therefore derives a minimum wavenumber susceptible to the NTSI, given by

$$k_{\min} \gtrsim \left(\frac{V_S}{c_s} \right) R^{-1} \,, \tag{9.28}$$

where V_S is the velocity of the inflowing gas in the frame of the shock, and R is the propagation distance over which zeroth order shock properties evolve. For a stationary slab R is equal to the external gas velocity V_E multiplied by the time t. We assume that the collision is significantly supersonic so that $V_S \approx V_E$. The minimum wavenumber is then

$$k_{\min} \gtrsim \frac{1}{c_s t} \,, \tag{9.29}$$

as expected. However we find this criterion tends to predict a very high minimum wavenumber for our simulations; after the first timestep of 0.005 Myr the predicted minimum wavenumber is $1040\,\mathrm{pc}^{-1}$. Even after the complete high-resolution simulation run time of 0.1 Myr, the predicted minimum wavenumber is $52\,\mathrm{pc}^{-1}$. This would suggest none of our simulations was susceptible to the NTSI; however, we show that our simulations are unstable to the NTSI at much lower wavenumbers.

In a similar way we can use our numerical simulation of the decay of the initial

perturbation to estimate when each of our monochromatic wavenumbers will have a rate of growth due to the NTSI greater than the rate of growth of the layer. We use equation 9.24 together with our computed layer positions to estimate the amplitude of the NTSI at time t. We can compare this rate of growth to the rate at which the layer grows, as determined later in our simulations. We find that the NTSI rate of growth never exceeds the layer rate of growth in our simulation time of 1.5 Myr.

It therefore seems that the growth rate of the NTSI does not need to initially exceed the growth rate of the layer. Once the NTSI has started to grow, it can increase its growth rate until it does exceed the growth rate of the layer, but our simple model does not capture this behaviour.

9.3 Simulating the collisions and basic analysis

Once we have created our set of simulations as described in Section 9.1.2, we evolve them in time using the SEREN code described in Section 3.1. The total evolution time for each simulation is given in Table 9.4. Every 0.05 Myr we produce a snapshot file containing the SPH particle positions, velocities, masses and densities.

In Appendix B we provide cross-section and column density plots of a representative sample of our simulations. We show each simulation at a range of snapshot times to show the evolution of each system.

Each snapshot is converted to a 256^3 grid using the SPLASH visualization tool as described in Section 5.1.3. We filter this grid, as described in Section 5.2, to identify which grid cells are part of the layer. This includes finding the average density across the redundant axis for a simulation with perturbations across only one dimension.

Once we have identified the regions of the layer, we calculate the centre-of-mass positions of the layer along the collision axis, as described in Section 5.3.1, and the surface density of the layer, as described in Section 5.3.2. We take a Fourier transform of these quantities, as described in Section 5.3.3, to produce a Fourier spectrum for each. For two-dimensional perturbations this includes taking the circular average of the two-dimensional Fourier transform.

For each simulation, we construct time series consisting of the Fourier spectrum of the centre-of-mass positions and the Fourier spectrum of the surface density from every snapshot. We then calculate growth rates as a function of wavenumber at all but the first and last timestep as described in Section 5.3.4. As we have multiple realizations for each simulation, we calculate the mean and standard deviation of each rate of growth.

9.3.1 Layer width and thickness

We calculate the layer width and thickness as described in Section 5.3.5. The thickness is the distance across the layer, including only regions identified as part of the layer and not including voids within the layer structure, while the width is simply the distance from the

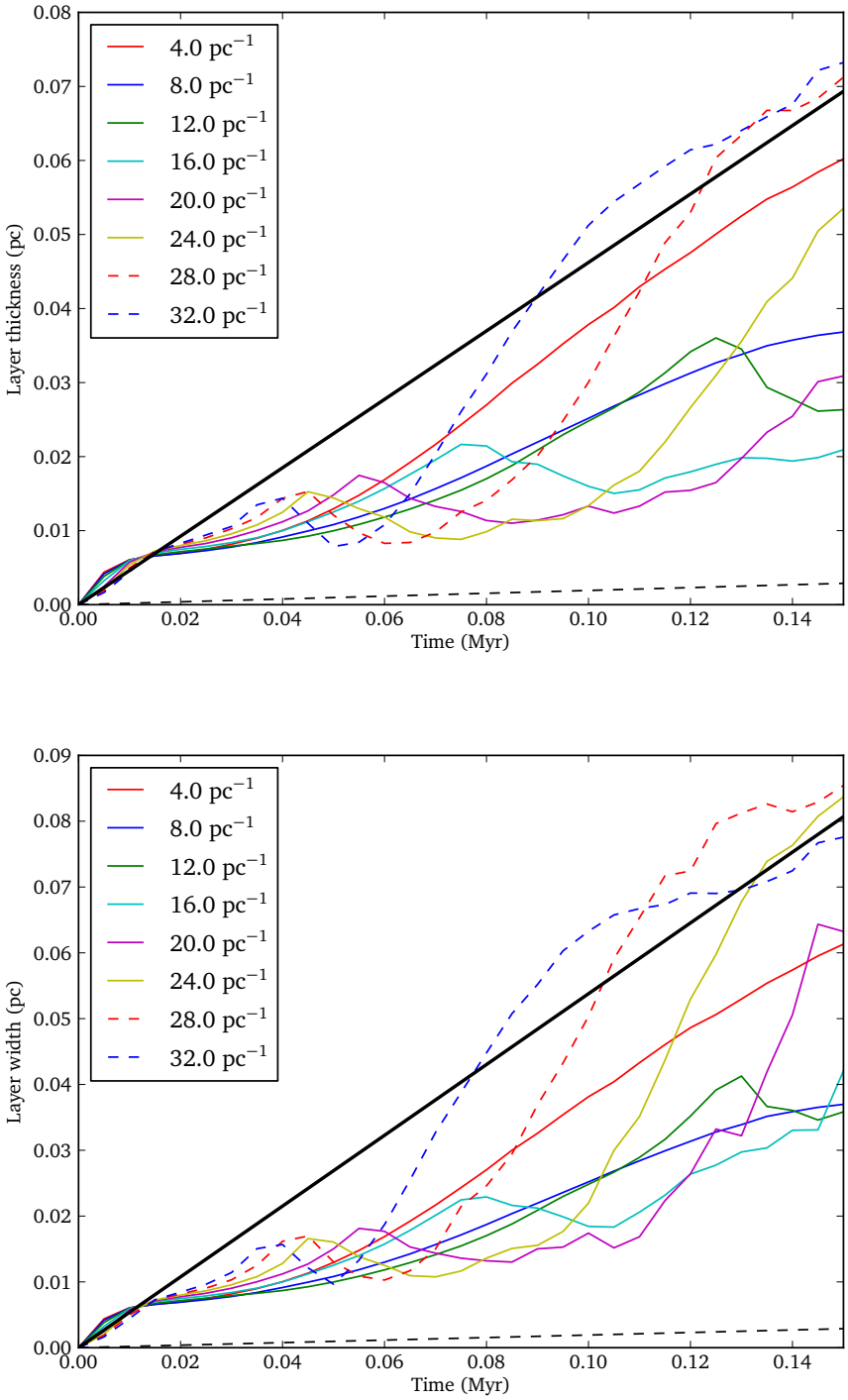


Figure 9.5 – Thickness and width, as defined in Section 5.3.5, with respect to time for the 1DLR simulations. Each coloured line represents the averaged results from the labelled set of realizations. Each set of realizations is fitted with a straight line; the solid black line indicates the average of those fits. The fits were constrained to pass through the origin. The dashed black line indicates the expected layer thickness for an idealized isothermal shock-confined layer.

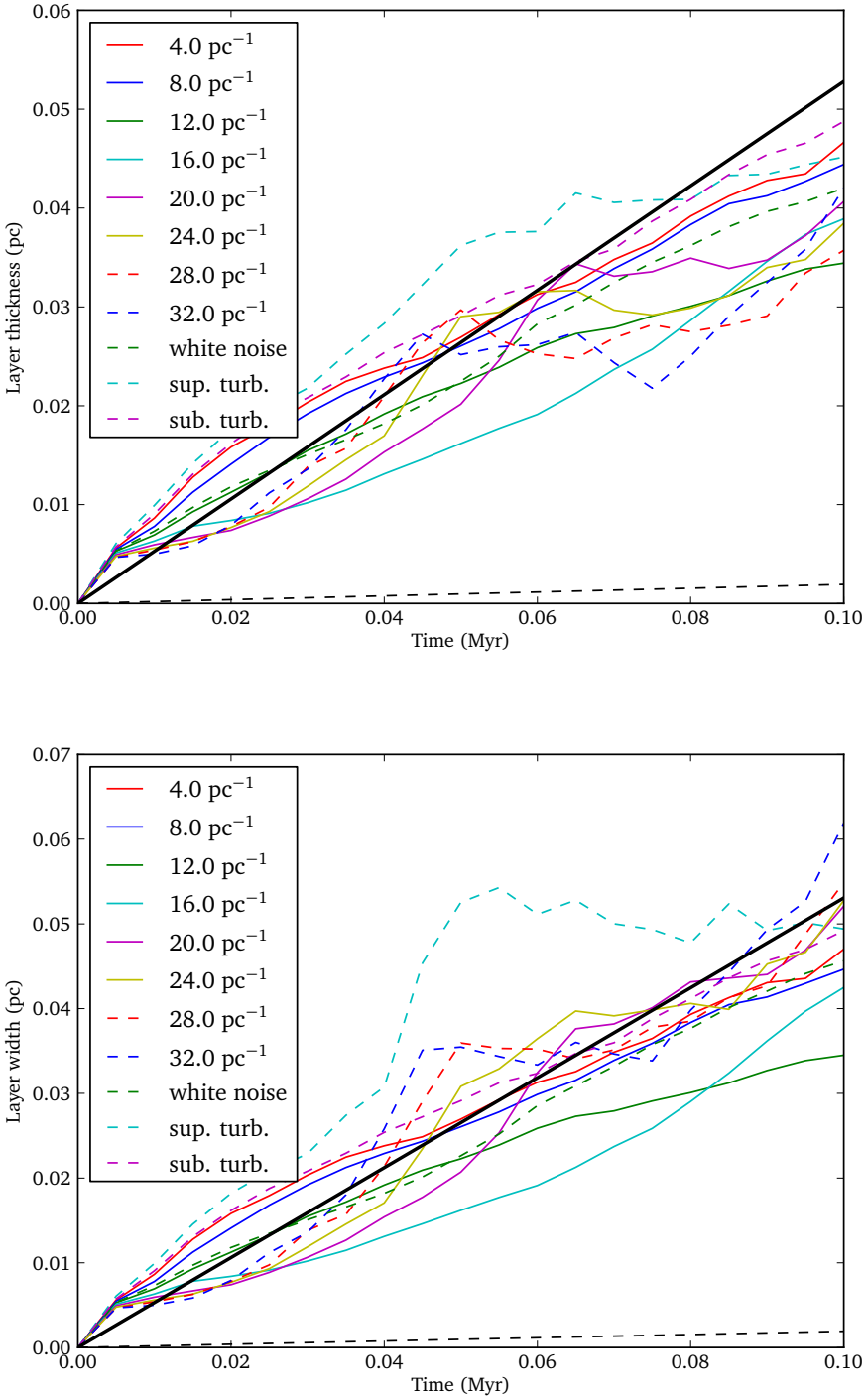


Figure 9.6 – Thickness and width, as defined in Section 5.3.5, with respect to time for the 1DHR simulations. Each coloured line represents the averaged results from the labelled set of realizations. The labels 'sub. turb.' and 'sup. turb.' indicate subsonic and supersonic turbulence. Each set of realizations is fitted with a straight line; the solid black line indicates the average of those fits. The fits were constrained to pass through the origin. The dashed black line indicates the expected layer thickness for an idealized isothermal shock-confined layer.

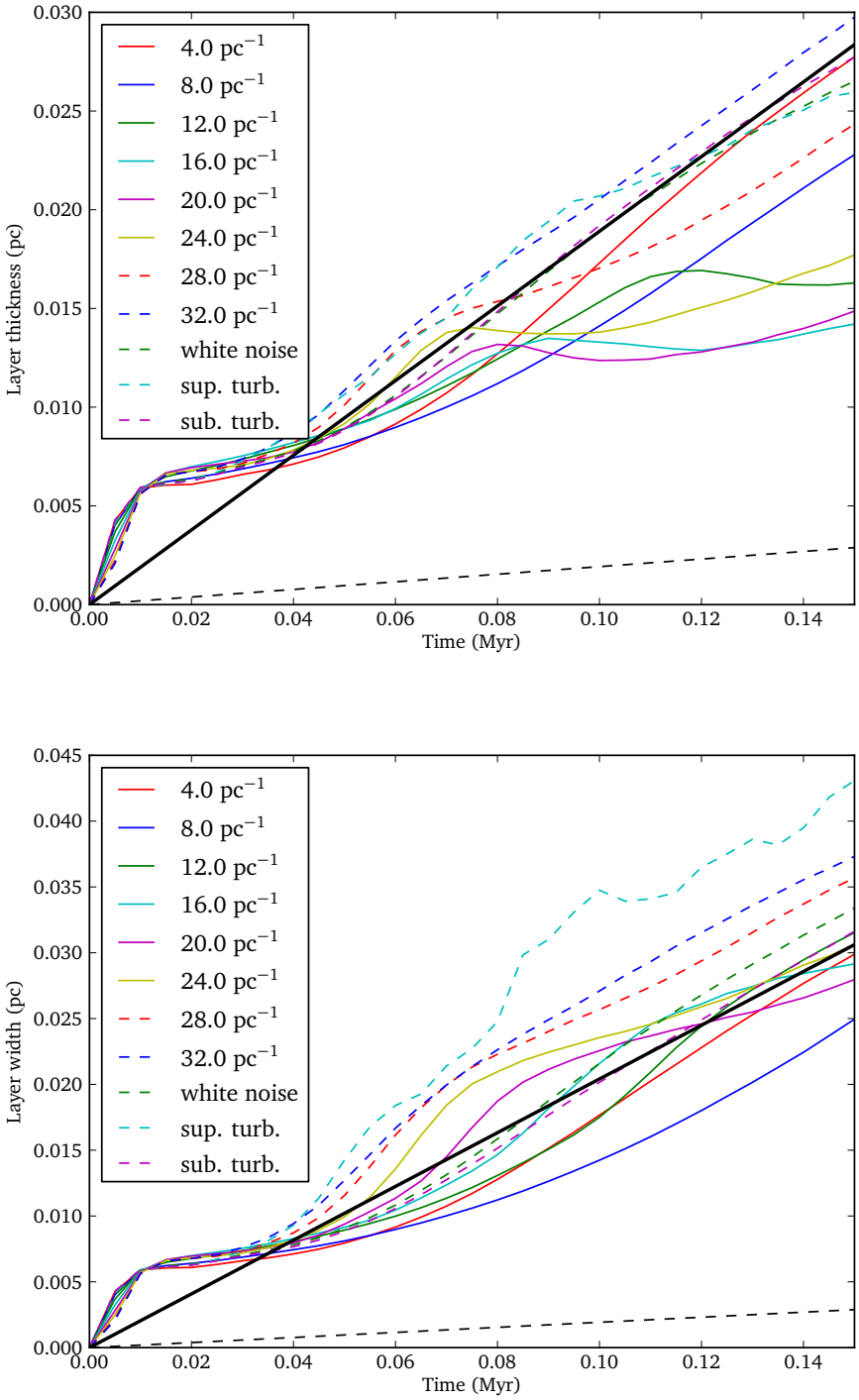


Figure 9.7 – Thickness and width, as defined in Section 5.3.5, with respect to time for the 2D simulations. Each coloured line represents the averaged results from the labelled set of realizations. The labels 'sub. turb.' and 'sup. turb.' indicate subsonic and supersonic turbulence. Each set of realizations is fitted with a straight line; the solid black line indicates the average of those fits. The fits were constrained to pass through the origin. The dashed black line indicates the expected layer thickness for an idealized isothermal shock-confined layer.

top to the bottom of the layer. Both are subsequently corrected for the bending angle of the layer.

We find that the filtering process works well to eliminate dense regions in turbulent simulations which are not part of the layer, as can be seen in Figure 12.8. This prevents a substantial overestimate of the width of the layer. However at higher wavenumbers we find that the layer can contain substantial voids, as can be seen in Figure 10.3. This causes a relatively low measurement of the thickness which does not reflect of the large bending angle of the layer.

Our main interest in the width or thickness of the layer is in estimating a maximum resolvable wavenumber. We wish to measure the total expanded size of the layer, including voids, to estimate the smallest wavelength perturbation a layer can have, and so we use the width in all further analysis.

We take the mean value of the width and thickness across the layer to give a single width and thickness for each snapshot. We then take the median value across each set of realizations. These are shown in Figure 9.5, 9.6 and 9.7 for the 1DLR, 1DHR and 2D simulations respectively. We fit a single line, passing through zero at zero time, to the results for each set of realizations, giving a rate of layer width or thickness increase $\dot{L}$ with respect to time. Finally we take a simple average value of $\dot{L}$ by averaging the fitted values for each set of realizations.

As described in Section 6.2.1, for an isothermal shock in the absence of instabilities the layer density should be equal to the background density ρ_0 multiplied by the square of the Mach number $\mathscr{M}$. The layer accretes surface density at a rate $2\rho_0 V_\mathrm{E}$. The layer thickness is then

$$L_\mathrm{layer} = \frac{2V_\mathrm{E}t}{\mathscr{M}^2}\,. \tag{9.30}$$

As shown in Figures 9.6, 9.5 and 9.7 the rate at which our layers grow is considerably in excess of this idealized model. There is not a strong correlation between the type or wavenumber of perturbation and the layer thickness growth rate, suggesting that the NTSI is not directly responsible. It is not clear whether this is due to a physical or numerical effect; however we note the layer grows more rapidly in the higher resolution simulation than in the lower resolution simulation. This suggests that the expansion of the layer is a physical effect which is being increasingly well-resolved as the ability to resolve smaller features and higher wavenumbers improves.

Using our fits to the rate at which the width of the layer grows, we can estimate the maximum resolvable wavenumber at each timestep. As described in Section 9.2.3, the wavelength of the perturbation cannot be smaller than twice the layer thickness. The maximum resolvable wavenumbers for the 1DLR, 1DHR and 2D simulations are shown in Figures 9.8 and 9.9.

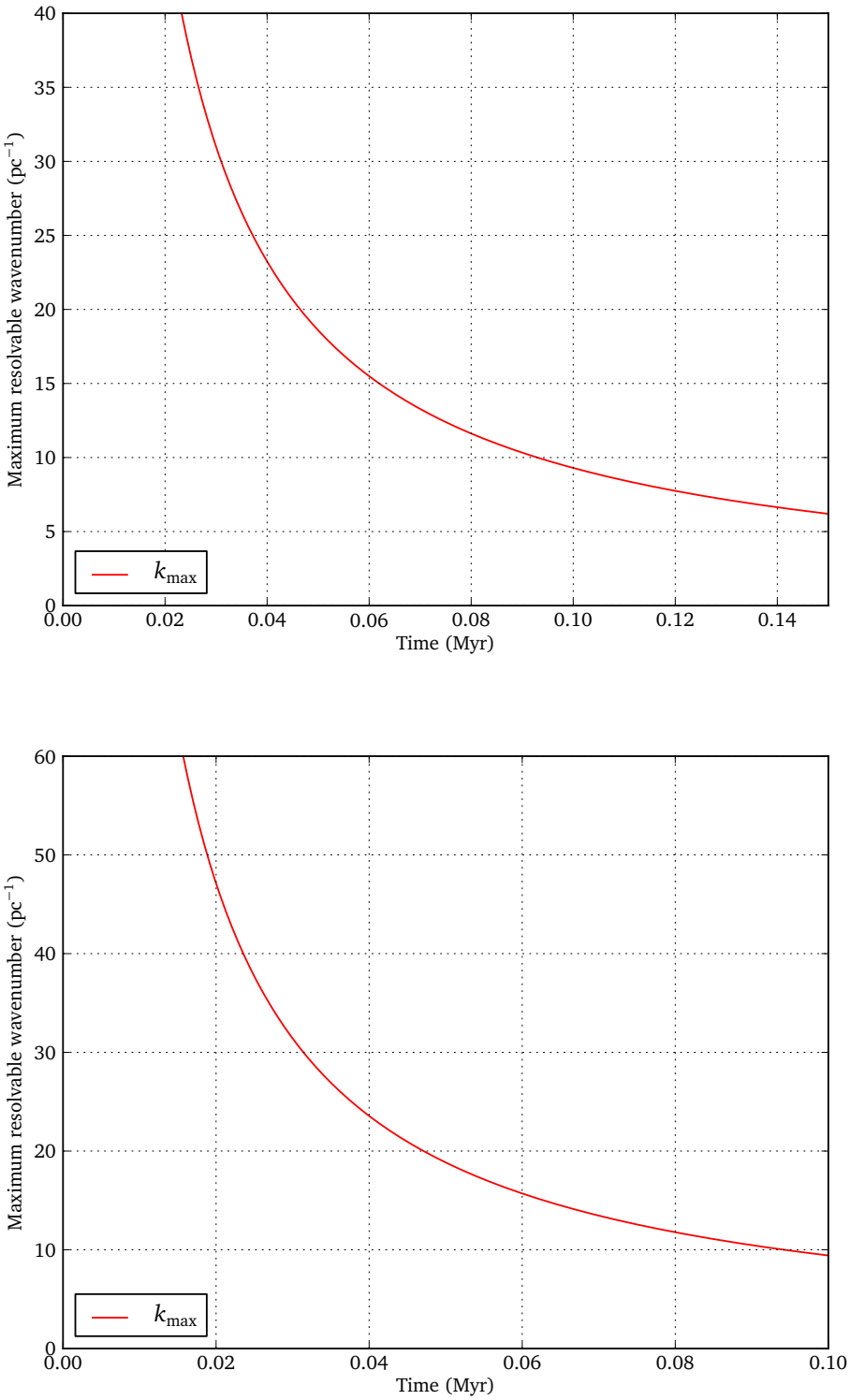


Figure 9.8 – Maximum resolvable wavenumber for one-dimensional simulations. The upper plot shows results from the 1DLR simulations; the lower plot shows the results from the 1DHR simulations.

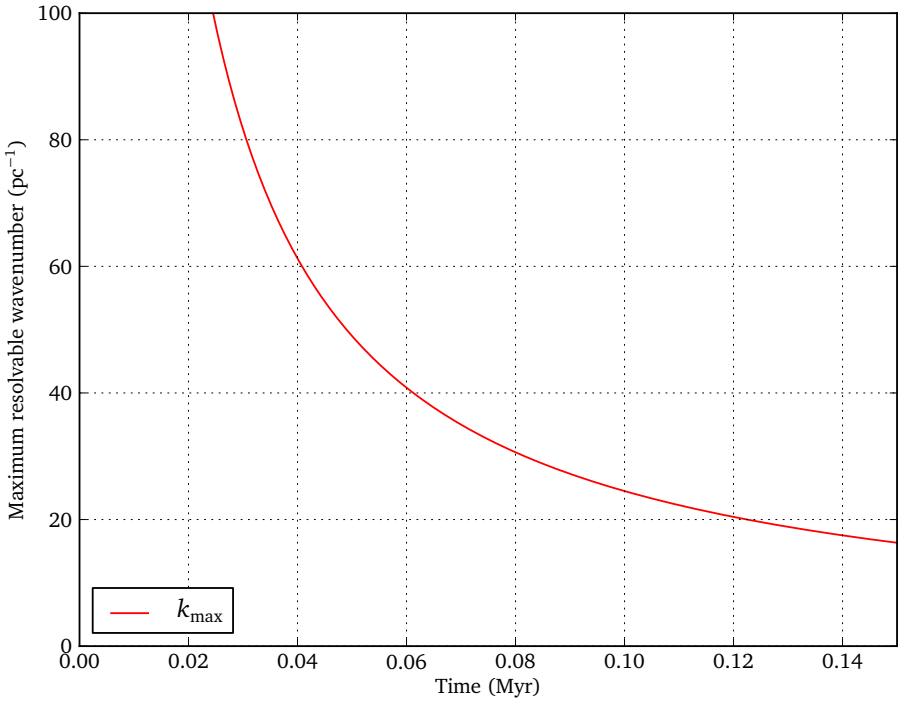


Figure 9.9 – Maximum resolvable wavenumber for two-dimensional simulations from the 2D simulation suite.

Chapter 10

NTSI – monochromatic perturbations

As described in Section 9.1.1, we conduct a suite of simulations with one-dimensional and two-dimensional monochromatic perturbations. Each simulation contains an initial velocity perturbation at a single wavenumber. We run five realizations per wavenumber with different settled SPH particle distributions.

In all our monochromatic simulations the amplitude of the triggered perturbation is much larger than the amplitude of any other perturbation. We therefore consider only the triggered perturbation in each monochromatic simulation. We consider the rates of growth of two quantities relating to the triggered perturbation.

The first is the centre-of-mass position, as described in Section 5.3.1. As can be seen in Figures 10.1 to 10.3, the centre-of-mass position is a good tracer of perturbation amplitude for our monochromatic simulations. We take the Fourier transform of this tracer for each timestep, and calculate the rates of growth of the perturbation amplitude at each timestep as described in Chapter 5. This rate of growth is approximately the velocity of the perturbation peak.

We also consider the column density, as described in Section 5.3.2. As can be seen most clearly in Figures 10.2 and 10.3, the column density acts as a tracer of the peaks of the perturbation. The NTSI creates growth in column density as shear in the layer moves extra mass towards the peak of the perturbation. For this reason the column density may be a better tracer for the NTSI, as it is not affected by the initial perturbation. However the column density is also affected by both gravity and pressure forces. Where the NTSI is present, the column density will approximate a Dirac comb function with a wavenumber twice that of the initial perturbation, as the column density is increased at both peaks and troughs. Taking a Fourier transform produces another Dirac comb with the same wavenumber; we use the amplitude of the first peak. We calculate the rate of growth of this amplitude at each timestep. The physical meaning of this rate of growth is less clearly defined than that of the rate of growth from layer position, but relates to the amplitude of the column density peaks.

We have five realizations at each wavenumber, and we average these quantities across

these realizations. We can then extract the rate of growth of the relevant wavenumber from each simulation to produce a combined dataset. We therefore have a rate of growth of centre-of-mass position for each of the wavenumbers we have triggered, where each value represents the average of five realizations with that perturbation. Similarly we have a rate of growth of column density corresponding to each of the wavenumbers we have triggered, except that the wavenumbers are doubled for the reasons explained above.

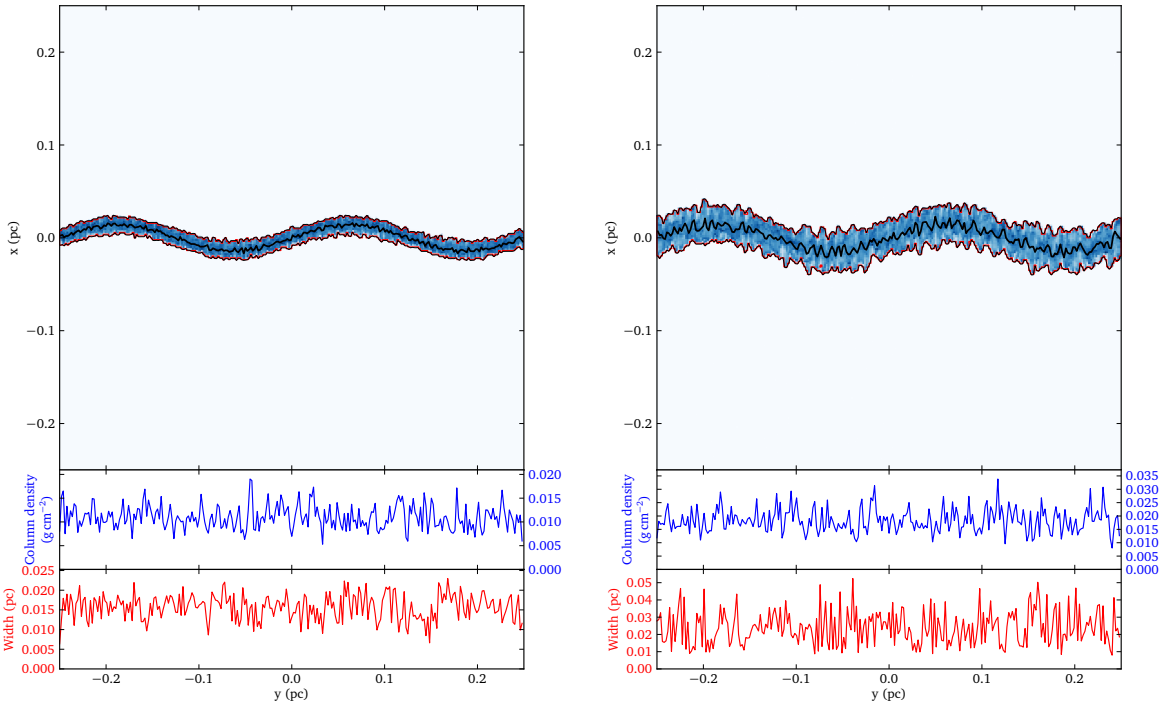
10.1 One-dimensional monochromatic perturbations

We first consider the results from one-dimensional simulations. As described in Section 9.1.1, we conduct a suite of simulations with one-dimensional monochromatic perturbations. Each simulation contains a single wavenumber from the set 4, 8, 12, 16, 20, 24, 28 and $32\,\mathrm{pc}^{-1}$. We run five realizations per wavenumber with different settled SPH particle distributions.

The upper plot of Figures 10.4 and 10.5 shows the rates of growth of centre-of-mass position for each wavenumber. The results from low-resolution and high-resolution simulations show the same patterns of growth. The initial rates of growth, showing the decay of the initially imposed perturbation, is close to that predicted in Section 9.2.1. As expected, the decay is slightly faster than predicted in our simple model, which does not account for the growth of the layer thickness. The smallest wavenumbers do not grow significantly in the simulation time; we can therefore state that the wavenumbers $k \lesssim 10\,\mathrm{pc}^{-1}$ are not susceptible to the NTSI.

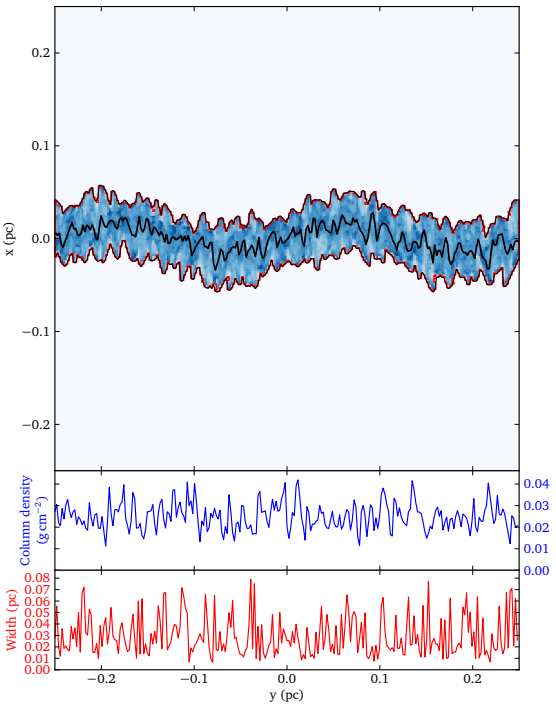
Wavenumbers which are susceptible to the NTSI begin to grow, with the largest wavenumbers growing first. However these largest wavenumbers are also the first to saturate and then rapidly decay. This is due to the growth of the layer, which prevents the growth of wavelengths shorter than twice the width of the layer. In our simulation only a very narrow window of time around approximately 0.02 Myr is not significantly affected by either the initial perturbation or the later saturation of the higher wavenumbers.

We note that for wavenumbers that do grow, the rate of growth of centre-of-mass position peaks at around $0.6\,\mathrm{km\,s}^{-1}$, noting that $1\,\mathrm{km\,s}^{-1} \approx 1\,\mathrm{pc\,Myr}^{-1}$. This is considerably greater than the sound speed of $0.188\,\mathrm{km\,s}^{-1}$, indicating that the motions of the layer are supersonic. This explains why we are able to excite the NTSI at much lower wavenumbers than predicted by Vishniac (1994), and as described in Section 9.2.3. In Vishniac's analysis the motion of the layer is not allowed to become supersonic, and the maximum amplitude of the perturbation is not allowed to exceed the wavelength to maintain the constraints on small bending angles. The result is to limit the maximum rate of growth to approximately the sound speed c_s, which we clearly exceed. We also find that the amplitude of perturbations does grow larger than the wavelength of the perturbation, yielding large bending angles.

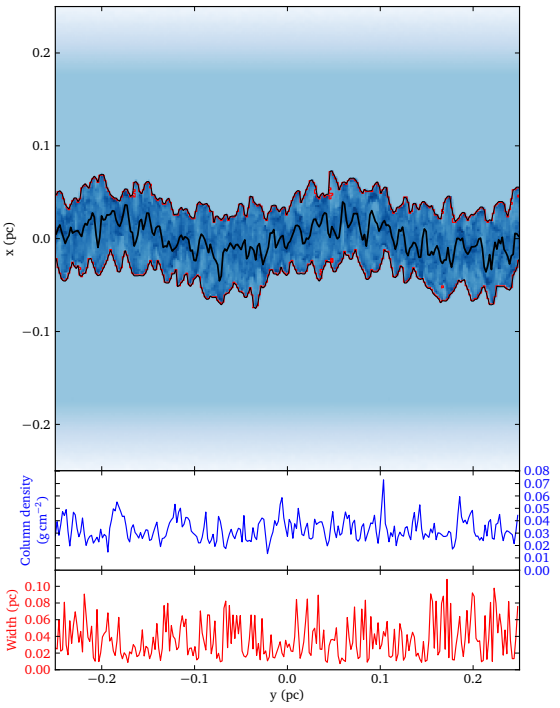


(a) $t = 0.02\,\mathrm{Myr}$

(b) $t = 0.04\,\mathrm{Myr}$

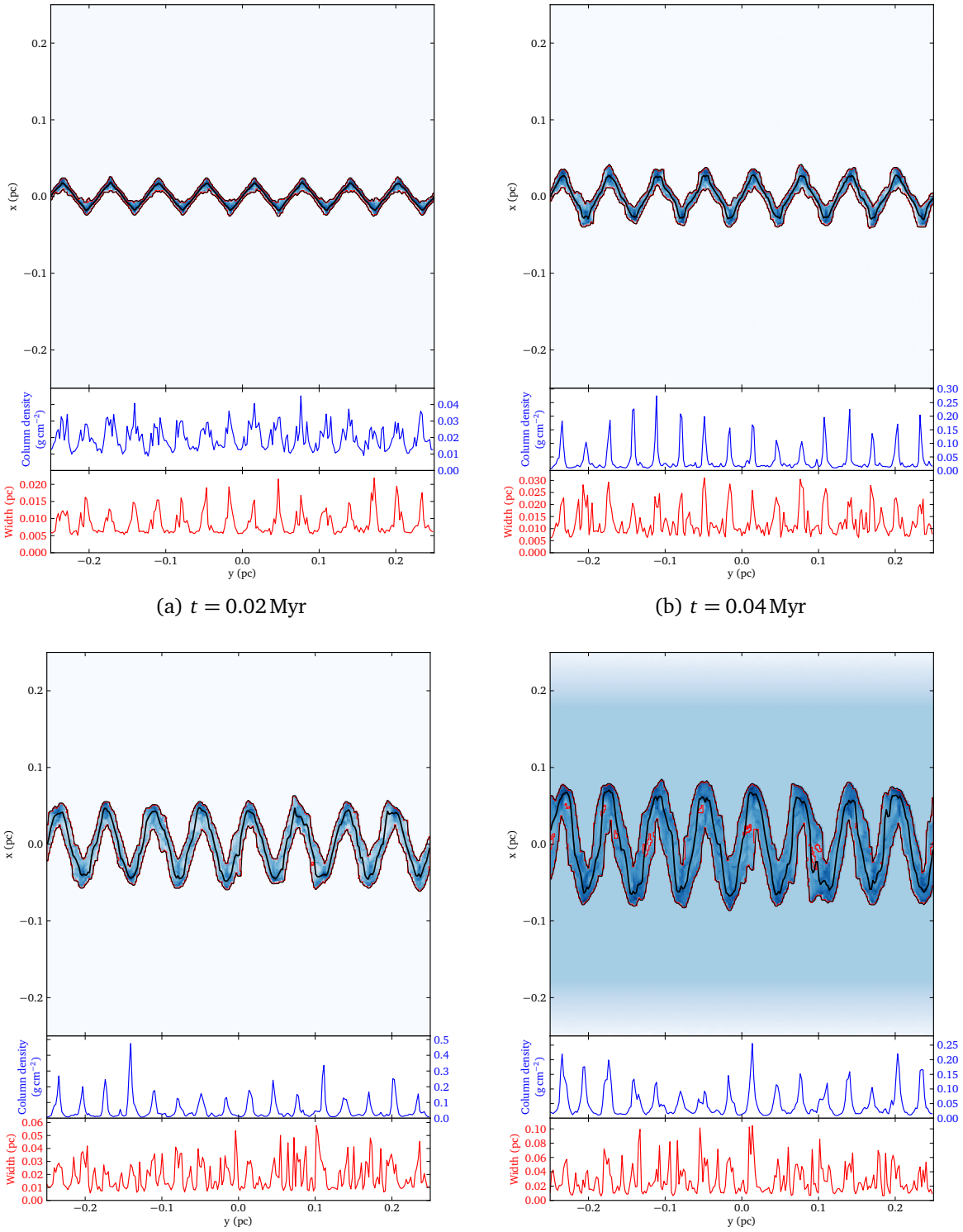


(c) $t = 0.06\,\mathrm{Myr}$



(d) $t = 0.08\,\mathrm{Myr}$

Figure 10.1 – Layer diagrams for 1DHR simulation of a monochromatic perturbation of wavenumber $4\,\mathrm{pc}^{-1}$. Colours indicate density but are not consistent across plots. The thick black line shows the centre-of-mass position of the layer. Red lines indicate the edges of groups of layer cells. The thin black lines show the upper and lower extent of the layer. The approach of the rarefaction wave is visible in the final diagram. Column density and layer width as a function of y position are also shown.



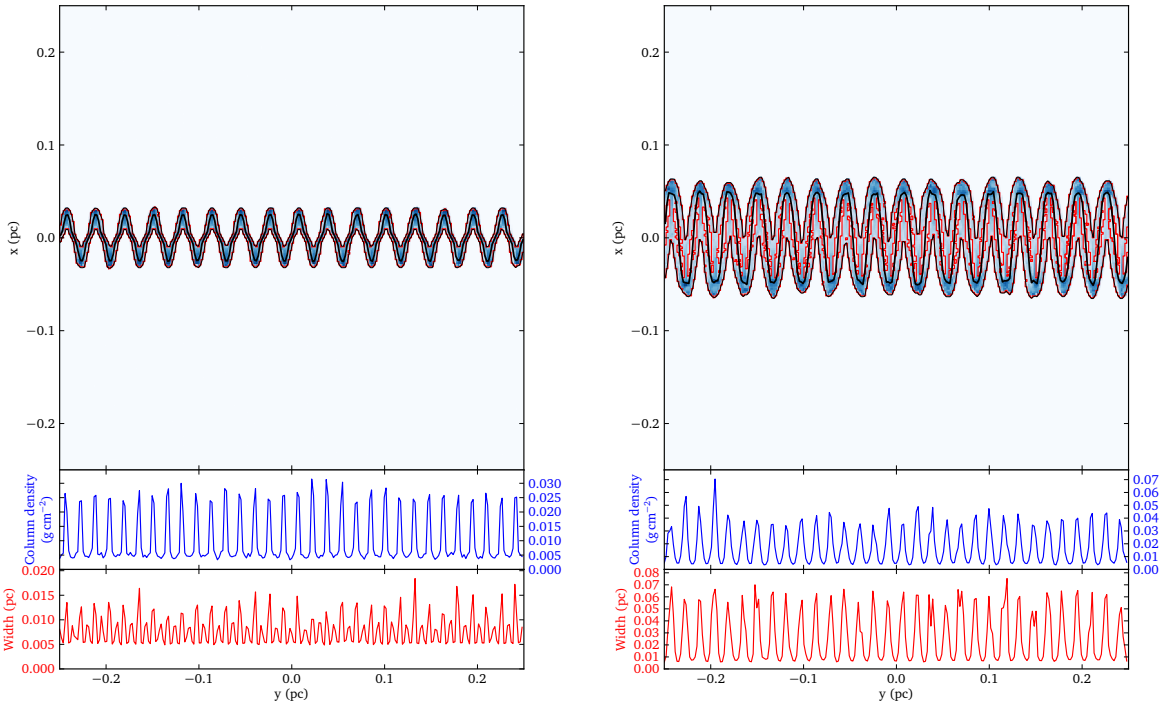
(a) $t = 0.02\,\mathrm{Myr}$

(b) $t = 0.04\,\mathrm{Myr}$

(c) $t = 0.06\,\mathrm{Myr}$

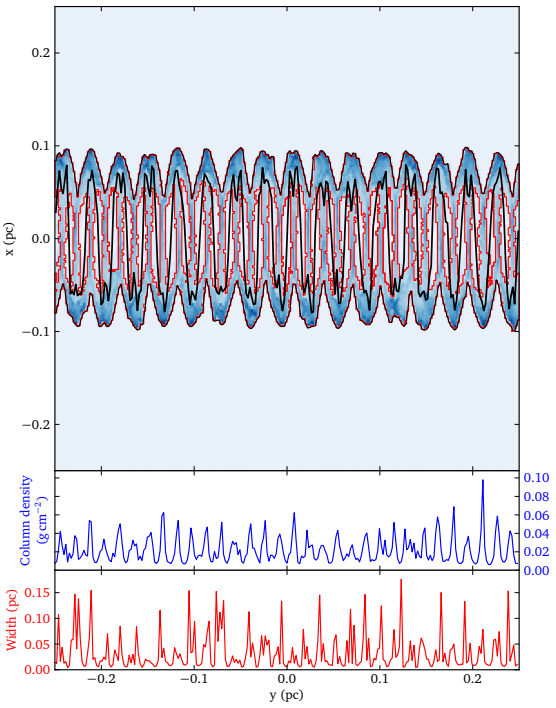
(d) $t = 0.08\,\mathrm{Myr}$

Figure 10.2 – Layer diagrams for 1DHR simulation of a monochromatic perturbation of wavenumber $16\,\mathrm{pc}^{-1}$. Colours indicate density but are not consistent across plots. The thick black line shows the centre-of-mass position of the layer. Red lines indicate the edges of groups of layer cells. The thin black lines show the upper and lower extent of the layer. The approach of the rarefaction wave is visible in the final diagram. Column density and layer width as a function of y position are also shown.

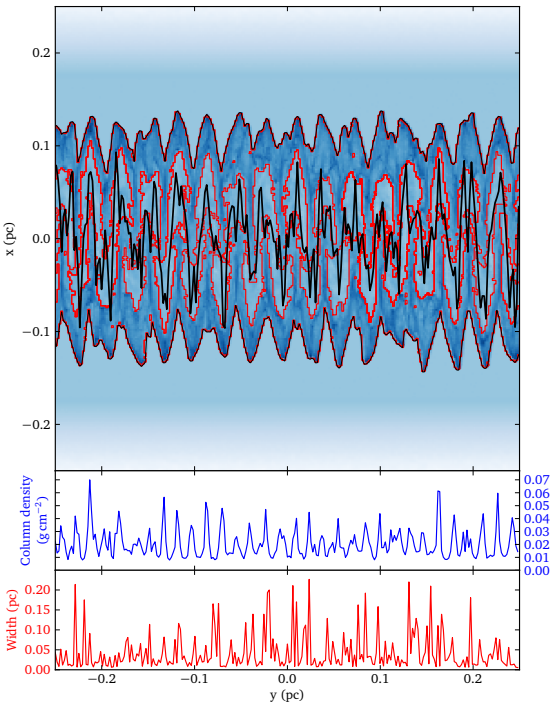


(a) $t = 0.02\,\mathrm{Myr}$

(b) $t = 0.04\,\mathrm{Myr}$



(c) $t = 0.06\,\mathrm{Myr}$



(d) $t = 0.08\,\mathrm{Myr}$

Figure 10.3 – Layer diagrams for 1DHR simulation of a monochromatic perturbation of wavenumber $32\,\mathrm{pc}^{-1}$. Colours indicate density but are not consistent across plots. The thick black line shows the centre-of-mass position of the layer. Red lines indicate the edges of groups of layer cells. The thin black lines show the upper and lower extent of the layer. The approach of the rarefaction wave is visible in the final diagram. Column density and layer width as a function of y position are also shown.

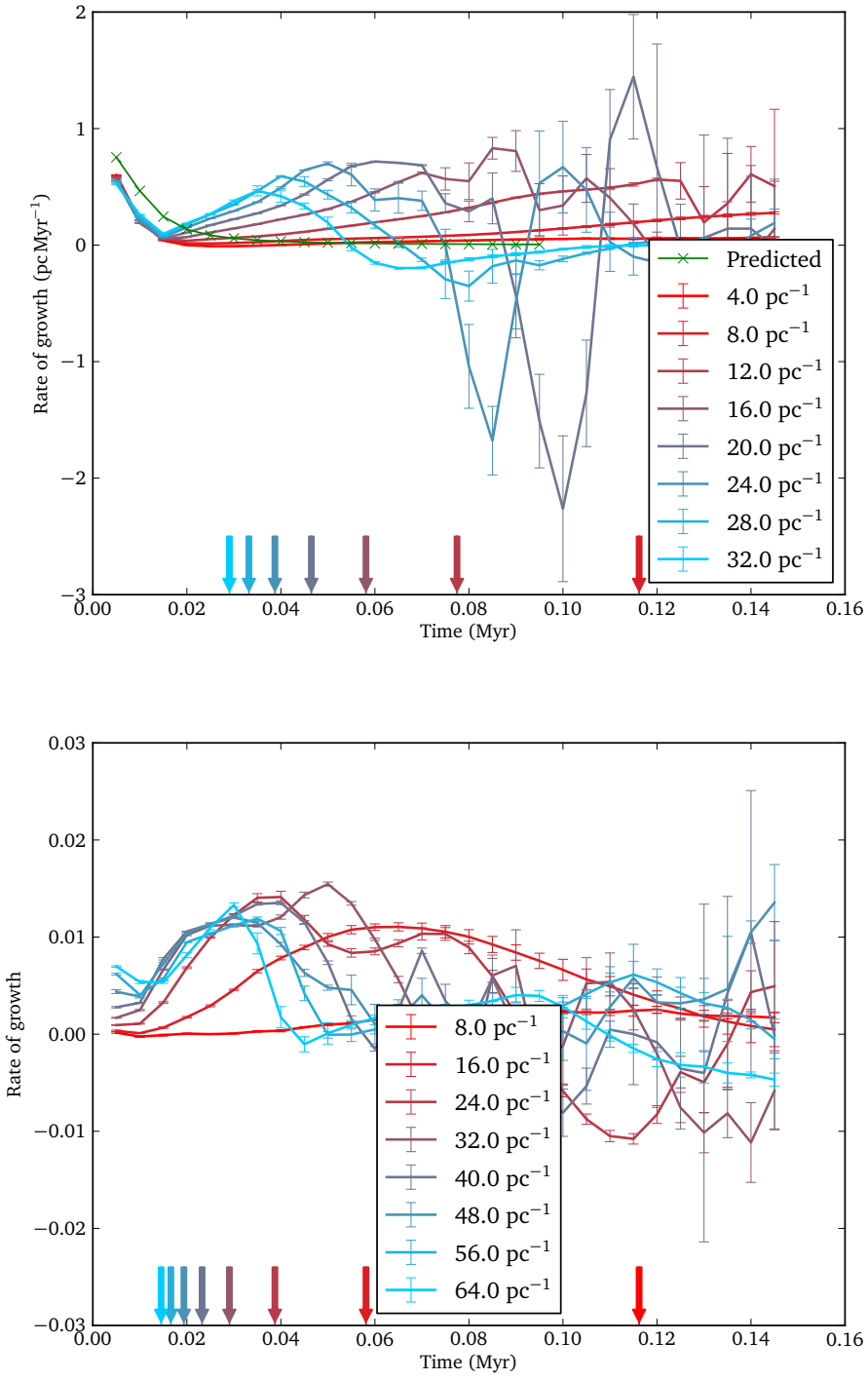


Figure 10.4 – The rates of growth for monochromatic perturbations with time for each triggered wavenumber from 1DLR simulations. The upper plot shows rates of growth of centre-of-mass position, while the lower plot shows the rates of growth of column density. The green line indicates the values predicted in Section 9.2.1. Each other line represents the rate of growth of the relevant wavenumber averaged over five realizations where only that wavenumber was triggered; error bars show the standard deviation across realizations. The arrows indicate the last resolvable time for the correspondingly coloured wavenumber, as defined in Section 9.2.3.

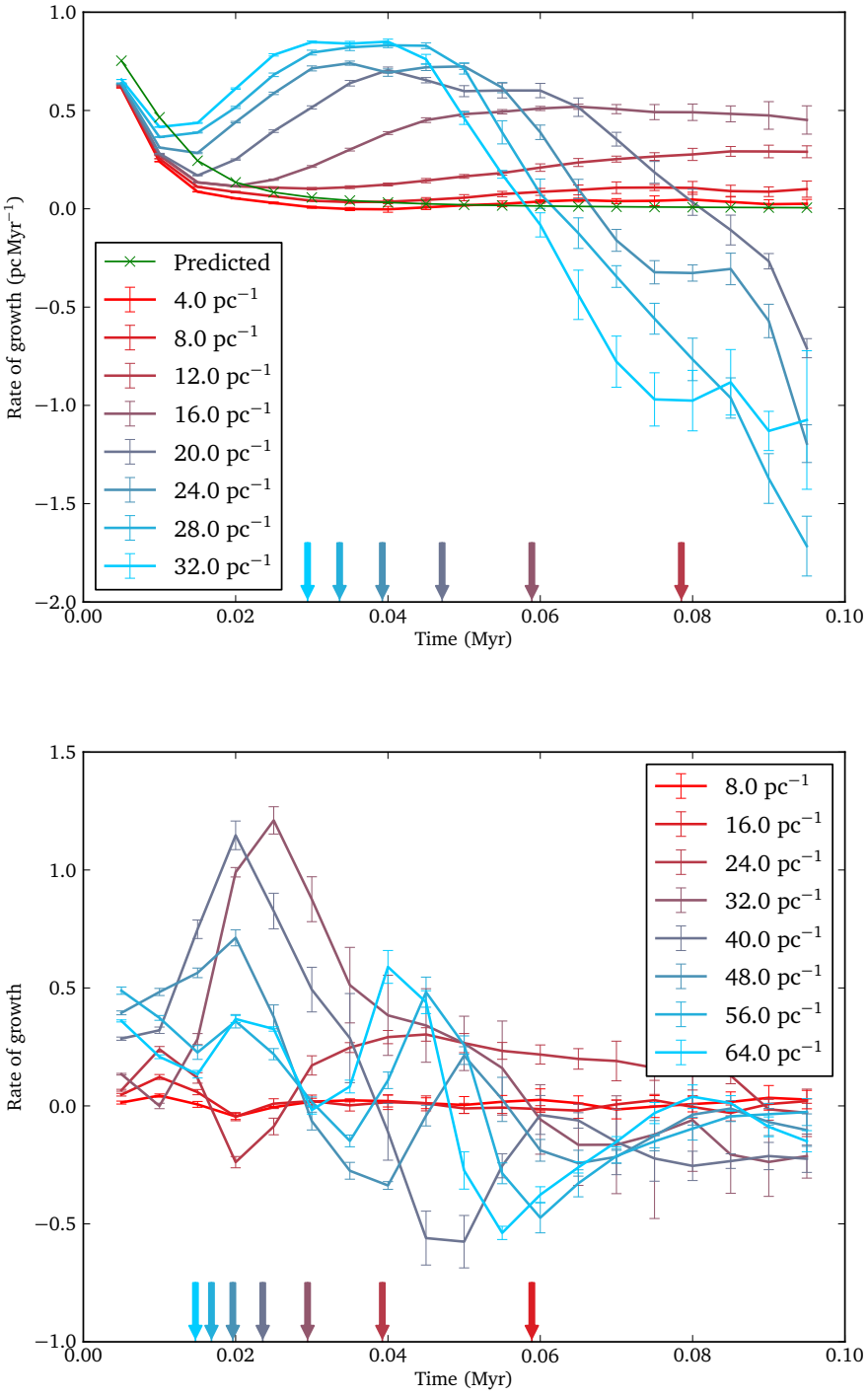


Figure 10.5 – The rates of growth for monochromatic perturbations with time for each triggered wavenumber from 1DHR simulations. The upper plot shows rates of growth of centre-of-mass position, while the lower plot shows the rates of growth of column density. The green line indicates the values predicted in Section 9.2.1. Each other line represents the rate of growth of the relevant wavenumber averaged over five realizations where only that wavenumber was triggered; error bars show the standard deviation across realizations. The arrows indicate the last resolvable time for the correspondingly coloured wavenumber, as defined in Section 9.2.3.

The lower plot of Figures 10.4 and 10.5 shows the rates of growth of column density for each wavenumber. Here we can see that the smallest wavenumbers do not create any significant perturbations in column density. This is expected as we have shown these wavenumbers are not susceptible to the NTSI, and the initial velocity perturbation should not cause perturbations in column density. However no clear pattern can be seen for the rest of the wavenumbers. The low-resolution simulations produce very little growth at any wavenumber, while even in the high-resolution simulations it is not obvious from the column density perturbations that higher wavenumbers grow earlier or faster than lower wavenumbers.

The problem with column density perturbations may be that they are transient. We note that, as wavenumbers become saturated at later times, the column density perturbations tend towards zero. Once formed, a column density perturbation has to be continually reinforced to survive. If the instability feeding it ceases, it will expand laterally under its own pressure forces, removing the perturbation. By contrast the position of a layer is a much more persistent quantity. The rate of growth of a perturbation in the layer will decay only slowly, as the layer accretes momentum, even after the forces creating that perturbation are removed.

We find throughout that data derived from the column density is noisy and does not show clear trends. For this reason, we do not continue using it, and restrict ourselves only to data derived from the centre-of-mass position of the layer.

Figures 10.6 and 10.7 show the same rates of growth of the centre-of-mass position, but as a function of wavenumber. The maximum wavenumber, as predicted in Section 9.2.3, is also shown. The approximate gradient of the slopes can be compared to the relationship

$$\omega \propto k^{3/2} \,, \tag{10.1}$$

where ω is the rate of growth,as predicted by Vishniac (1994). We find that at the earliest timestep, the distribution of growth rate is approximately equal across all wavenumbers. However, the growth rate of the lower wavenumbers rapidly decreases, increasing the gradient. The gradient then appears to become steeper than the predicted relationship. We obtain the gradient by fitting straight lines to appropriate ranges of wavenumbers.

Our estimated maximum resolvable wavenumber appears to be quite accurate. Wavenumbers above the maximum resolvable wavenumber rapidly fall below the slope of the other wavenumbers at the same timestep. For each timestep we fit power law slopes to the rates of growth as a function of wavenumber, including only wavenumbers lower than the maximum resolvable wavenumber. Any points with negative growth rates cannot be included in the power law fit and are disregarded. The index α of the power law fit at each timestep is plotted in Figure 10.8, together with the number of points used

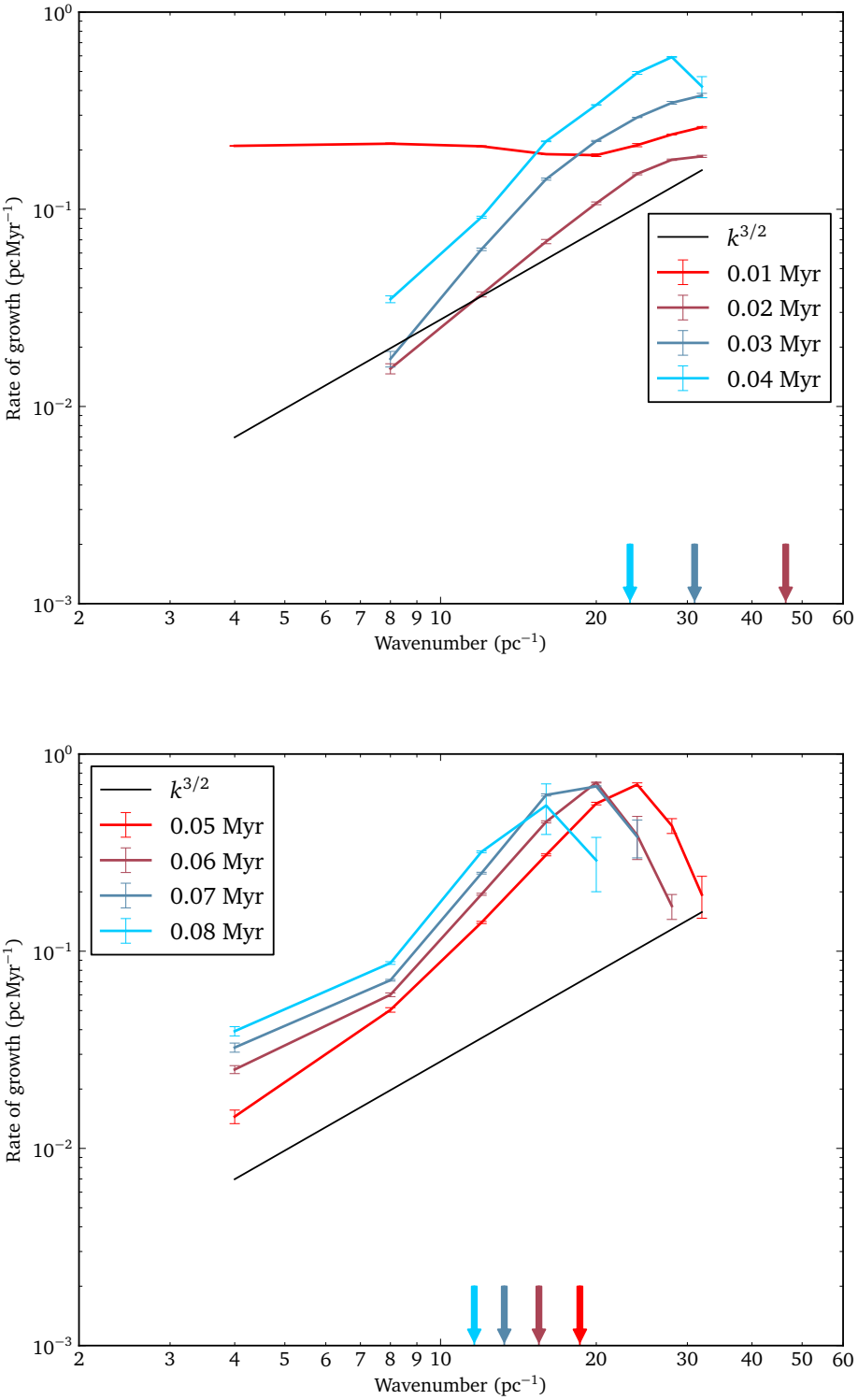


Figure 10.6 – The rates of growth of centre-of-mass position for monochromatic perturbations as a function of wavenumber from 1DLR simulations. Each line represents results from a different timestep; error bars show the standard deviation across realizations. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep, as defined in Section 9.2.3.

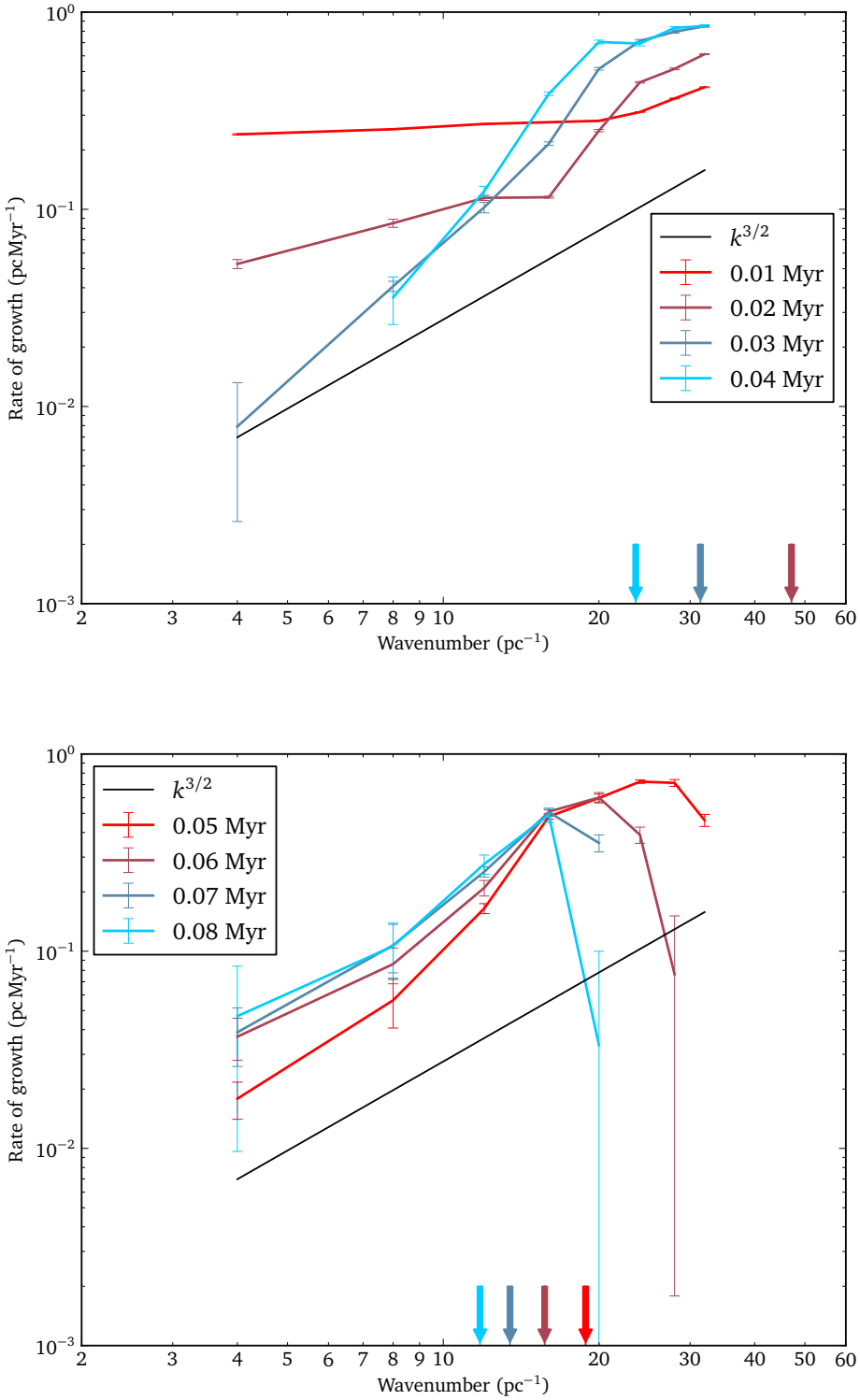


Figure 10.7 – The rates of growth of centre-of-mass position for monochromatic perturbations as a function of wavenumber from 1DHR simulations. Each line represents results from a different timestep; error bars show the standard deviation across realizations. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep, as defined in Section 9.2.3.

to construct each fit; later timesteps are heavily affected by the reduction in maximum resolvable wavenumber.

For both the low-resolution and high-resolution simulations, the power law index is initially low. This is expected; at early times the initial velocity perturbation, which has equal amplitude for all wavenumbers, is dominant. The power law index then rapidly rises. For the low-resolution simulations, the growth rates settles at $\alpha \sim 2$, which is somewhat higher than the predicted value of $\alpha = 1.5$. For the high-resolution simulations, the index continues to rise, reaching $\alpha \sim 3$, before falling back down to lower values as the reliability of the fits decreases.

Examination of Figures 10.4 and 10.5 can help explain these trends. After approximately 0.02 Myr, the largest wavenumbers start to saturate and decrease. These decreases correspond with the time at which these wavenumbers are no longer resolvable, and are therefore omitted from the fits. This leaves the smallest wavenumbers, which have approximately zero rates of growth, and a decreasing number of growing wavenumbers which in turn become unresolvable and are discounted. As a result, the later evolution of the power index should be considered extremely suspect.

We therefore restrict our analysis to the earliest times where the NTSI has become dominant but no wavenumbers have yet become saturated or unresolvable; this is approximately 0.025 Myr to 0.035 Myr for the low-resolution simulations and 0.02 Myr to 0.03 Myr for the high-resolution simulations. For the low-resolution simulations we see an index of approximately 2, which is still higher than the predicted value, and for the high-resolution simulations we find an index of between 1.3 and 2.4. We have therefore demonstrated that higher wavenumbers grow faster, as predicted by Vishniac (1994), but we have not confirmed the exact index.

We can compare the results of low-resolution and high-resolution simulations. In general we find good agreement, although as expected the low-resolution simulations show less small-scale detail. Since the limiting factor in resolving wavenumbers is the physical thickness of the layer and not the numerical resolution, the ability to resolve higher wavenumbers is comparable between low-resolution and high-resolution simulations. This validates our two-dimensional simulations, which are performed only at low resolution due to computational limitations.

10.2 Two-dimensional monochromatic perturbations

We conduct a suite of simulations with two-dimensional monochromatic perturbations. We do not have the computational resources to run two-dimensional simulations at the same resolution as our 1DHR simulations. Instead, we run our 2D simulation suite at the same resolution as the 1DLR simulation suite.

Each simulation contains a single wavenumber perturbation. We excite the product of

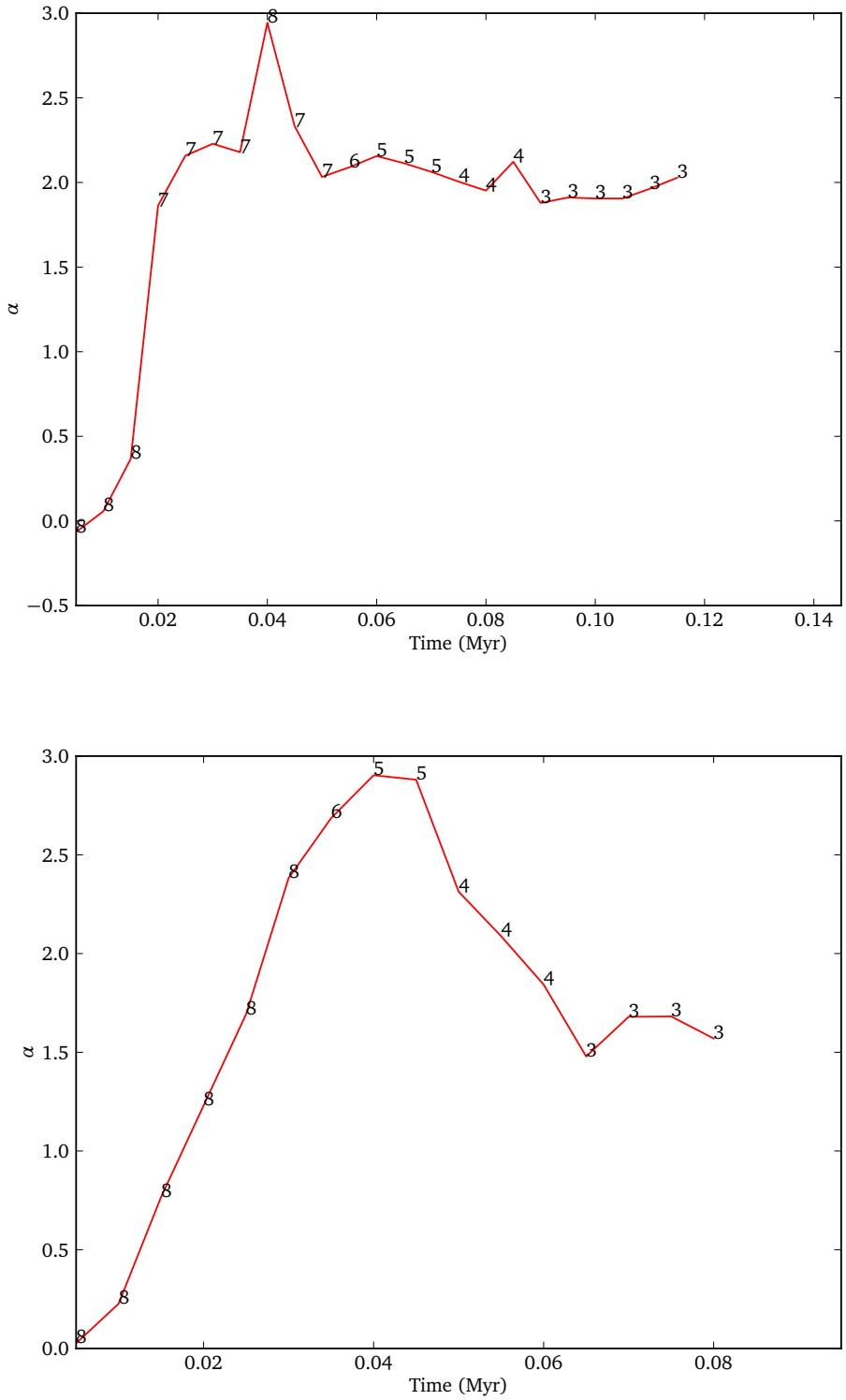


Figure 10.8 – Indices of power law fits to rates of growth of centre-of-mass position. The upper plot shows results from the monochromatic 1DLR simulations; the lower plot shows the results from the monochromatic 1DHR simulations. The number of points used to construct each fit is also shown; fits constructed with fewer points are less reliable.

two sinusoidal perturbations along the y and z axes respectively which have the same wavenumber k_{2D}, as described in Section 8.2.2. The corresponding Fourier wavevectors $\boldsymbol{k}(k_y, k_z)$ are the wavevectors $\boldsymbol{k}(\pm k_{2D}, \pm k_{2D})$, with approximately equal amplitude in each wavevector. The equivalent of the one-dimensional wavenumber k is given by the magnitude of the wavevector, such that

$$k = \sqrt{2}\, k_{2D} \,. \tag{10.2}$$

For k_{2D} we use the set of wavenumbers 4, 8, 12, 16, 20, 24, 28 and $32\,\mathrm{pc}^{-1}$. This gives the set of wavenumbers 5.66, 11.31, 16.97, 22.63, 33.94, 39.60 and $45.25\,\mathrm{pc}^{-1}$ for k. As before, we run five realizations per wavenumber with different settled SPH particle distributions. For each simulation we extract the Fourier amplitude of the relevant wavenumber $A(k_{2D}, k_{2D})$ directly from the two-dimensional Fourier field, and so we do not need to perform a circular average.

The upper plot of Figure 10.9 shows the rates of growth of centre-of-mass position for each wavenumber. The initial rates of growth are less than that predicted in Section 9.2.1. This differs from the one-dimensional simulations where the initial rates of growth are only slightly less than predicted. Perturbations in two dimensions lose an initially imposed velocity perturbation faster than perturbations in one dimension. This particularly affects large wavenumbers; small wavenumbers decay more slowly.

Our model assumes that there are no transverse motions carrying momentum away from the peak of a perturbation. This may be more reasonable for a one-dimensional perturbation than for a two-dimensional perturbation. For a one-dimensional perturbation only motion along one axis will reduce the velocity of the peak. For a two-dimensional perturbation, transverse motions along two axes will carry away momentum and so reduce the velocity of the peak. The total absolute momentum of the perturbation, found by integrating the absolute value of the initial velocity perturbation over the plane of the layer, is lower than for the one-dimensional perturbations. This makes it more likely that transverse motions will be able to carry away momentum from the peak. The result is that the absolute value of rates of growth for two-dimensional monochromatic perturbations are lower than for the one-dimensional case.

As expected the higher wavenumbers grow more rapidly than the lower wavenumbers. The smallest wavenumber of $5.66\,\mathrm{pc}^{-1}$ does not grow significantly during the simulation time, while the $11.3\,\mathrm{pc}^{-1}$ wavenumber does begin to grow. All higher wavenumbers grow rapidly before saturating at a few times the sound speed and subsequently decaying. This matches the results for the one-dimensional perturbations, where only wavenumbers greater than approximately $10\,\mathrm{pc}^{-1}$ were susceptible to the NTSI.

We calculate the last resolvable time as described in Section 9.2.3. We find that the

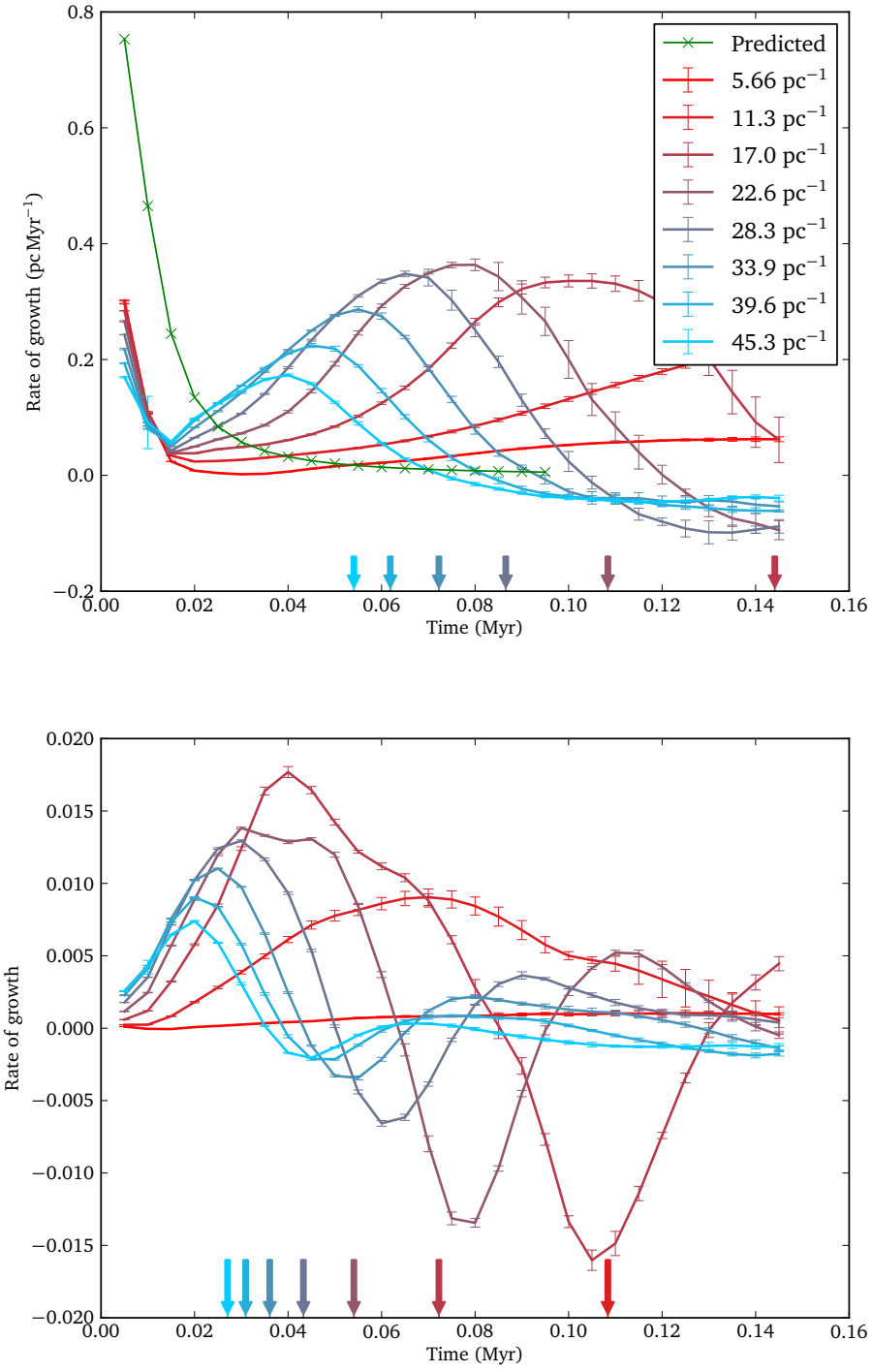


Figure 10.9 – The rates of growth for monochromatic perturbations with time for each triggered wavenumber from 2D simulations. The upper plot shows rates of growth of centre-of-mass position, while the lower plot shows the rates of growth of column density. The green line indicates the values predicted in Section 9.2.1. Each other line represents the rate of growth of the relevant wavenumber averaged over five realizations where only that wavenumber was triggered; error bars show the standard deviation across realizations. The arrows indicate the last resolvable time for the correspondingly coloured wavenumber, as defined in Section 9.2.3.

rates of growth decay slightly earlier than when they are predicted to be unresolvable. One-dimensional perturbations appear well-resolved until the theoretical limit, suggesting that two-dimensional perturbations are more difficult to resolve than one-dimensional perturbations.

As for the one-dimensional simulations, there is a small window of time around 0.02 Myr where the rates of growth of the perturbations are unaffected by either the initial velocity perturbation or by the decay of rates of growth due to wavenumbers becoming unresolvable. At later times we have to discard the rates of growth of the higher wavenumbers when performing further analysis.

For one-dimensional monochromatic perturbations, the initial rate of growth of the centre-of-mass position is larger than the sound speed, implying that the bending motion of the layer is supersonic. This is also true for these 2D simulations, but to a lesser degree due to the smaller rates of growth. The fastest rates of growth are approximately twice the sound speed.

The lower plot of Figure 10.9 shows the rates of growth of each wavenumber relating to the column density. As expected, the lowest wavenumber does not grow significantly, as it is not susceptible to the NTSI. The other wavenumbers produce a positive rate of growth of column density, which first rises from zero and later decays. The intermediate wavenumbers grow the fastest, as the highest wavenumbers saturate before they are able to grow significantly.

Column density perturbations are transient; once the instability creating them saturates and decays they will expand and disappear. This can be seen in Figure 10.9. Each peak in rate of growth is matched by an equivalent negative peak, representing the expansion of the perturbation. By comparison, a bending mode may stop growing, but the layer will remain bent in the absence of other disturbances. As for the one-dimensional case, we find that the rates of growth of column density are noisy, and do not produce clear trends. We therefore do not continue analysing them in this section.

Figure 10.10 shows the rates of growth of the centre-of-mass position as a function of wavenumber. The maximum resolvable wavenumber, as predicted in Section 9.2.3, is also shown. The rates of growth can be compared to the relationship

$$\tau^{-1} \propto k^{3/2} \,, \tag{10.3}$$

predicted by Vishniac (1994) for a one-dimensional perturbation. We find that at the earliest timestep, the distribution of growth rate is approximately equal across all wavenumbers. However, the growth rate of the lower wavenumbers rapidly decreases, increasing the gradient. The gradient appears to be close to the predicted relation.

For each timestep we fit power law slopes to the rates of growth as a function of

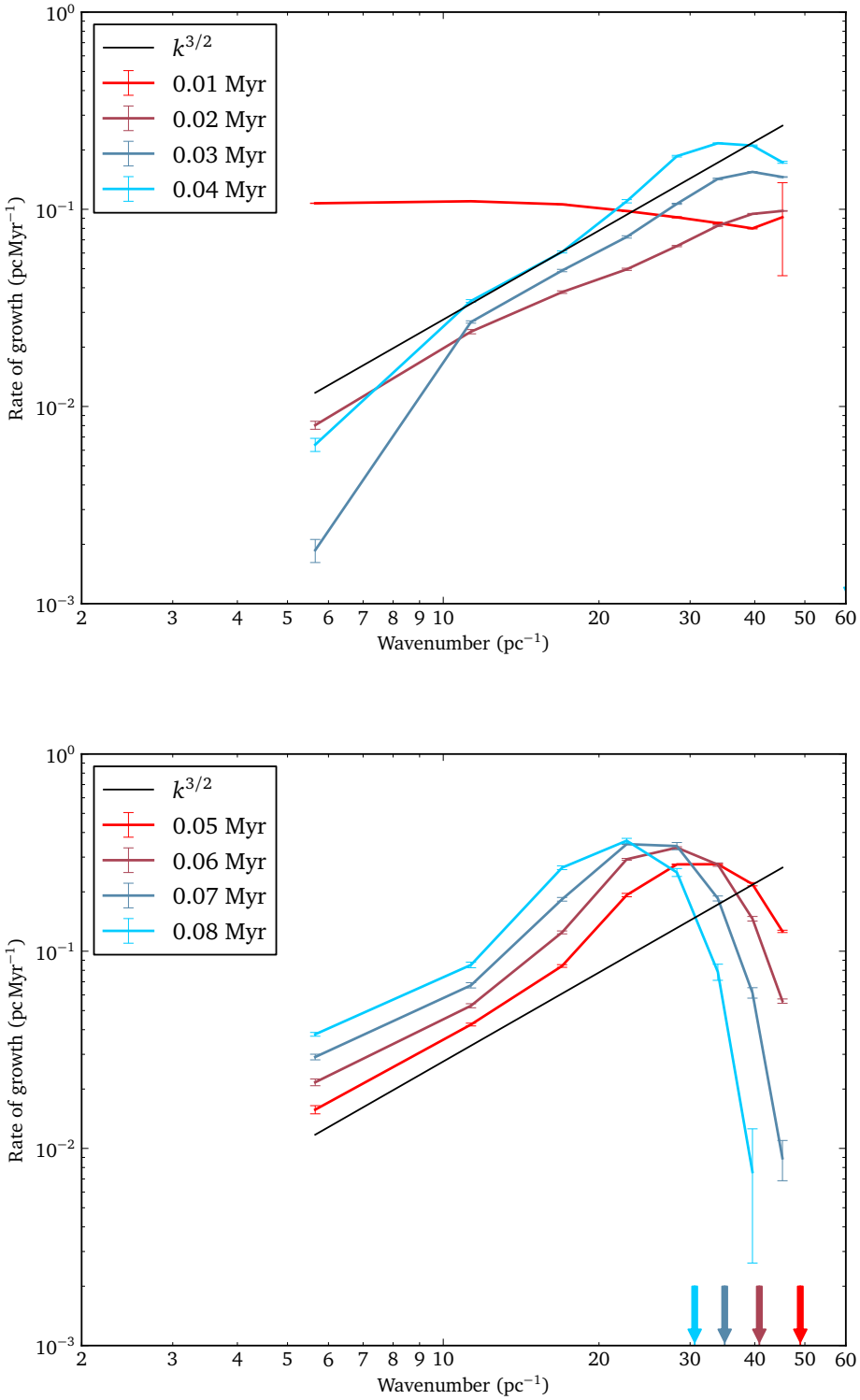


Figure 10.10 – The rates of growth of centre-of-mass position for monochromatic perturbations as a function of wavenumber from 2D simulations. Each line represents results from a different timestep; error bars show the standard deviation across realizations. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep, as defined in Section 9.2.3.

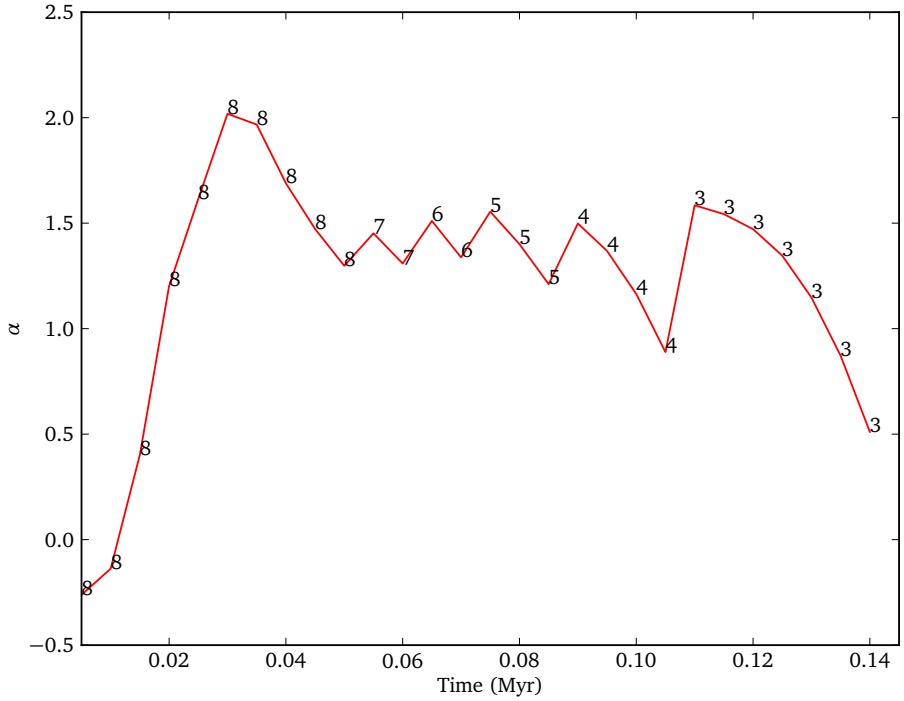


Figure 10.11 – Indices of power law fits to rates of growth of centre-of-mass position for monochromatic 2D simulations. The number of points used to construct each fit is also shown; fits constructed with fewer points are less reliable.

wavenumber, including only wavenumbers lower than the estimated maximum resolvable wavenumber. Any points with negative growth rates cannot be included in the power law fit and are disregarded. The index α of the power law fit at each timestep is plotted in Figure 10.11, together with the number of points used to construct each fit; later timesteps are heavily affected by the reduction in maximum resolvable wavenumber.

For the centre-of-mass position rates of growth, the power law index is initially low. This is the same as for the one-dimensional simulations, and is due to the initial velocity perturbation which is of equal amplitude for all wavenumbers. The power law index then rapidly rises. After initially reaching $\alpha \sim 2$, the index flattens out at approximately the predicted value of $\alpha = 1.5$. Despite our failure to accurately match the predicted power law index of rates of growth of centre-of-mass position for the one-dimensional simulations, we appear to have confirmed the relation for two-dimensional perturbations. One difference that may explain this is that the one-dimensional perturbations have a higher total absolute momentum, as a greater fraction of the layer is close to a 'peak'. The perturbation is therefore stronger, and may rapidly saturate the instability, preventing the NTSI from growing for a sufficient time to measure the power law index.

Chapter 11

NTSI – white noise perturbations

We run simulations of the NTSI which include white noise perturbations, as described in Section 9.1. We run high-resolution simulations with one-dimensional white-noise perturbations, and low-resolution simulations with two-dimensional perturbations. The white noise perturbations are set up as described in Section 8.2.3. We generate white noise perturbations between the wavenumbers $2\,\mathrm{pc}^{-1}$ to $40\,\mathrm{pc}^{-1}$ for one-dimensional perturbations. We use the same range for the wavenumbers k_y and k_z of the two-dimensional perturbations.

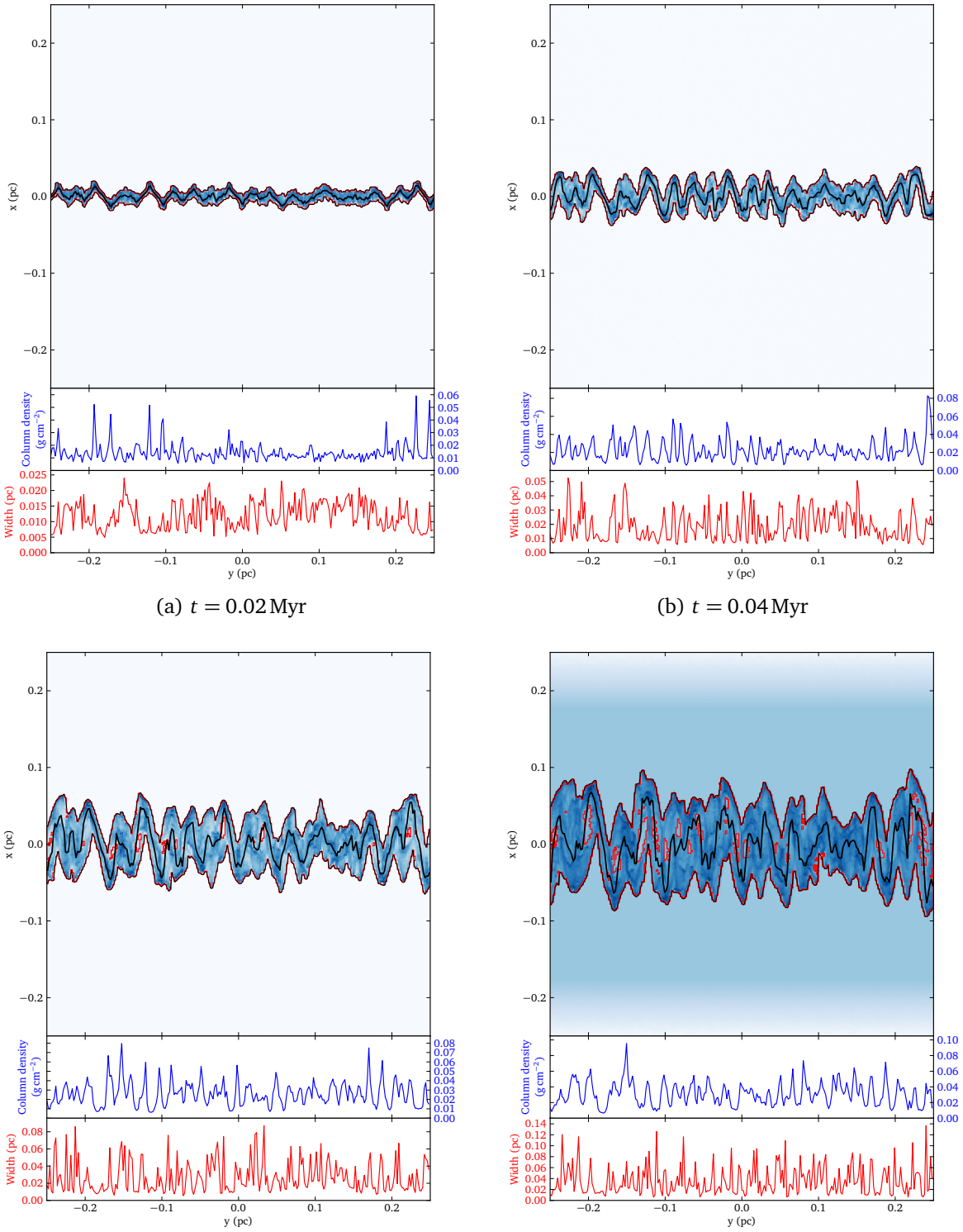
The amplitude of each mode is random and drawn from a uniform distribution between 0 and 0.2 times the collision velocity. This means that the amplitude of each mode is less than for the monochromatic simulations, where the amplitude of the excited mode is 0.5 times the collision velocity. The phases are randomly selected from a uniform distribution between 0 and 2π. We create ten realizations of one-dimensional perturbations and ten realizations of two-dimensional perturbations. All of the quantities that we analyse are averaged across the ten relevant realizations in each case. We also calculate the standard deviation for each quantity.

Figure 11.1 shows an example of a simulation with white noise perturbations. This shows that the centre-of-mass position is a good tracer of the bending modes of the layer, and that the width is a good measure of size of the layer. However, it is not obvious to see that the column density traces the bending modes of the layer.

11.1 One-dimensional white noise perturbations

We first examine the one-dimensional results from the 1DHR simulation suite. Figure 11.2 shows the rates of growth of centre-of-mass position for our simulations with white noise. At the first timestep the rates of growth are similar at all wavenumbers lower than the highest excited wavenumber of $40\,\mathrm{pc}^{-1}$. Beyond this wavenumber the rates of growth are considerably reduced.

At early timesteps, the slope of the rates of growth steepens as the rates of growth for small wavenumbers decrease and the rates of growth for high wavenumbers increase. The



(a) $t = 0.02\,\mathrm{Myr}$

(b) $t = 0.04\,\mathrm{Myr}$

(c) $t = 0.06\,\mathrm{Myr}$

(d) $t = 0.08\,\mathrm{Myr}$

Figure 11.1 – Layer diagrams for 1DHR simulation of a white noise perturbation. Colours indicate density but are not consistent across plots. The thick black line shows the centre-of-mass position of the layer. Red lines indicate the edges of groups of layer cells. The thin black lines show the upper and lower extent of the layer. The approach of the rarefaction wave is visible in the final diagram. Column density and layer width as a function of y position are also shown.

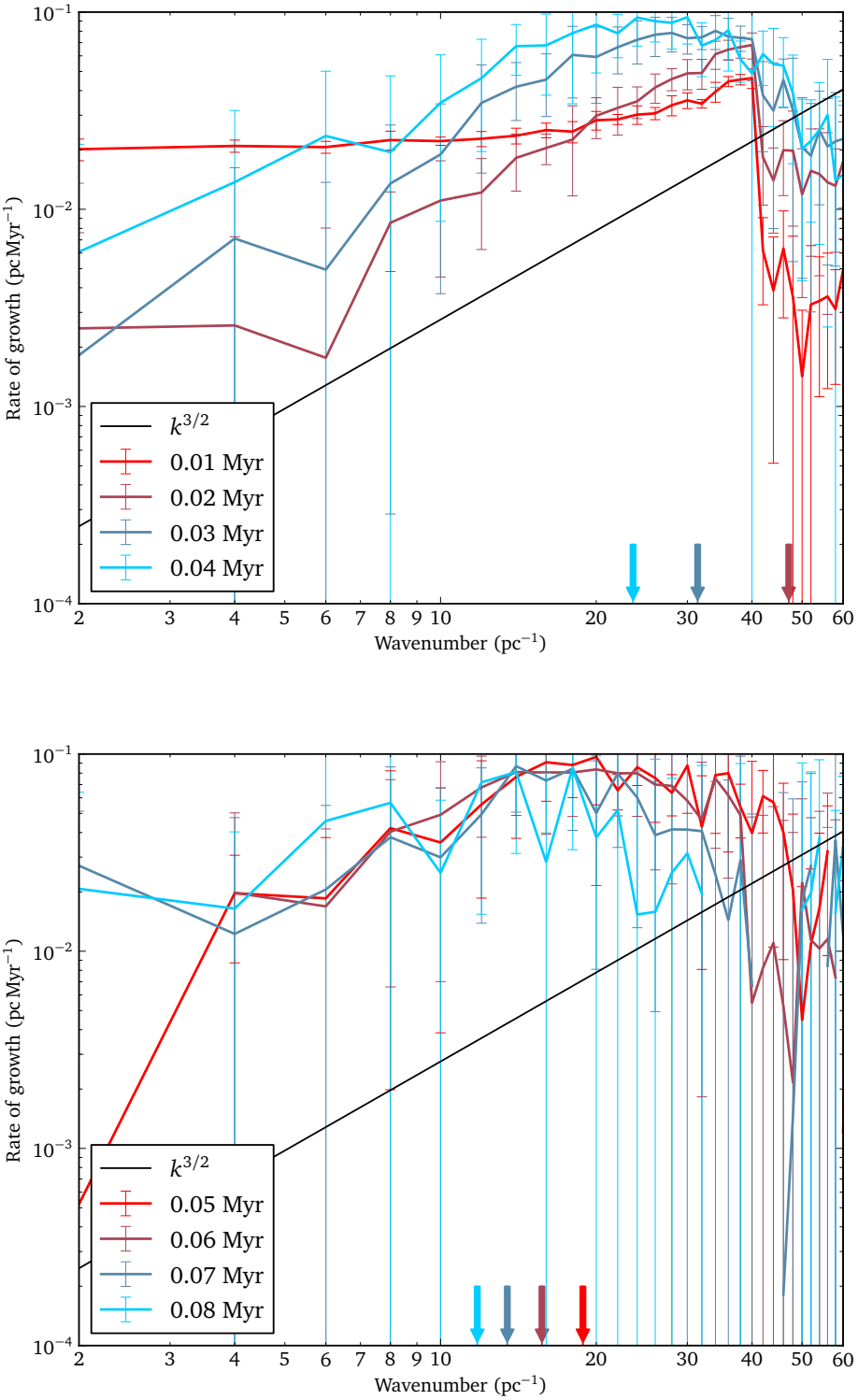


Figure 11.2 – The rates of growth of centre-of-mass position for white noise perturbations as a function of wavenumber from 1DHR simulations. Each line represents results from a different timestep; error bars show the standard deviation across realizations. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep, as defined in Section 9.2.3.

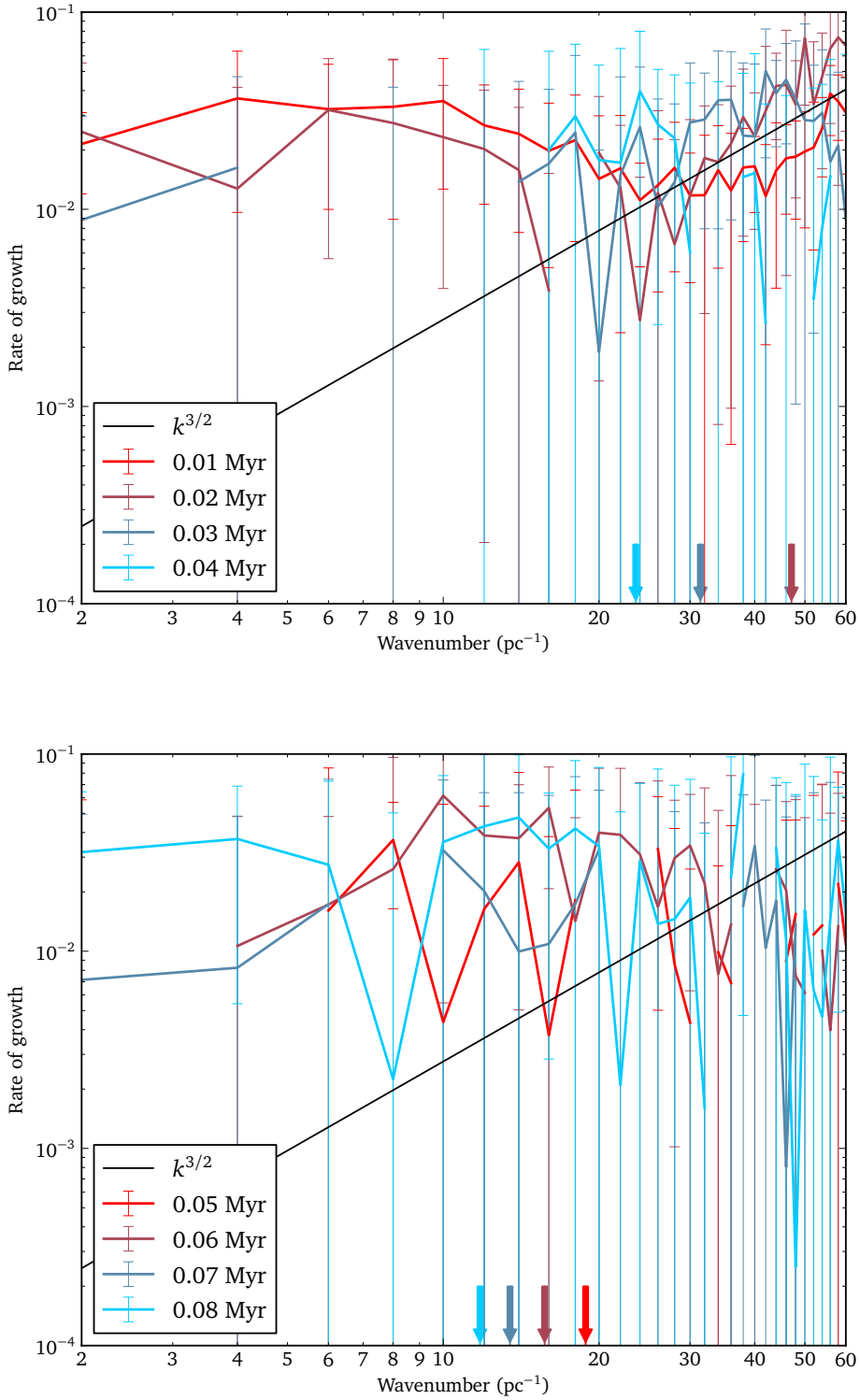


Figure 11.3 – The rates of growth of column density for white noise perturbations as a function of wavenumber from 1DHR simulations. Each line represents results from a different timestep; error bars show the standard deviation across realizations. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep, as defined in Section 9.2.3.

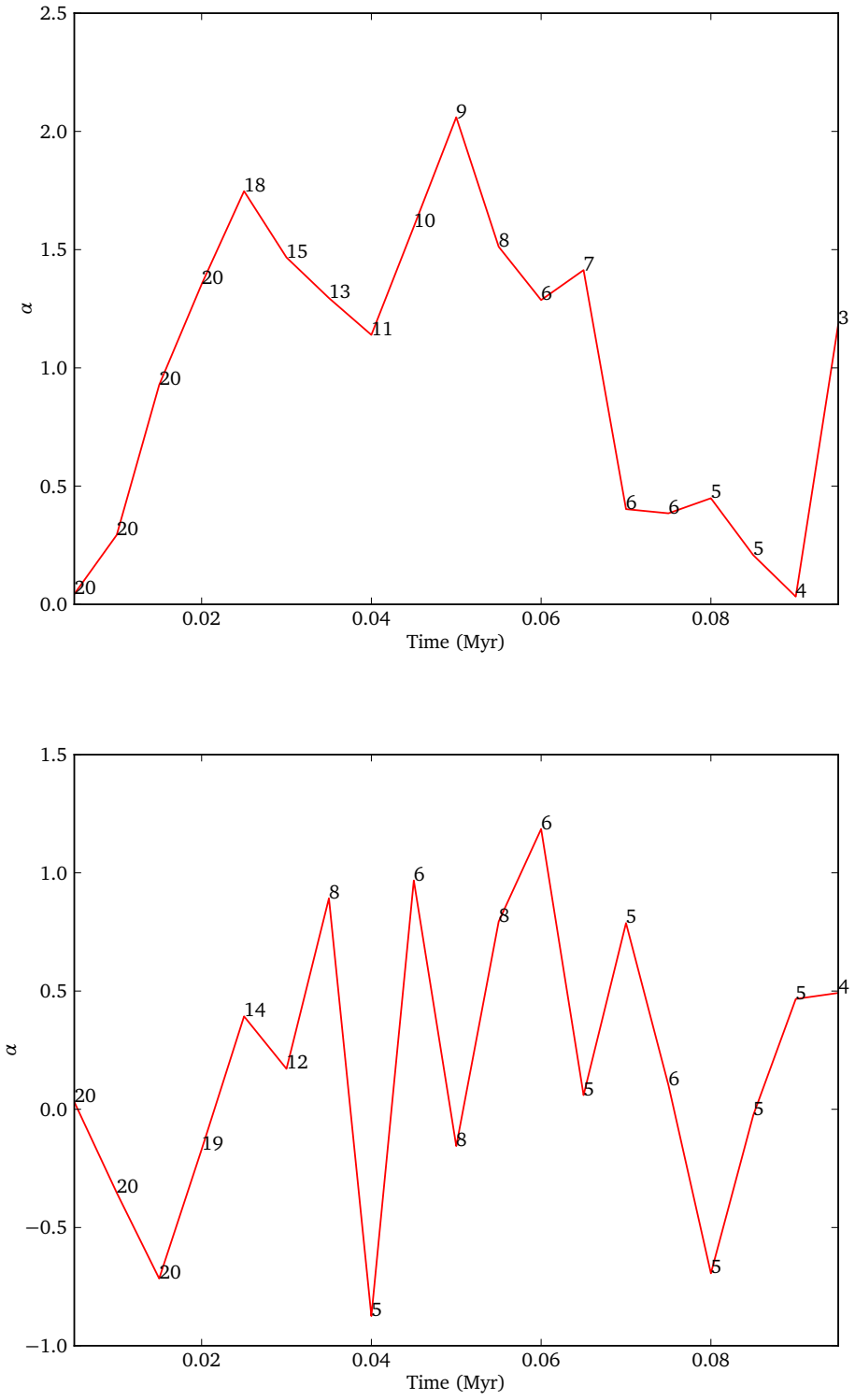


Figure 11.4 – Indices of power law fits to rates of growth for white noise 1DHR simulations. The upper plot shows centre-of-mass position rates of growth; the lower plot shows column density rates of growth. The number of points used to construct each fit is also shown; fits constructed with fewer points are less reliable.

transition between increase and decrease occurs at approximately $10\,\mathrm{pc}^{-1}$. In Section 10.1, we found this was the minimum wavenumber susceptible to the NTSI. As the simulation progresses, the wavenumbers higher than $40\,\mathrm{pc}^{-1}$ also become excited. At intermediate timesteps, the rates of growth as a function of wavenumber does appear to fit the relation predicted by Vishniac (1994). At later timesteps the higher wavenumbers become unresolvable at the predicted wavenumbers.

Figure 11.3 shows the rates of growth for column density. We find that the rate of growth is approximately constant across all wavenumbers, although the results are very noisy. There does not appear to be any clear trend with respect to either wavenumber or time. We are therefore unable to draw conclusions from these results.

We fit power laws to the rates of growth of centre-of-mass position and column density. We only include wavenumbers lower than both the maximum resolvable wavenumber at each timestep and the highest excited wavenumber of $40\,\mathrm{pc}^{-1}$. Figure 11.4 shows the indices of these power laws with respect to time. For the centre-of-mass position, the index is initially close to zero. This is expected, as the initial perturbation is white noise with a uniform distribution of amplitudes as a function of wavenumber. The index then rises to approximately the predicted value of 1.5. At later times the index begins to fall, but at these later times the higher wavenumbers have become unresolvable and thus the power law fits are less reliable.

For the column density rates of growth, the power law indices are noisy and erratic, varying unpredictably between zero and one. The column density does not appear to provide useful results in this context.

11.2 Two-dimensional white noise perturbations

We now examine the two-dimensional results from the 2D simulation suite. Figure 11.5 shows the rates of growth of centre-of-mass position as a function of wavenumber. As for the one-dimensional case, at the earliest timestep the rates of growth are approximately equal at all excited wavenumbers. At higher wavenumbers, the later evolution is also similar to the one-dimensional case; the rates of growth for the higher wavenumbers increases to match the predicted slope.

At lower wavenumbers the behaviour is quite different. Below $10\,\mathrm{pc}^{-1}$, the lowest wavenumber susceptible to the NTSI, the rates of growth immediately become negative. This decrease is more dramatic than in the one-dimensional case, falling well below the predicted relation of Vishniac (1994). At later timesteps, the lowest wavenumbers then begin to grow, reaching rates of growth comparable to those at wavenumbers where the NTSI is present.

As these lower wavenumbers are below the minimum wavenumber susceptible to the NTSI, it is not clear what is causing these modes to grow. If we consider only the

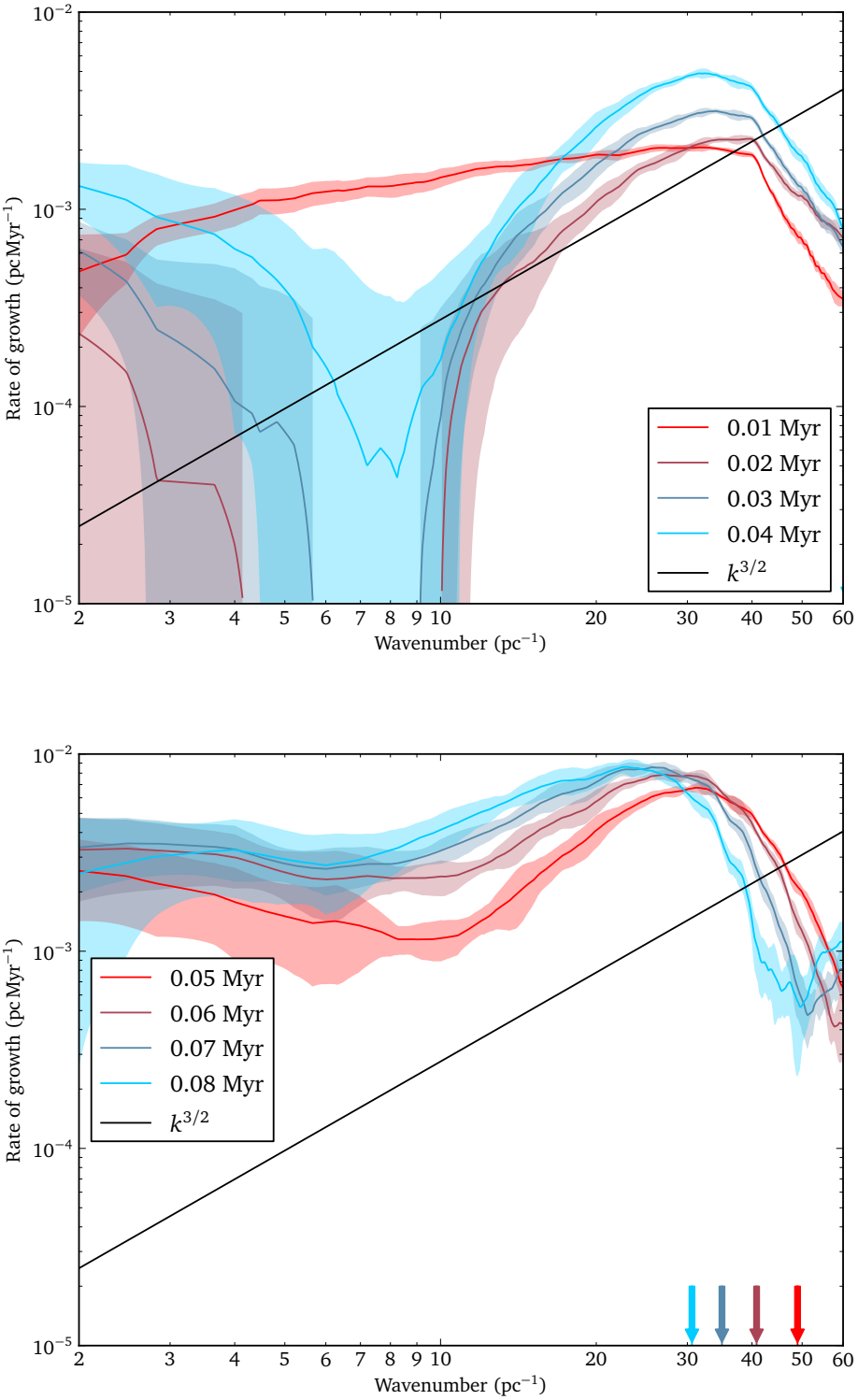


Figure 11.5 – The rates of growth of centre-of-mass position for white noise perturbations as a function of wavenumber from 2D simulations. Each line represents results from a different timestep; shaded regions show the standard deviation across realizations. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep, as defined in Section 9.2.3.

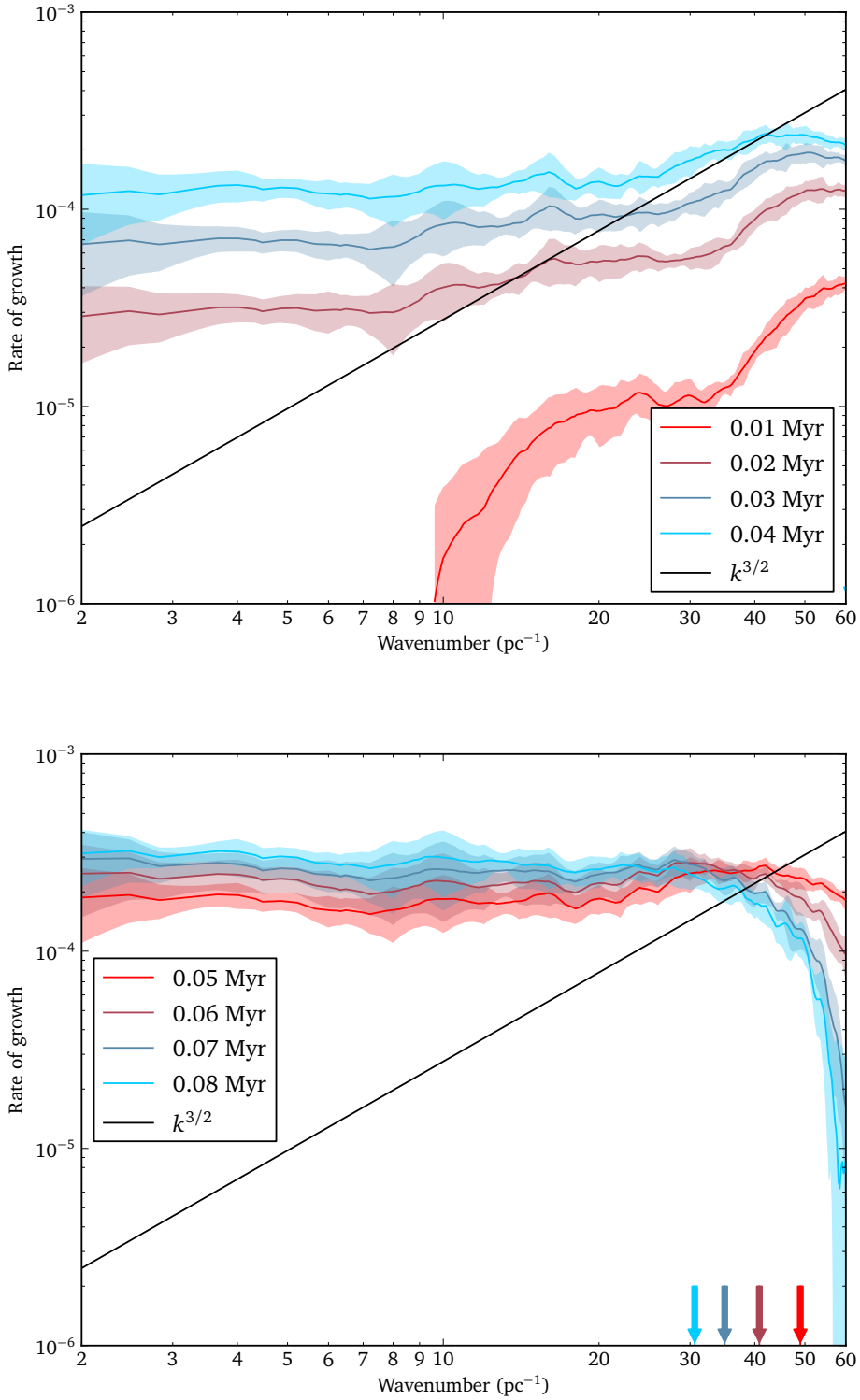


Figure 11.6 – The rates of growth of column density for white noise perturbations as a function of wavenumber from 2D simulations. Each line represents results from a different timestep; shaded regions show the standard deviation across realizations. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep, as defined in Section 9.2.3.

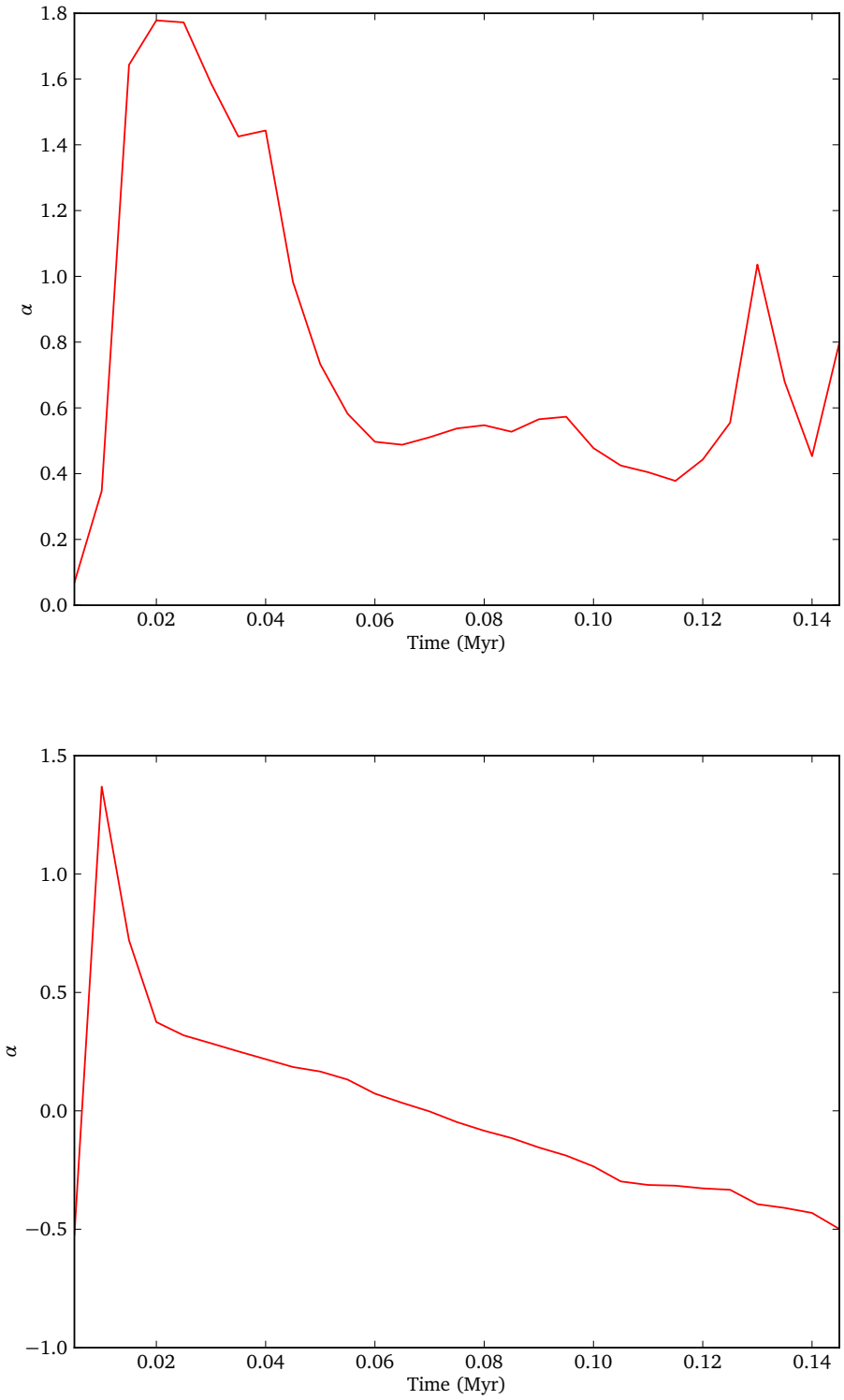


Figure 11.7 – Indices of power law fits to rates of growth for white noise 2D simulations. The upper plot shows centre-of-mass position rates of growth; the lower plot shows column density rates of growth.

wavenumbers above the lowest susceptible wavenumber and below the highest resolvable wavenumber, the rates of growth do appear to approximately match the theoretical prediction. However when we fit power laws to the rates of growth, the excess rates of growth at small wavelengths will tend to reduce the power law index.

We also consider the rates of growth of column density, as shown in Figure 11.6. These are less noisy than for the one-dimensional case. Initially the highest wavenumbers grow fastest. However, at later times the rate of growth is approximately a constant for all wavenumbers lower than the maximum resolvable wavenumber, and may have saturated. As the prediction of Vishniac (1994) relates to the bending of the layer, there is no reason to assume that the surface density perturbations will follow a similar relation.

Figure 11.7 shows the indices of power law fits to the rates of growth from centre-of-mass position and column density. For the fits to centre-of-mass position, the index begins close to zero as expected. This then rapidly rises, overshooting the predicted value of 1.5, before falling to a fairly-constant index of approximately 0.6. Examination of Figure 11.5 can help explain this. The higher wavenumbers do appear to approximately match the predicted relation. However, at lower wavenumbers the rates of growth decrease more than expected, raising the power law index to its peak of approximately 1.8. These low wavenumbers then grow more than expected, and so the power law index decreases.

For fits to column density, the power law index first rises to approximately 1.4 before slowly decreasing towards, and beyond, zero. The reason for the slow decay in power law index is not clear. It also fails to match the results of the one-dimensional simulation, although the power law indices for the one-dimensional simulations are very noisy.

Chapter 12

NTSI – turbulent perturbations

We run a series of simulations with turbulent velocity fields, as described in Section 9.1. We use two types of turbulent velocity field, differing only by their average velocity. One set uses a subsonic average velocity of Mach 0.5, while the other uses a supersonic average velocity of Mach 4, as described in Section 9.1.1. We run simulations as part of the 1DHR suite. We consider these simulations to contain one-dimensional perturbations in the layer, as the velocity field only varies significantly along the collision axis, which does not directly affect the layer, and along one dimension across the layer. The other dimension across the layer is ignored due to the narrowness of the simulation box along the z axis. We also run simulations as part of the 2D suite, which we consider two-dimensional perturbations in the layer by the same reasoning.

The simulation box for the 1DHR simulations is much narrower along the z axis than along the y axis, although both axes are periodic. For perturbations with a wavenumber larger than the z axis of the box, the variation across the z axis can be ignored. For perturbations with a wavenumber smaller than the z axis this approximation is no longer true, and our method of taking average densities across the z axis is no longer representative of the true distribution of perturbations.

The z axis of the 1DHR simulations is 0.01 pc, as described in Section 9.1.1. This corresponds to a wavenumber of $100\,\mathrm{pc}^{-1}$. We do not generally consider such high wavenumbers as they are not resolvable at most times in our simulations. Our approximation of a narrow periodic box as producing one-dimensional perturbations in the layer is therefore valid for our purposes.

Our turbulent velocity field has an amplitude spectrum with a power law index of -2, as described in Section 8.2.4. This means that the strongest perturbations are found at the lowest wavenumbers. We may therefore observe growth at these wavenumbers which is due to the initial turbulent perturbation. These perturbations may also have an effect for longer than the initial velocity perturbations in other simulations.

When adding monochromatic or white noise perturbations, the perturbations are confined to the region close to the collision. These regions enter the layer within a

relatively short time, and should rapidly decay as predicted in Section 9.2.1. However the turbulent perturbations are distributed throughout the simulation volume, and will not decay significantly before entering the layer. The layer will therefore experience a continuous input of perturbations, with the lowest wavenumbers being the strongest.

12.1 One-dimensional subsonic turbulent perturbations

Figure 12.1 shows an example of a 1DHR simulation with subsonic turbulence. This shows that the centre-of-mass position is a good tracer of the bending modes of the layer, and that the width is a good measure of size of the layer. However, it is not obvious to see that the column density traces the bending modes of the layer. We can see that subsonic turbulence does not create strong density enhancements in the inflowing gas.

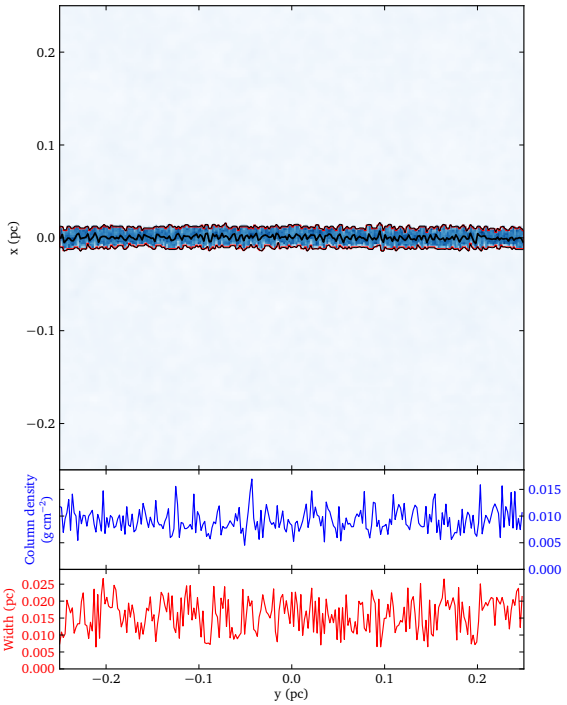
Figure 12.2 shows the rates of growth of centre-of-mass position for the one-dimensional simulations of subsonic turbulence. At earlier timesteps, the rates of growth appear slightly higher at both low and high wavenumbers compared to intermediate wavenumbers. The growth at low wavenumbers may be due to the continuous input of turbulent perturbations. There is a hint of rates of growth approaching the predicted power law relation of Vishniac (1994) at wavenumbers between the minimum wavenumber susceptible to the NTSI of $10\,\mathrm{pc}^{-1}$ and the maximum resolvable wavenumber at each timestep. At later timesteps the rates of growth are approximately equal at all wavenumbers, although they are quite noisy.

Figure 12.3 shows the rates of growth of column density for the one-dimensional simulations of subsonic turbulence. The rates of growth are very noisy, particularly at low wavenumbers. There is no clear trend with increasing wavenumber. Figure 12.4 shows the indices of power law fits to rates of growth. The fits to both centre-of-mass position and column density are quite noisy. The index α of the power law fit is approximately zero in both cases.

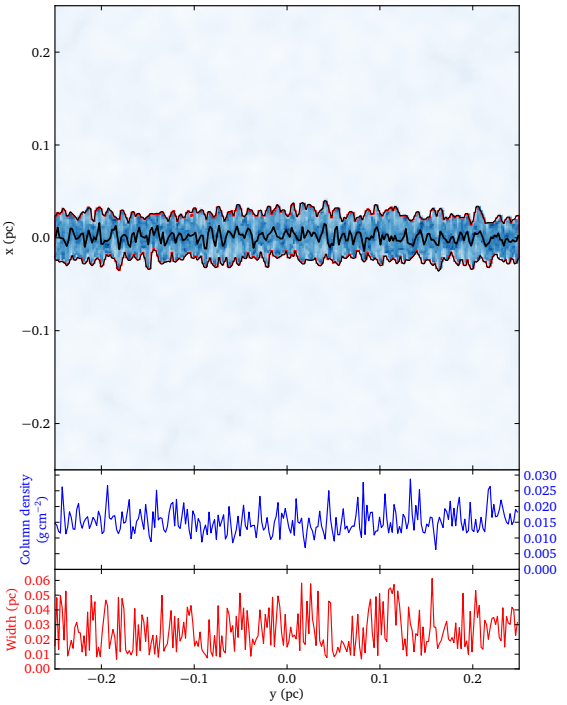
12.2 Two-dimensional subsonic turbulent perturbations

We now consider the results from 2D simulations with subsonic turbulence. Figure 12.5 shows the rates of growth of centre-of-mass position. There are three different regimes of wavenumber visible. At low wavenumbers, the rates of growth decreases with increasing wavenumber, with the highest rates of growth at low wavenumbers. This is probably due to the continuous input of turbulent perturbations.

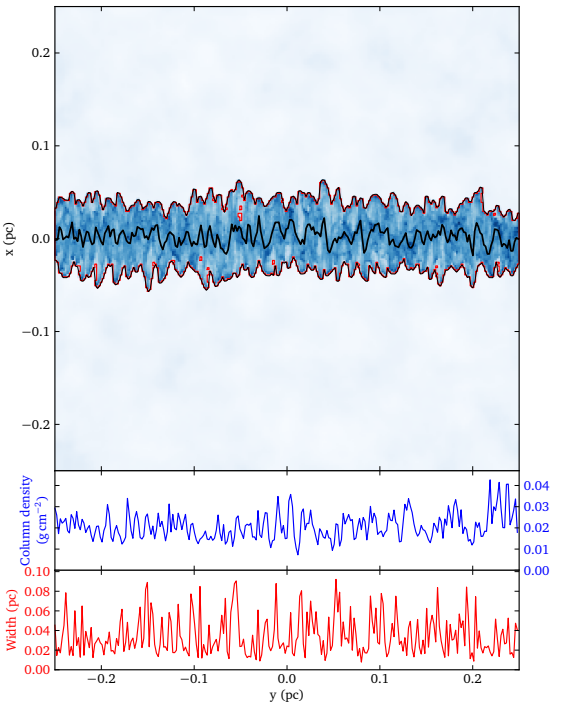
At intermediate wavenumbers the rates of growth increase with increasing wavenumber. Although the rate of increase does not quite reach the relation predicted by Vishniac (1994), the increase is probably due to the NTSI triggering the growth of higher wavenumbers. The transition wavenumber between the effects of turbulence and the growth of the NTSI is approximately the minimum wavenumber susceptible to the NTSI,



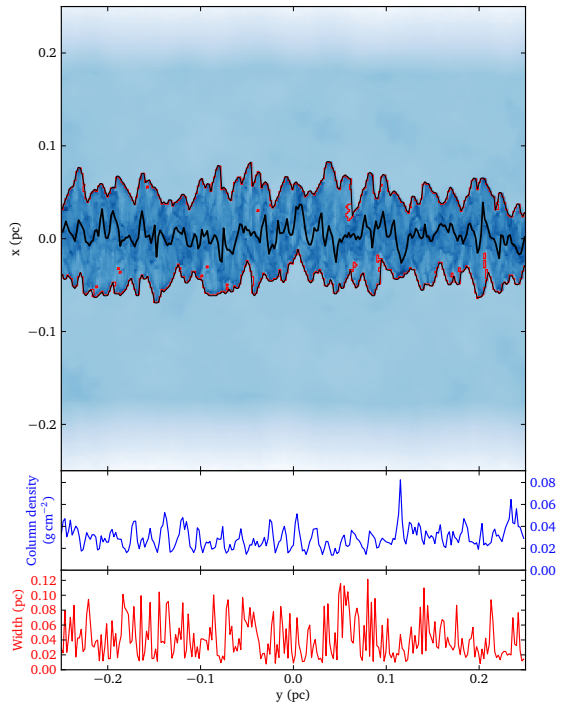
(a) $t = 0.02\,\mathrm{Myr}$



(b) $t = 0.04\,\mathrm{Myr}$



(c) $t = 0.06\,\mathrm{Myr}$



(d) $t = 0.08\,\mathrm{Myr}$

Figure 12.1 – Layer diagrams for 1DHR simulation with subsonic turbulence. Colours indicate density but are not consistent across plots. The thick black line shows the centre-of-mass position of the layer. Red lines indicate the edges of groups of layer cells. The thin black lines show the upper and lower extent of the layer. The approach of the rarefaction wave is visible in the final diagram. Column density and layer width as a function of y position are also shown.

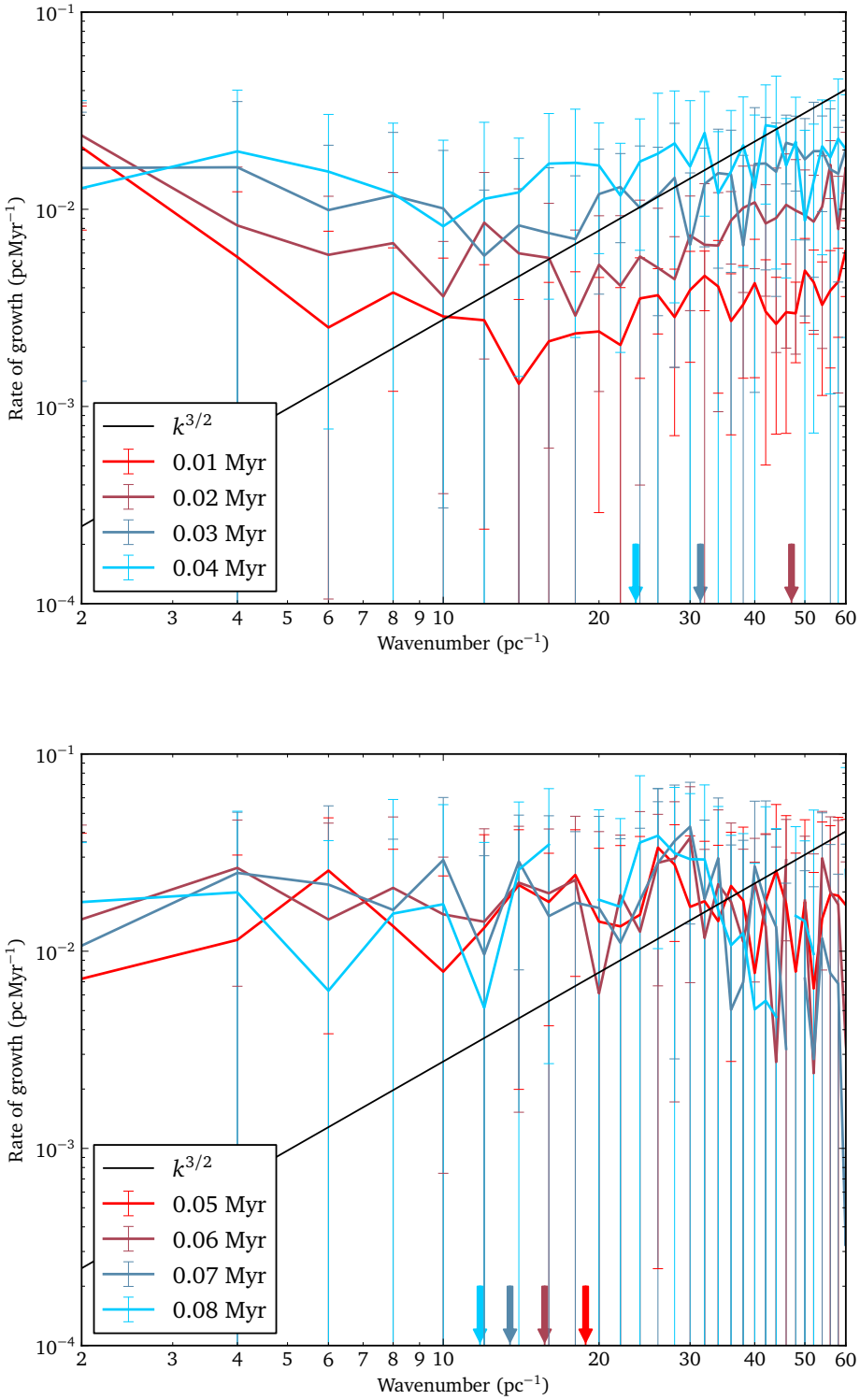


Figure 12.2 – The rates of growth of centre-of-mass position for subsonic turbulent perturbations as a function of wavenumber from 1DHR simulations. Each line represents results from a different timestep; error bars show the standard deviation across realizations. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep, as defined in Section 9.2.3.

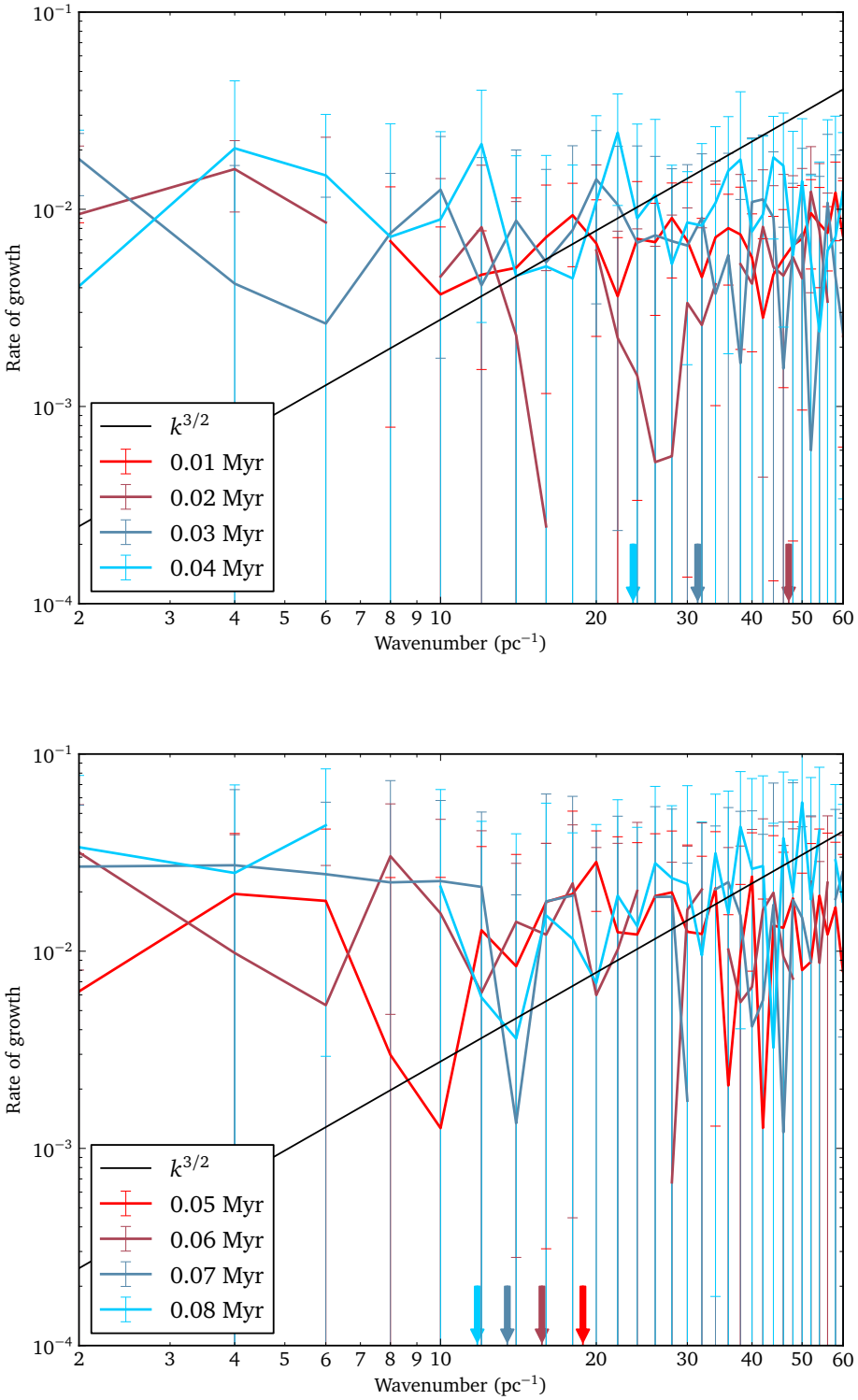


Figure 12.3 – The rates of growth of column density for subsonic turbulent perturbations as a function of wavenumber from 1DHR simulations. Each line represents results from a different timestep; error bars show the standard deviation across realizations. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep, as defined in Section 9.2.3.

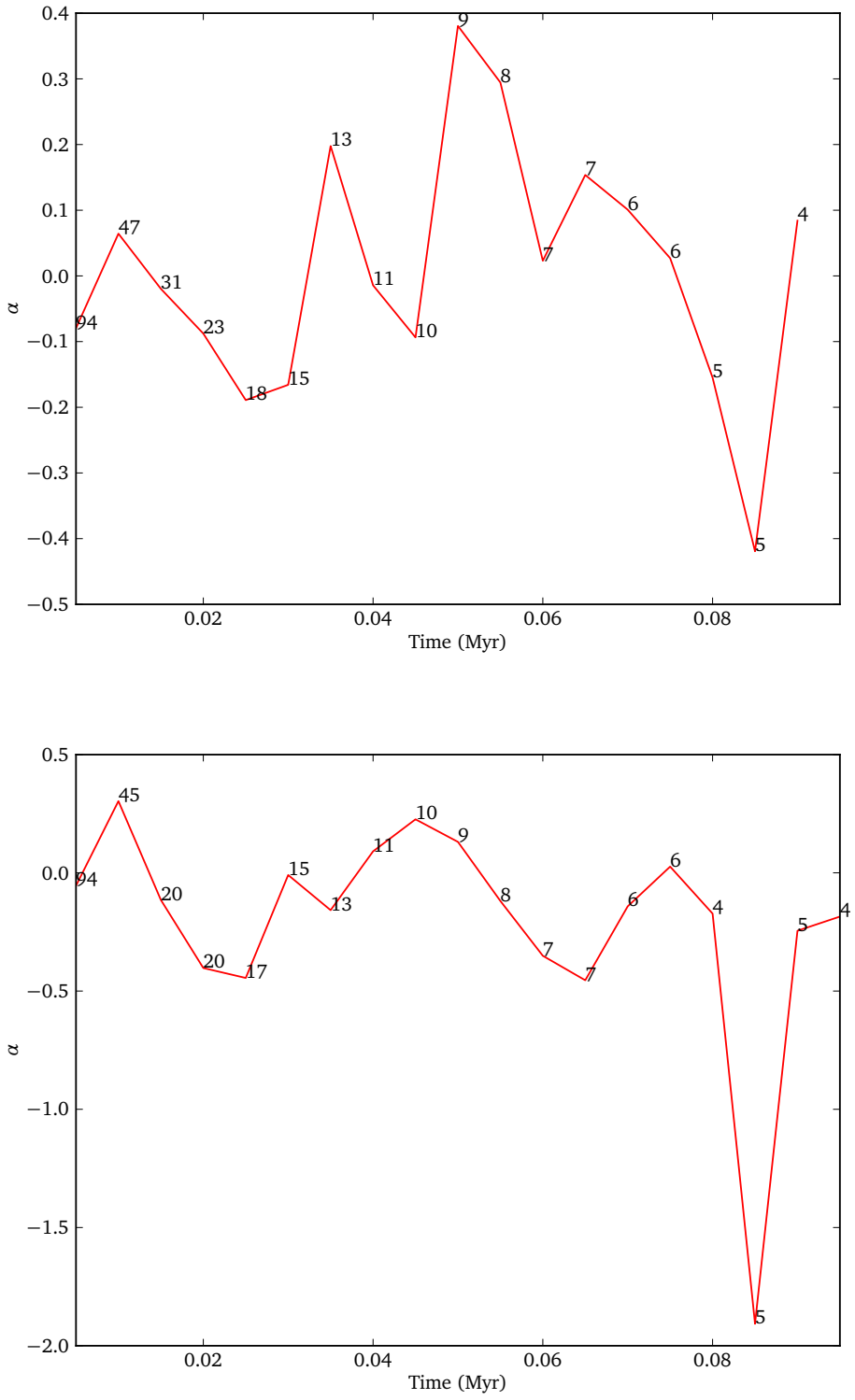


Figure 12.4 – Indices of power law fits to rates of growth for subsonic turbulent 1DHR simulations. The upper plot shows centre-of-mass position rates of growth; the lower plot shows column density rates of growth. The number of points used to construct each fit is also shown; fits constructed with fewer points are less reliable.

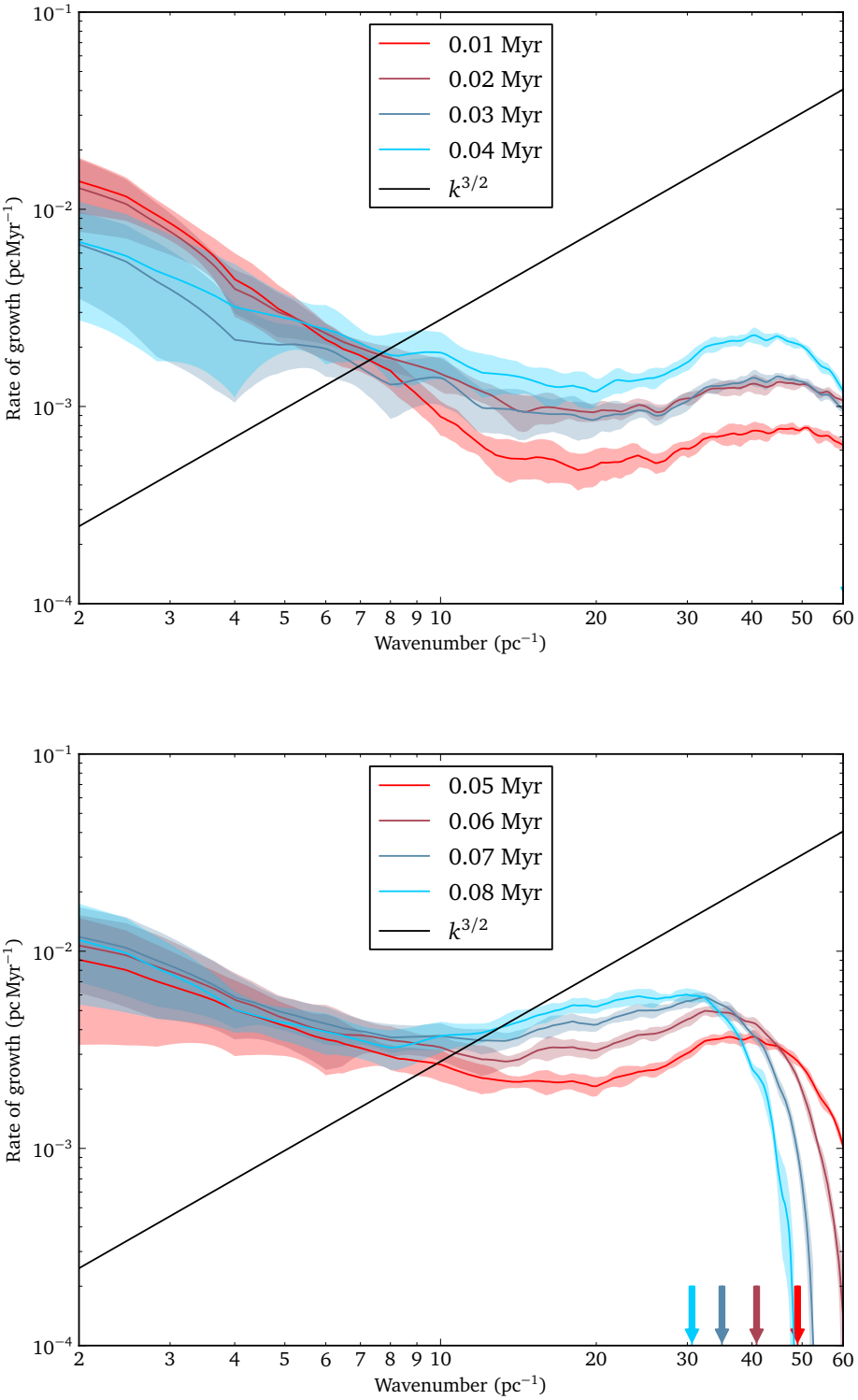


Figure 12.5 – The rates of growth of centre-of-mass position for subsonic turbulent perturbations as a function of wavenumber from 2D simulations. Each line represents results from a different timestep; shaded regions show the standard deviation across realizations. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep, as defined in Section 9.2.3.

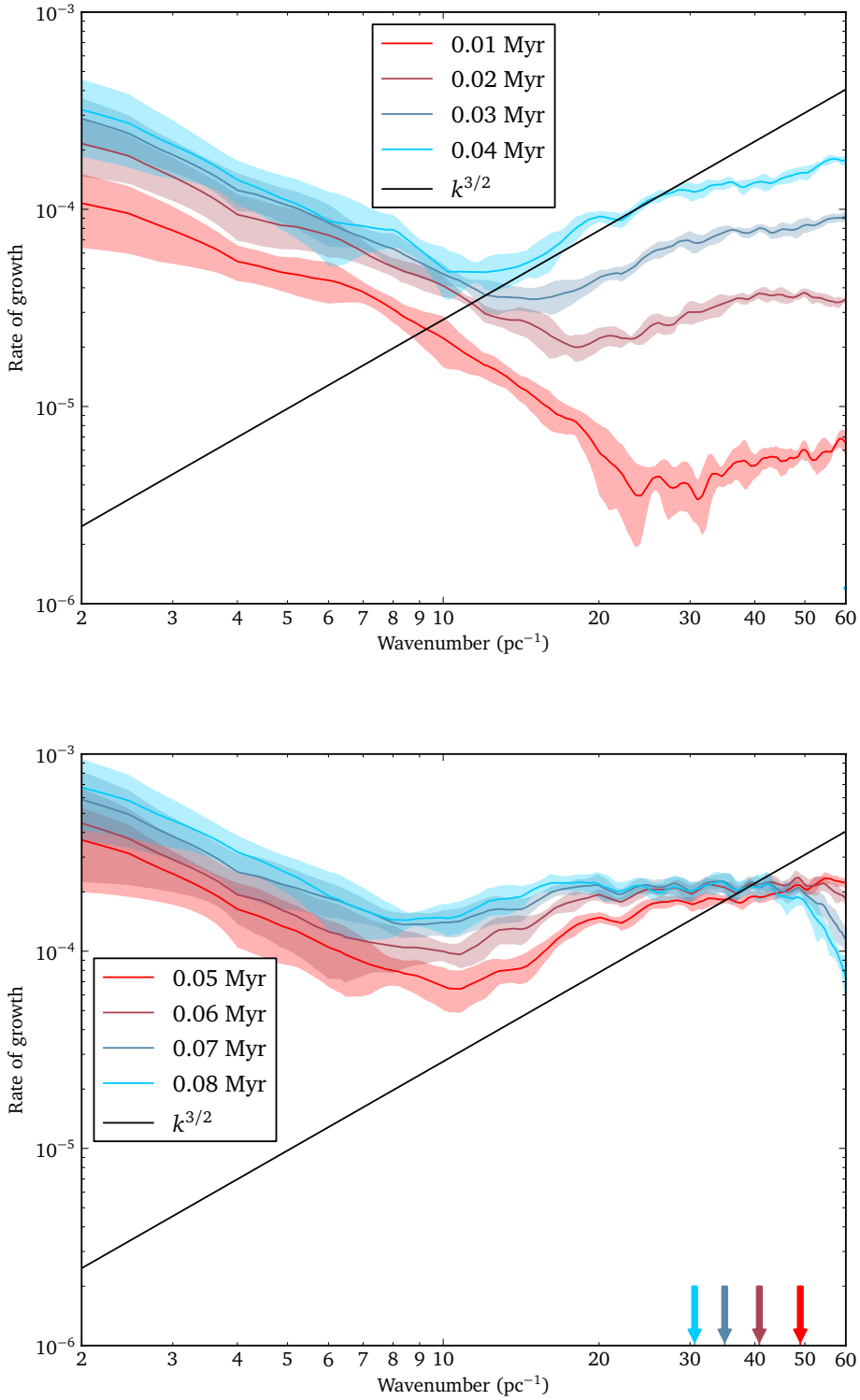


Figure 12.6 – The rates of growth of column density for subsonic turbulent perturbations as a function of wavenumber from 2D simulations. Each line represents results from a different timestep; shaded regions show the standard deviation across realizations. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep, as defined in Section 9.2.3.

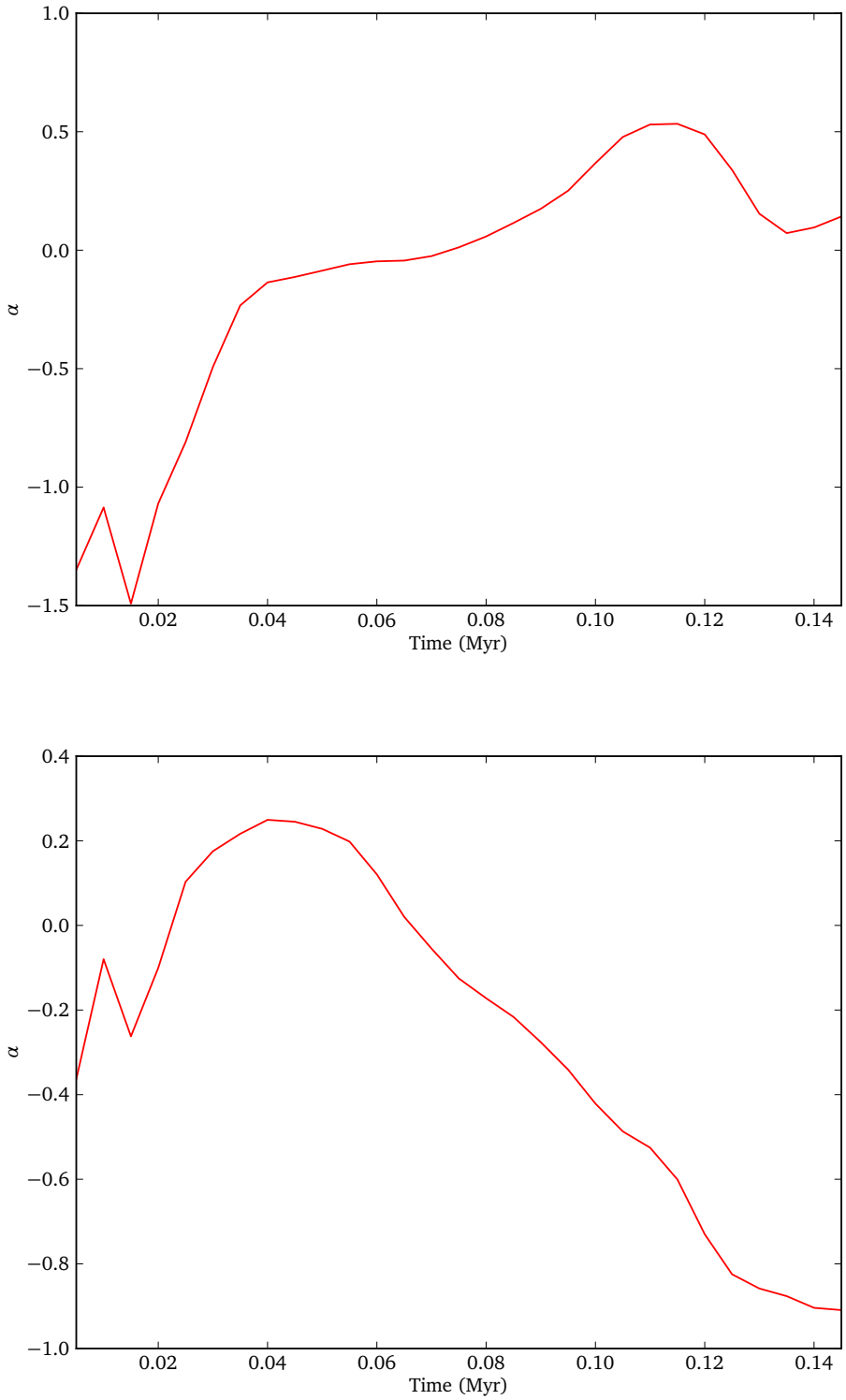


Figure 12.7 – Indices of power law fits to rates of growth for subsonic turbulent 2D simulations. The upper plot shows centre-of-mass position rates of growth; the lower plot shows column density rates of growth.

which is $10\,\mathrm{pc}^{-1}$. At the highest wavenumbers the rates of growth decreases rapidly. This is due to the wavenumber exceeding the maximum resolvable wavenumber of bending mode perturbations, and occurs at approximately the predicted wavenumber.

Figure 12.6 shows the rates of growth of column density. The rates of growth of column density are fairly similar to those for the centre-of-mass position. We see the same three regimes corresponding to turbulence, the growth of the NTSI and unresolvable wavenumbers. The major difference is that the transition between turbulence and the NTSI begins at much smaller wavenumbers, and evolves towards the minimum wavenumber susceptible to the NTSI. This may be because it takes longer to build column density perturbations through the NTSI than simply to bend the layer.

Figure 12.7 shows the indices of power law fits to the rates of growth of centre-of-mass position and column density. These fit the rates of growth for all wavenumbers lower than the maximum resolvable wavenumbers, and so do not reflect the complexity of the rates of growth. However, they may be a useful first approximation. For the fits to centre-of-mass position, the power law index α initially rises to approximately zero, and remains fairly constant apart from briefly peaking at 0.5.

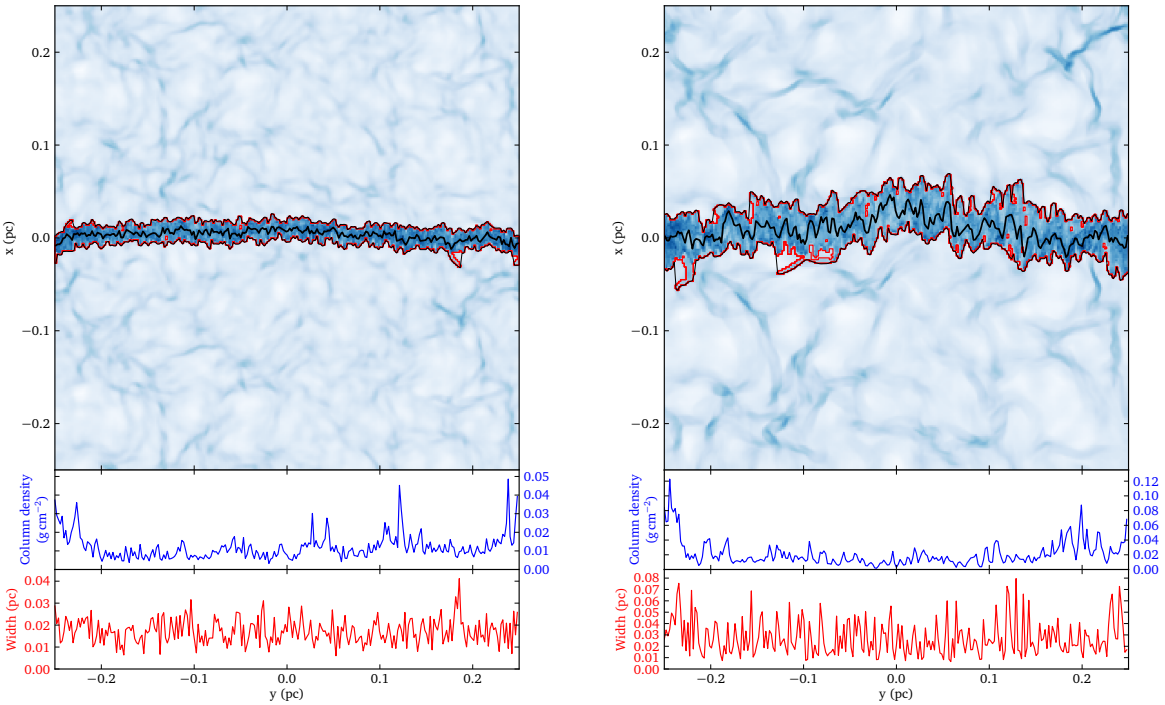
The fits to column density find the that the power law index quickly peaks at approximately 0.2, before falling slowly to -1.0. This indicates that the growth in column density at later times is concentrated in low wavenumber perturbations, due to the continuous input of turbulence.

12.3 One-dimensional supersonic turbulent perturbations

Figure 12.8 shows an example of a 1DHR simulation with supersonic turbulence. This shows that the centre-of-mass position is a good tracer of the bending modes of the layer, and that the width is a good measure of size of the layer. However, it is not obvious to see that the column density traces the bending modes of the layer. The supersonic turbulence creates large density enhancements in the inflowing gas. However, the filtering process described in Section 5.3.5 works well to restrict the analysis to the layer, rather than including dense regions in the inflowing gas.

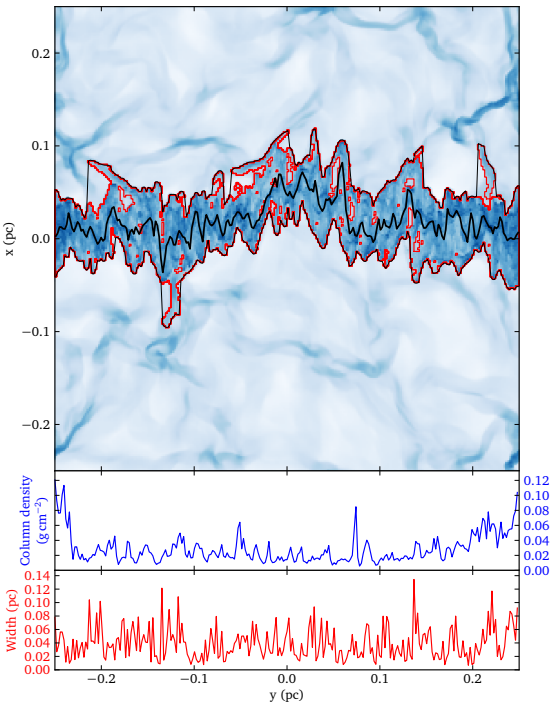
For supersonic turbulence we can expect the continuous input of turbulence to be more significant than for the subsonic turbulence. Figure 12.9 shows the rates of growth of centre-of-mass position for 1DHR simulations with supersonic turbulence. The rates of growth decrease with increasing wavenumber, with no evidence of the NTSI. The rates of growth are also larger than for the subsonic turbulence. The one-dimensional results, as for the subsonic turbulence, are also quite noisy.

Figure 12.10 shows the rates of growth of column density. Again, we see decreasing rates of growth with increasing wavenumber. This suggests that the strong turbulence is suppressing the growth the NTSI. Figure 12.11 shows the indices of power law fits to the

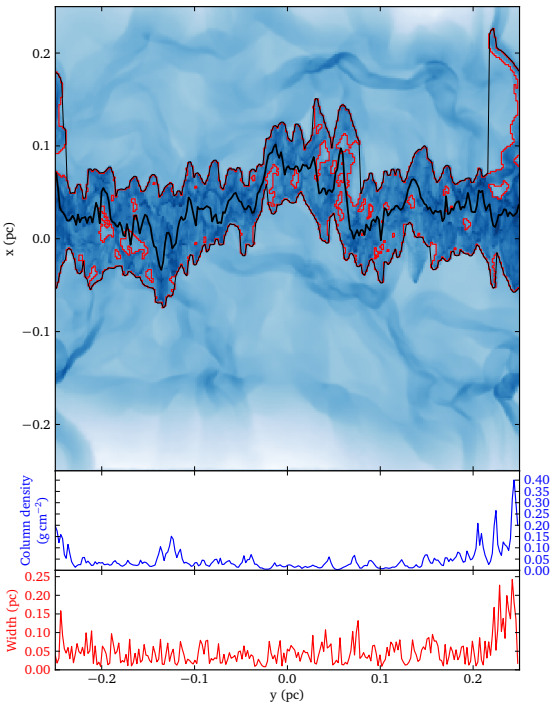


(a) $t = 0.02\,\mathrm{Myr}$

(b) $t = 0.04\,\mathrm{Myr}$



(c) $t = 0.06\,\mathrm{Myr}$



(d) $t = 0.08\,\mathrm{Myr}$

Figure 12.8 – Layer diagrams for 1DHR simulation with supersonic turbulence. Colours indicate density but are not consistent across plots. The thick black line shows the centre-of-mass position of the layer. Red lines indicate the edges of groups of layer cells. The thin black lines show the upper and lower extent of the layer. The approach of the rarefaction wave is visible in the final diagram.

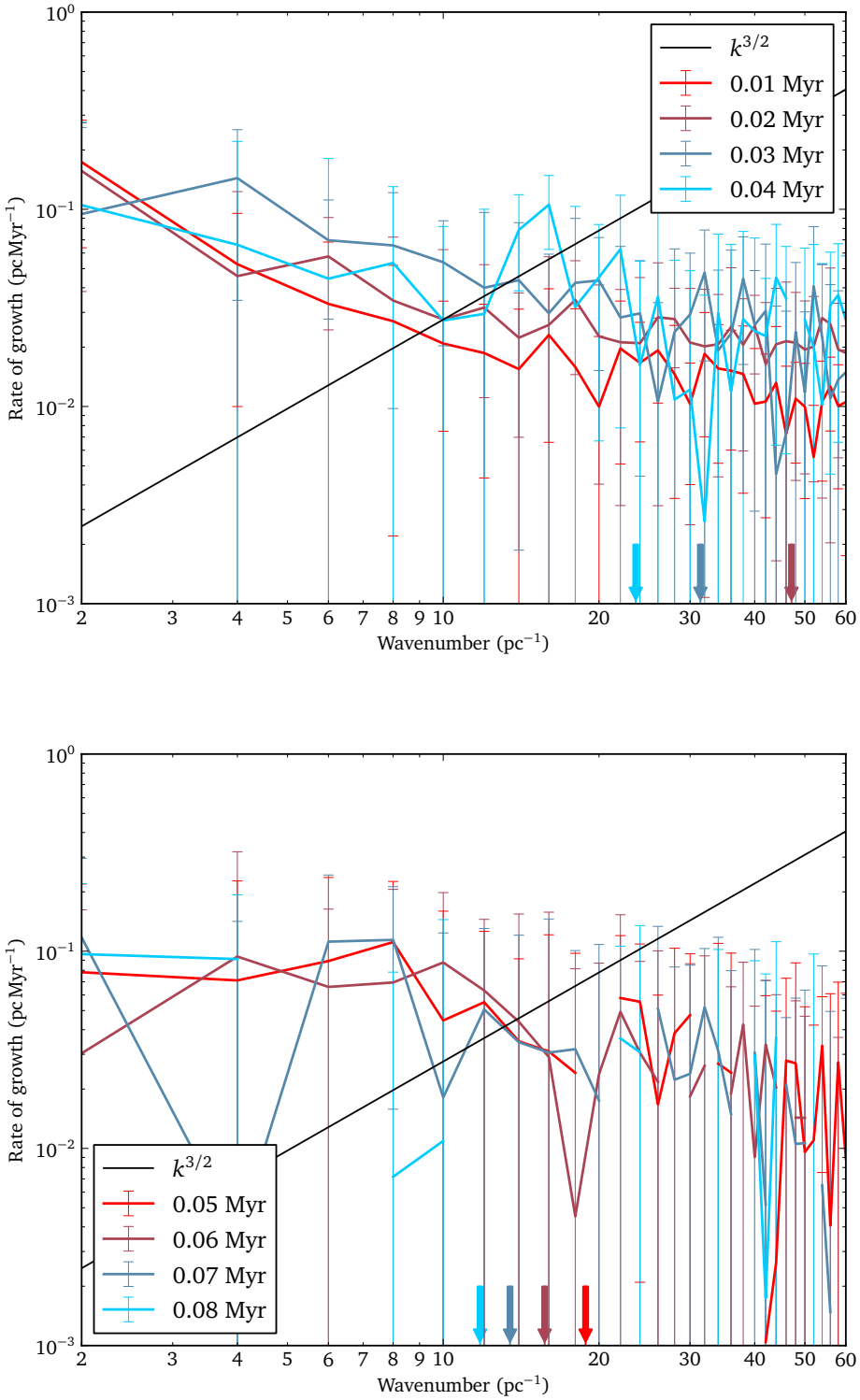


Figure 12.9 – The rates of growth of centre-of-mass position for supersonic turbulent perturbations as a function of wavenumber from 1DHR simulations. Each line represents results from a different timestep; error bars show the standard deviation across realizations. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep, as defined in Section 9.2.3.

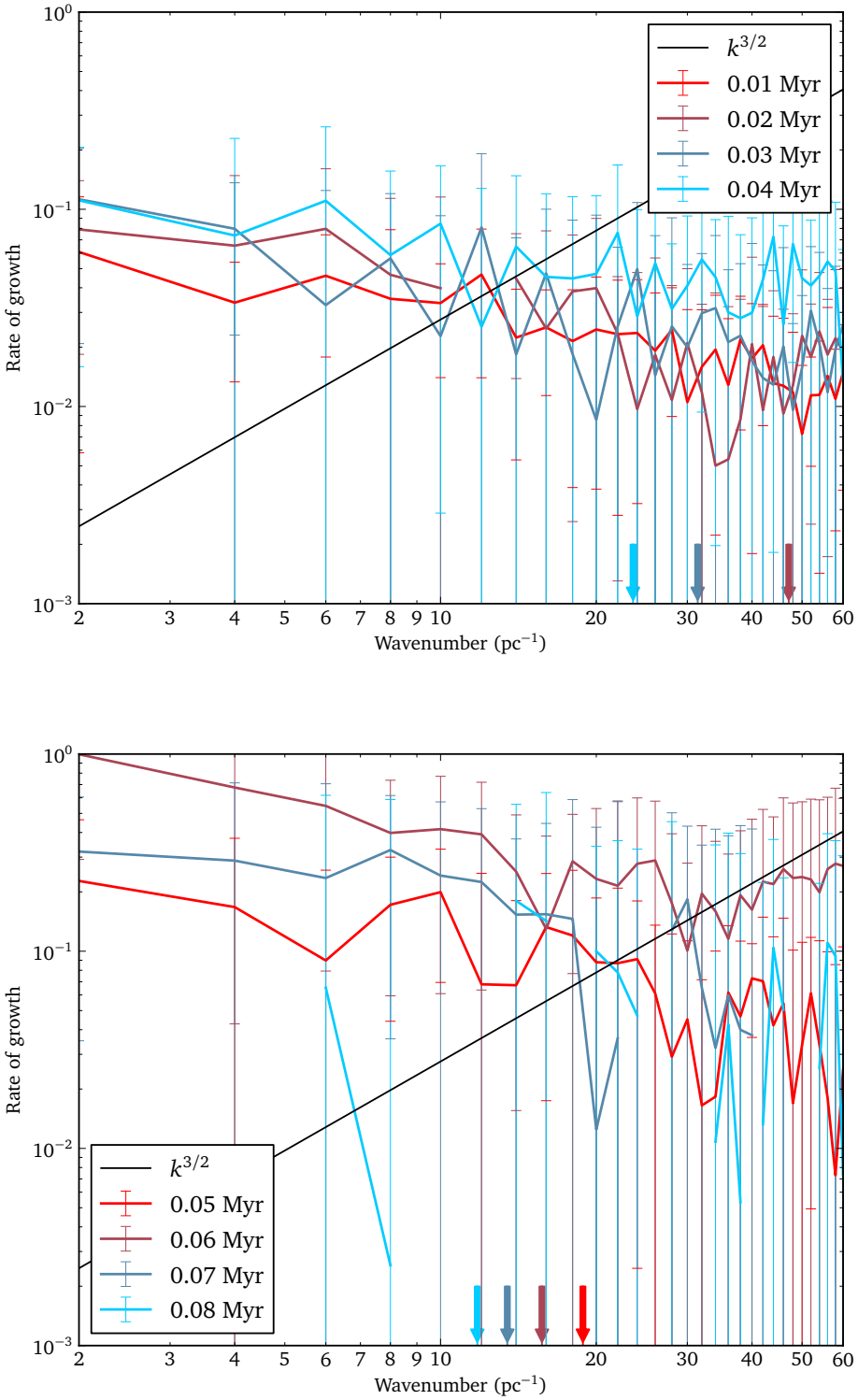


Figure 12.10 – The rates of growth of column density for supersonic turbulent perturbations as a function of wavenumber from 1DHR simulations. Each line represents results from a different timestep; error bars show the standard deviation across realizations. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep, as defined in Section 9.2.3.

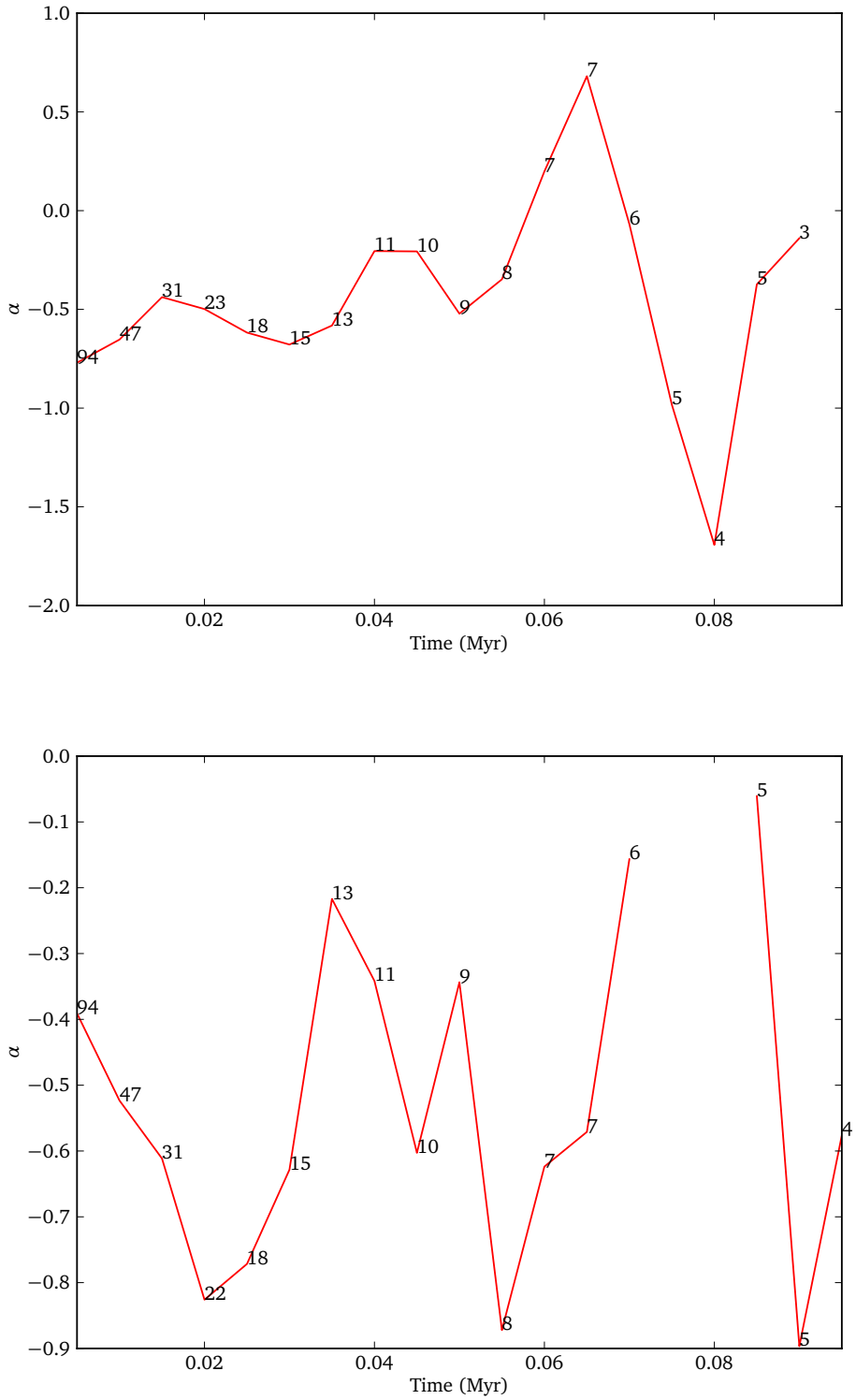


Figure 12.11 – Indices of power law fits to rates of growth for supersonic turbulent 1DHR simulations. The upper plot shows centre-of-mass position rates of growth; the lower plot shows column density rates of growth. The number of points used to construct each fit is also shown; fits constructed with fewer points are less reliable.

rates of growth of centre-of-mass position and the column density. In both cases there is some variation with time, but the indices tend to be approximately -0.5. A negative index such as this is not expected in simulations where the NTSI is dominant.

12.4 Two-dimensional supersonic turbulent perturbations

We examine the simulations with supersonic turbulence from the 2D simulation suite. Figure 12.12 shows the rates of growth of centre-of-mass position. As for the one-dimensional case, the rates of growth decrease rapidly with increasing wavenumber. There is no evidence of the NTSI acting at any wavenumber.

The situation is slightly different for the rates of growth of column density, as shown in Figure 12.13. The overall trend is similar, with the rate of growth decreasing with increasing wavenumber. However, at early timesteps there is some evidence of the NTSI increasing rates of growth at higher wavenumbers. At later times these increases disappear.

Figure 12.14 shows the indices of power law fits to the rates of growth. For fits to centre-of-mass position, the power law index α begins at approximately -1.5 before rising and peaking at approximately -0.5. This is less steep than the original turbulent power law index of -2, showing there is not a direct relation between the spectrum of perturbations in the turbulence and those created in the layer.

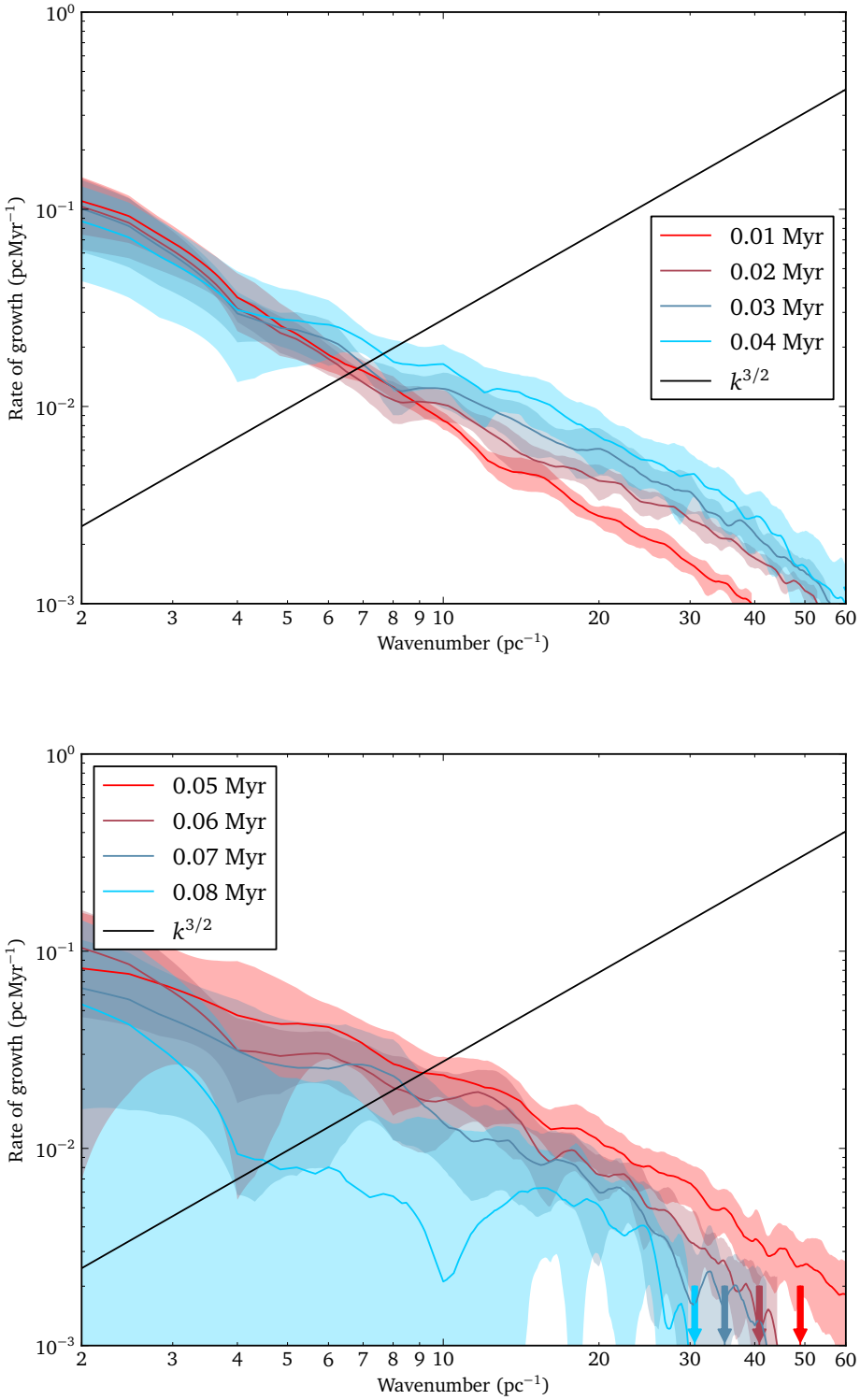


Figure 12.12 – The rates of growth of centre-of-mass position for supersonic turbulent perturbations as a function of wavenumber from 2D simulations. Each line represents results from a different timestep; shaded regions show the standard deviation across realizations. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep, as defined in Section 9.2.3.

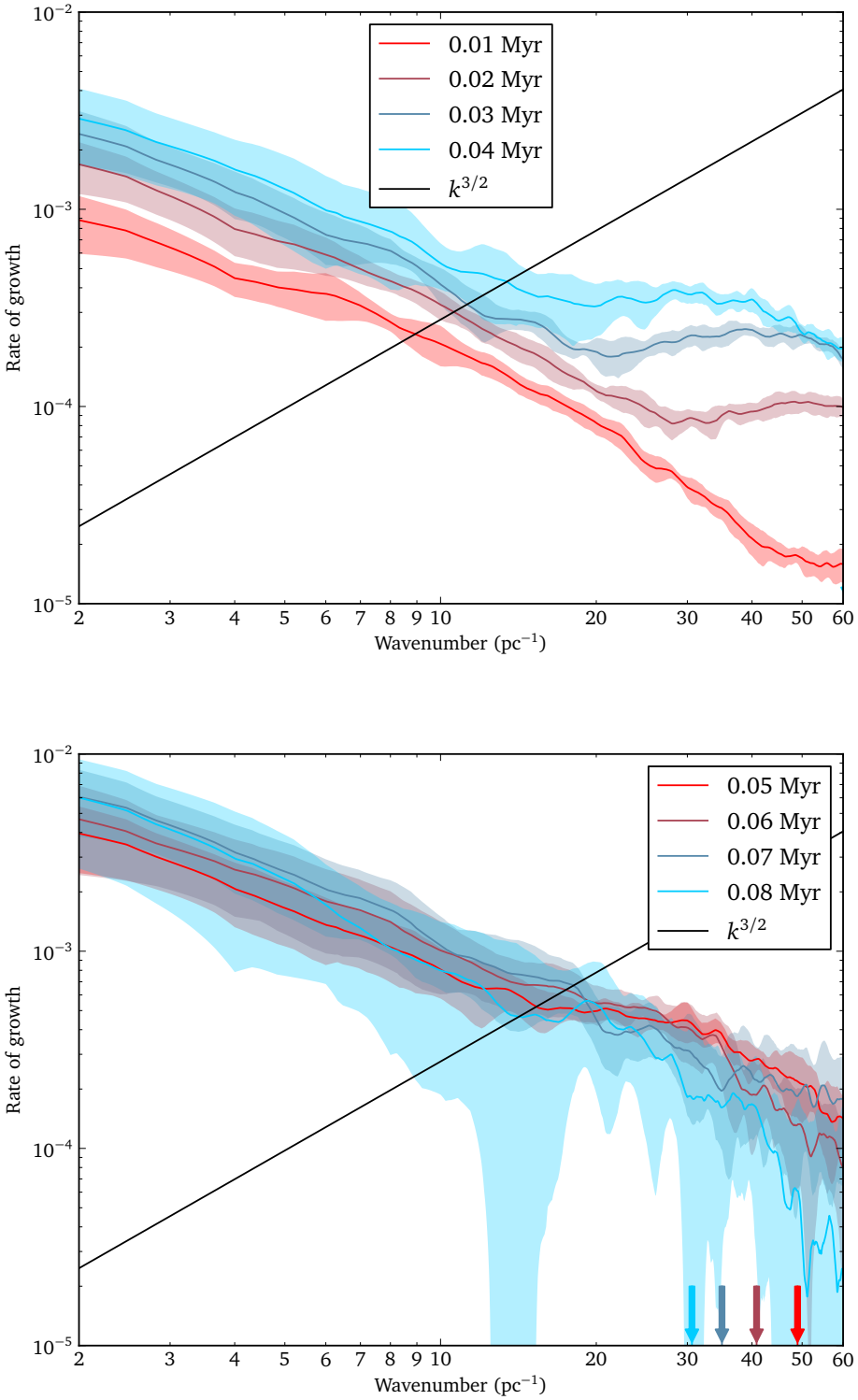


Figure 12.13 – The rates of growth of column density for supersonic turbulent perturbations as a function of wavenumber from 2D simulations. Each line represents results from a different timestep; shaded regions show the standard deviation across realizations. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep, as defined in Section 9.2.3.

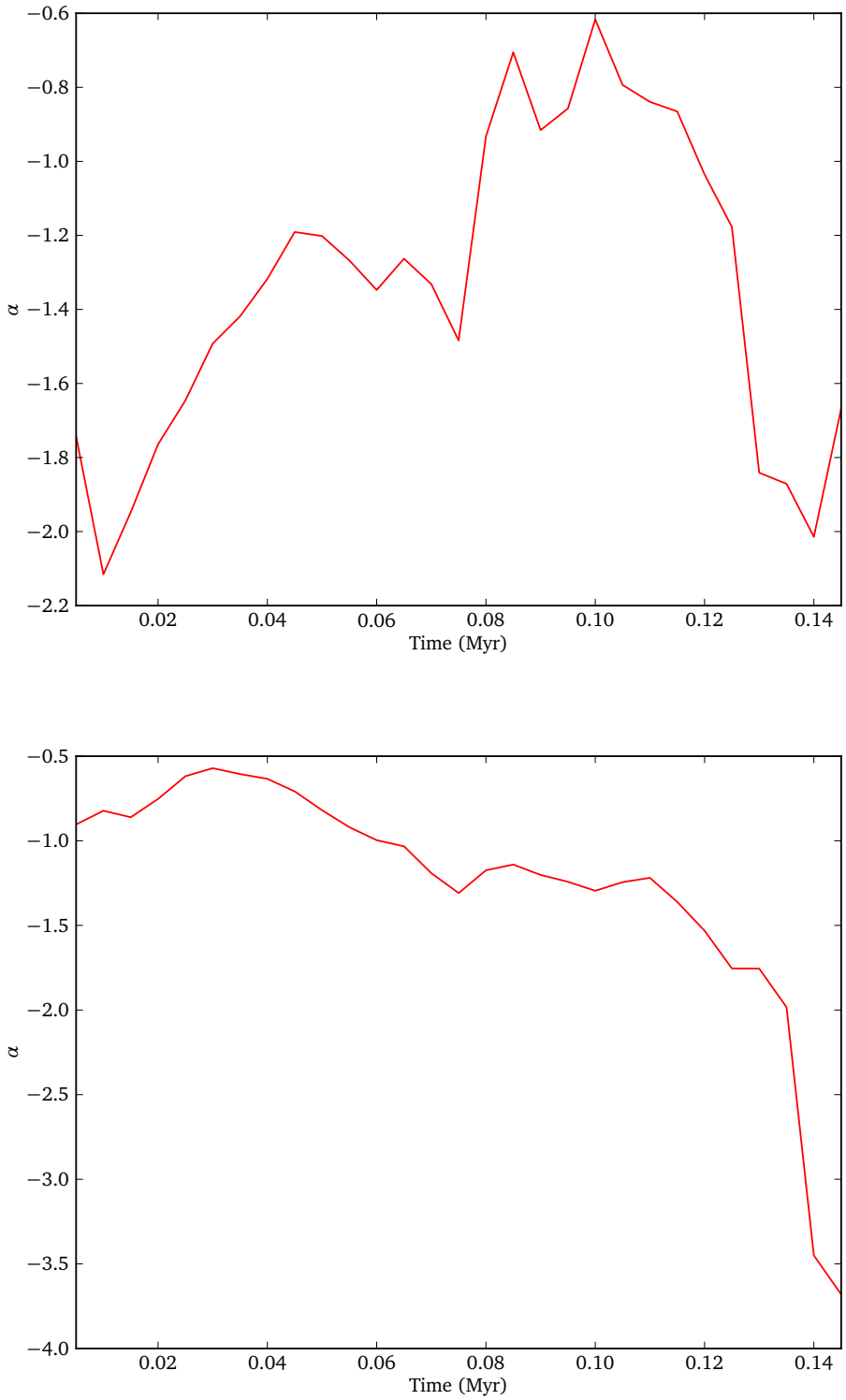


Figure 12.14 – Indices of power law fits to rates of growth for supersonic turbulent 2D simulations. The upper plot shows centre-of-mass position rates of growth; the lower plot shows column density rates of growth.

Chapter 13

NTSI – conclusions

There are two main aims of our simulations of the non-linear thin shell instability, described in Chapters 9, 10, 11 and 12. The first is to confirm the prediction of Vishniac (1994) that the rate of growth of the NTSI is given by

$$\omega \propto k^{3/2} \,. \tag{13.1}$$

This prediction was derived for one-dimensional monochromatic perturbations. Our second aim is to empirically determine similar relations for the NTSI in more complex situations, such as two-dimensional perturbations and perturbations triggered by white noise or turbulence.

First we consider whether our simulations are numerically converged. We perform simulations of one-dimensional monochromatic perturbations at both low and high resolution, as defined in Section 9.1.1. Comparison of the results of these two suites shows that although the higher-resolution simulations show more fine detail, the overall growth of the NTSI is not significantly altered. We also perform two-dimensional simulations, but limited computational resources mean that we can only run these at low resolution. We have shown that the low-resolution one-dimensional simulations produce reasonable results. As the two-dimensional simulations are the same low resolution, we assume that these simulations also produce reasonable results.

Vishniac (1994) assumed that the motions of the layer would remain subsonic, and used this to determine minimum wavenumbers unstable to the NTSI. We find that the perturbations in the layer regularly exceed the sound speed, and that we excite the NTSI at wavenumbers considerably lower than the predicted limit. Instead, we find a limiting wavenumber of approximately $10\,\mathrm{pc}^{-1}$ for our simulations.

We measure the width of the layer with time to determine a maximum resolvable wavenumber, as described in Section 9.3.1. We find that all simulation suites produce a similar rate of layer growth, which does not depend strongly on the what type of perturbations are present. The rate of layer growth is much larger than that predicted for an idealized isothermal shock. This does not seem to be due to the NTSI, as this rate of

growth does not depend strongly on the wavenumber of monochromatic perturbations, unlike the NTSI.

In general we find that rates of growth do begin to decrease, sometimes rapidly, for wavenumbers above the maximum resolvable wavenumber predicted from the growth of the layer. For 1DLR and 1DHR simulations, this prediction appears to be quite accurate, while for 2D simulations we find that rates of growth sometimes decay at wavenumbers slightly lower than the predicted maximum.

We have modelled the decay of an initial monochromatic velocity perturbation in Section 9.2.1. The model considers the total momentum of the initial perturbation and accreted material of a small column through the layer. This column is taken through a peak in the initial velocity perturbation. We find that for one-dimensional perturbations the initial decay in the rates of growth does approximately match the predicted values. The fit is not so good for two-dimensional perturbations, which decay faster than predicted.

This is likely to be caused by the transverse motion of material away from the peaks of the velocity perturbation, which is easier for a two-dimensional perturbation than a one-dimensional perturbation. Such transverse motion is not considered in our simple model, but will carry momentum away from the peaks. For a one-dimensional perturbation the transverse motion can only be along one axis to decrease the rate of growth of the perturbation, whereas the transverse motion can be along two axes for two-dimensional perturbations.

We have fitted power laws at each timestep to the rates of growth as a function of wavenumber. We have fitted these for the rates of growth of both centre-of-mass position and the column density. However, the centre-of-mass position tends to produce clearer results and more obvious trends, and so in the rest of this section the 'index' refers to the index of a power law fit to the rates of growth of centre-of-mass position. Similarly, all further references to rates of growth refer to the rates of growth of centre-of-mass position. We note that for the monochromatic simulations and the white noise simulations we expect the initial index to be zero, as we add equal amplitudes for the monochromatic perturbations and random amplitudes drawn from a uniform distribution for the white noise perturbations.

Figures 13.1 and 13.2 summarize the indices of power law fits to the centre of mass position for 1DHR and 2D simulations. We would expect that the high-resolution one-dimensional simulations of monochromatic perturbations should recover the predicted relation of Vishniac (1994). However we find that these simulations do not produce a clear relation between rate of growth and wavenumber; the index of power law fits α increases from close to zero to approximately 3 before falling again to 1.5.

These simulations may fail to produce the expected result due to the strength of the initial perturbations. The NTSI in these simulations rapidly saturates for all higher

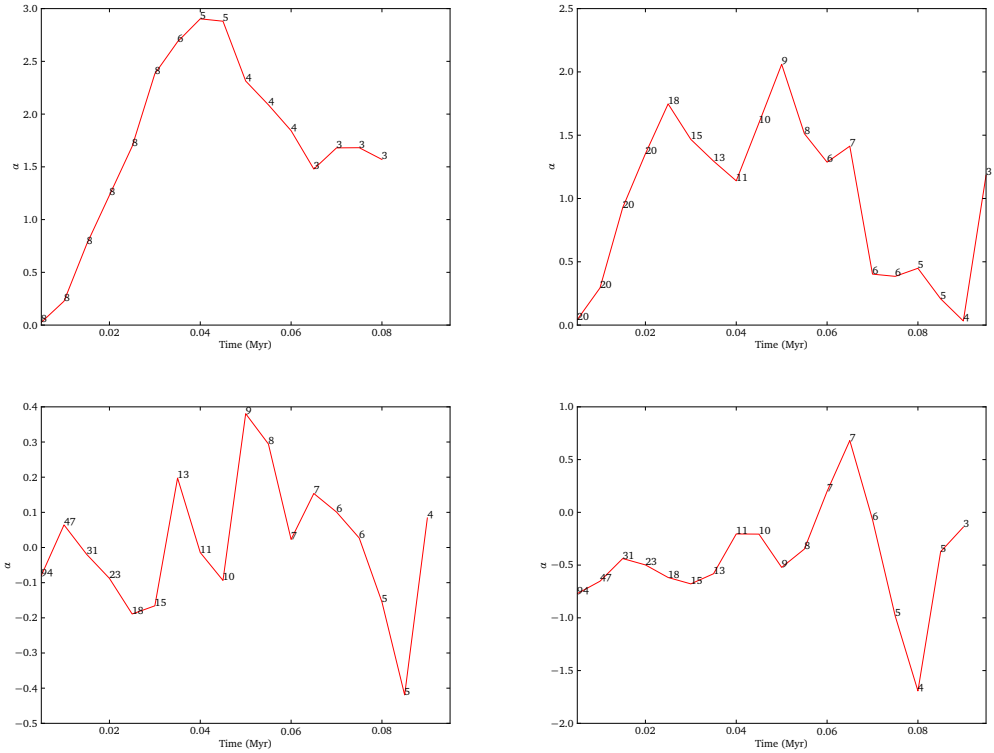


Figure 13.1 – Indices of power law fits to rates of growth for 1DHR simulations. From top left to bottom right are the fits for monochromatic perturbations, white noise perturbations, subsonic turbulence and supersonic turbulence.

wavenumbers, as can be seen in Figure 10.5. The two-dimensional monochromatic perturbations have the same peak amplitude as a fraction of collision velocity, but as the perturbation is two-dimensional the total absolute momentum of the perturbation is smaller, as less of the layer is close to a peak. In two dimensions the index reaches the expected value of 1.5. It is interesting that the NTSI appears to follow the same relation as the one-dimensional NTSI when excited in a single mode.

We see the same index of 1.5 for simulations with one-dimensional white noise. Each mode of the initial velocity perturbation has a much smaller amplitude than in the monochromatic case, allowing all modes to initially grow without interference. The power law index, while noisy, does appear consistent with the predicted relation at early times. In two-dimensions this is no longer true; the index varies considerably with time making it difficult to draw conclusions. However, the index in this case may be misleading. Examination of Figure 11.5 shows that the situation is more complex. Low wavenumbers grow more rapidly than would be expected from the NTSI for reasons that are not clear. At higher wavenumbers the rates of growth do appear to increase with increasing wavenumber at approximately the predicted rate. The excess growth at low wavenumbers decreases the index of the power law fit.

Simulations including turbulence are fundamentally different to our other simulations.

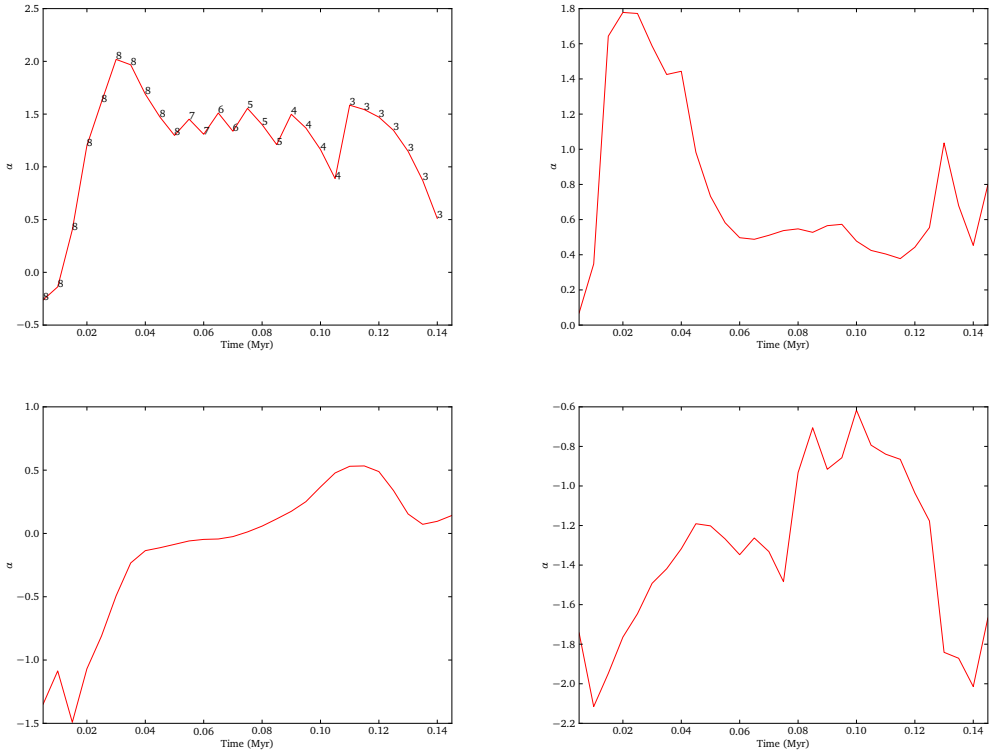


Figure 13.2 – Indices of power law fits to rates of growth for 2D simulations. From top left to bottom right are the fits for monochromatic perturbations, white noise perturbations, subsonic turbulence and supersonic turbulence.

Our other simulations include initial velocity perturbations only in a narrow zone close to the collision plane. This zone is rapidly accreted onto the layer, and the initial perturbations then decay over a short time period. For the simulations with turbulence, the turbulent velocity field covers all the inflowing gas in the simulation. The layer therefore experiences a continuous input of perturbations. The amplitude spectrum of the turbulent velocity has a power law index of -2, which gives the lowest wavenumber perturbations the greatest amplitude. This is in contrast to the growth of the NTSI, where the higher wavenumbers grow more rapidly.

We find that the indices of the power law fits alone do not convey a full description of the effect of turbulence on the NTSI. For both subsonic and supersonic turbulence, we find that the low wavenumbers grow, with decreasing rate of growth as the wavenumber increases. These perturbations are being driven directly by the turbulence as material is accreted to the layer. For subsonic turbulence, we find that this trend reverses at approximately the minimum wavenumber susceptible to the NTSI. Above this wavenumber the NTSI causes the rate of growth to increase with wavenumber up to the maximum resolvable wavenumber. Above this the rates of growth decrease rapidly as expected.

For supersonic turbulence the rate of growth of the low wavenumbers is much greater

than for the subsonic turbulence, due to the increased strength of the turbulence. As a result, the NTSI is not able to compete with the turbulence, and the rate of growth decreases with increasing wavenumber. The index of power law fits is correspondingly negative; approximately -0.5 for the one-dimensional case and between -1.5 and -0.5 for the two-dimensional case.

These results are similar for both one-dimensional and two-dimensional simulations with turbulence, although the one-dimensional simulations appear noisier. The two-dimensional simulations include a circular average and so includes contributions from a greater number of wavenumbers; this may help improve the results.

For simulations of colliding flows where the flows contain a turbulent velocity field, subsonic turbulence may excite the NTSI but will also directly excite low-wavenumber modes, and so confuse a naive power law fit. Supersonic turbulence will directly drive perturbations, but may completely suppress the NTSI.

We conclude that although we were not able to demonstrate the prediction of Vishniac (1994) for high-resolution one-dimensional monochromatic perturbations or two-dimensional white noise, we were able to show consistency with the prediction for two-dimensional monochromatic perturbations and one-dimensional white noise perturbations. We show that the NTSI in two dimensions may follow the same relationship as for the NTSI in one dimension. Finally, we demonstrate that the NTSI can be triggered by turbulence, but that turbulence will also directly excite perturbations, and strong turbulence can suppress the NTSI. Our results can be compared to those obtained from more complex simulations of colliding flows to help identify whether the NTSI is present.

Chapter 14

More realistic cloud–cloud collisions

Having simulated the non-linear thin shell instability in colliding flows in Chapter 9, we turn our attention to a more realistic simulation of cloud–cloud collision. We simulate the supersonic collision of spherical clouds. As before, the collision will produce a shock-compressed layer, but as we now include self-gravity this layer may be gravitationally unstable.

Our clouds are not initially in equilibrium, and we do not intend them to represent long-lived objects. Instead they can be viewed as transient collections of gas, formed through a combination of gravity and turbulence, which collide. While such collections are likely to contain significant density variations, we do not model these for simplicity. Furthermore, they are unlikely to be significantly centrally condensed, making a uniform density cloud a reasonable first approximation.

Computational limitations mean that we can only explore a limited area of parameter space. We therefore consider only head-on collisions of identical clouds, all of which have the same mass and the same uniform density. We limit ourselves to simulating six collisions which differ only in collision velocity.

14.1 Initial Conditions

We solve the energy equation using the method described in Section 3.4. All gas in the spheres is initially set to 10 K. At relatively low temperatures, this method assumes an average particle mass of $2.35 m_{\mathrm{H}}$, where m_{H} is the mass of a hydrogen atom, and the ratio of specific heats γ is 1.4. This gives a sound speed at 10 K of approximately $0.22\,\mathrm{km\,s^{-1}}$.

14.1.1 Size, mass and density

We begin by producing a settled box of SPH particles as described in Section 8.1. We carve a sphere of 10^7 particles from this box. We stretch each sphere to a radius r_{s} of 2.24 pc, and set particle masses to $5 \times 10^{-5}\,\mathrm{M_\odot}$, giving a total mass of $500\,\mathrm{M_\odot}$ per cloud. The density of each sphere is therefore $7.23 \times 10^{-22}\,\mathrm{g\,cm^{-3}}$. As described in Appendix A,

the Jeans mass is approximately

$$M_{\mathrm{J}} = \sqrt{\frac{375}{4\pi G^3 \rho}} c_{\mathrm{s}}^3 \,. \tag{14.1}$$

For our clouds, the Jeans mass is approximately $65\,\mathrm{M}_{\odot}$, and the cloud is initially unstable to gravitational collapse. The freefall time is

$$t_{\mathrm{ff}} = \sqrt{\frac{3\pi}{32 G \rho}} \,, \tag{14.2}$$

which for our clouds is approximately 2.48 Myr. Our simulations will therefore be significantly affected by the global collapse of the clouds only if the simulations are run for a significant fraction of this time.

The collision of two gas clouds creates a shock-compressed layer between them. Iwasaki & Tsuribe (2008) show that this layer will fragment through the gravitational instability only if the original clouds contain significantly more than one Jeans mass. As our clouds contains significantly more than one Jeans mass we expect the shock-compressed layer formed in the collision to be unstable to the gravitational instability described in Chapter 6.

14.1.2 Turbulent velocity field

We superimpose a velocity field of divergence-free subsonic turbulence, as described in Section 8.2.4, with an average speed of $0.1\,\mathrm{km\,s^{-1}}$ and an amplitude spectrum $A(k) \propto k^{-2}$. This amplitude spectrum is in agreement with both observations of velocity dispersion in molecular clouds (Larson 1981) and numerical simulations of supersonic turbulence (Porter, Pouquet & Woodward 1992). It is also the power spectrum of a step-function shock. Since Fourier transforms are additive, a box containing a large number of shocks will also have $\alpha = -2$. However, the reverse is not true: taking the Fourier spectrum of field with a power spectrum of $\alpha = -2$ will not usually produce a box full of shocks, as the location of the shocks is encoded in the phase information and not in the power spectrum.

While comparing the supersonic turbulence of the interstellar medium to our weak, subsonic turbulence is not necessarily relevant, the exact form of the turbulence should not be important. We are using the turbulence only to seed small perturbations at a range of scales. These should in turn seed instabilities in the shock-compressed layer. The turbulence is too weak to provide significant support to the spheres against gravitational collapse.

Collision	Collision speed $V_{\rm E}$ / km s^{-1}	Mach number $\mathscr{M}$
I	0.5	2.25
II	1.0	4.49
III	1.5	6.74
IV	2.0	8.99
V	3.0	13.48
VI	5.0	22.47

Table 14.1 – The range of collision speeds between clouds. The relative speed between each cloud is twice the value in this table. Mach numbers are calculated using Equation 14.3 and a sound speed of 0.22 km s^{-1}.

14.1.3 Low-density cooling

The method described in section 3.4 to solve the energy equation is used here primarily to track the thermal evolution of dense optically thick hydrostatic cores that have formed by gravitational collapse. It is not intended to track the thermal evolution of low-density diffuse gas, and for simplicity the method assumes dust and gas are thermally coupled. Since this is not true at densities of $n(\mathrm{H_2}) \lesssim 10^5\,\mathrm{cm^{-3}}$, this means the method overestimates dust cooling in low density gas, and is effectively isothermal at low densities. The analytic models of the gravitational instability and of the NTSI both assume an isothermal, and therefore rapidly cooling shock.

Marinho & Lépine (2000) collided larger, lower density clumps at the same speed as our fastest collision. They used an optically thin cooling function. Their clouds were of slightly lower density, and thus should cool even more slowly than our clumps. Although they find initial heating to a few thousand Kelvin in the shock front, they find that the dense gas in the layer rapidly cools to 10 K to 20 K. We therefore consider the implicit assumption of isothermality at low densities valid for our simulations.

14.1.4 Collision velocities

We duplicate each cloud to produce a 'left' and 'right' cloud, which we place so they are initially touching at the origin, with the left cloud centred at $(-r_{\rm s}, 0, 0)$ and the right cloud centred at $(r_{\rm s}, 0, 0)$. A velocity $V_{\rm E}$ is added to every particle in the left cloud, and a corresponding velocity $-V_{\rm E}$ is added to every particle in the right cloud. Table 14.1 lists the collision velocities we simulate, together with the equivalent Mach number. We label these collisions as collisions I, II, III, IV, V and VI, where collision I is the lowest-velocity collision and collision VI is the highest-velocity collision.

When discussing cloud velocities we refer to the cloud speeds in the centre of mass frame. The relative collision velocity between the clouds is twice as large. We use this definition as we are interested in the speed of the collision with respect to the stationary

layer that forms between the clouds. We define the Mach number of the collision $\mathscr{M}$ as

$$\mathscr{M} = \frac{V_{\rm E}}{c_{\rm s}} , \tag{14.3}$$

where $V_{\rm E}$ is the individual cloud velocity and $c_{\rm s}$ is the sound speed in the clouds.

If the collision is approximately isothermal, then as described in Section 6.2.1 the density is expected to be a factor of $\mathscr{M}^2$ higher than in the clouds. For our fastest collision this gives a density in the layer of $3.6 \times 10^{-19}\,{\rm g\,cm^{-3}}$, using the Mach numbers calculated in Table 14.1. If the collision is not isothermal then the layer will heat and the maximum density reached will be lower, making this density an upper estimate of the possible density in the layer. At these low densities gas remains cold due to rapid cooling, making it a reasonable assumption that the collision is approximately isothermal.

14.2 Timescale of the gravitational instability

The collision will create a dense shock-compressed layer. The fragmentation of such a layer through the gravitational instability is treated in Section 6.2. Gravitational fragmentation is predicted to begin in a time

$$t_{\rm transition} \approx \frac{1.2}{\sqrt{2\pi G \rho_{\rm ext} \mathscr{M}}} . \tag{14.4}$$

For our simulations, this is

$$t_{\rm transition} \approx \frac{2.2\,{\rm Myr}}{\sqrt{\mathscr{M}}} . \tag{14.5}$$

The fragmentation will become non-linear in a time

$$t_{\rm nl} \approx \frac{2.4\,\delta_0^{-0.1}}{\sqrt{2\pi G \rho_{\rm ext} \mathscr{M}}} , \tag{14.6}$$

where δ_0 is the amplitude of the initial perturbations. We do not know the size of initial perturbations, and instead estimate $\delta_0 \sim 1$. As the non-linear time is a very weak function of the amplitude, this should still give us an order of magnitude estimate. For our simulations, this non-linear time is

$$t_{\rm nl} \approx \frac{4.4\,{\rm Myr}}{\sqrt{\mathscr{M}}} . \tag{14.7}$$

The estimated transition times and non-linear times for our simulations are shown in Table 14.2.

Iwasaki & Tsuribe (2008) find that the first mode to become non-linear, and therefore

Mach number $\mathscr{M}$	$t_{\rm transition}$ (Myr)	$t_{\rm nl}$ (Myr)	$\lambda_{\rm frag}$ (pc)	$k_{\rm frag}$ (pc^{-1})
2.25	1.455	2.909	0.86	1.16
4.49	1.023	2.057	0.61	1.64
6.74	0.840	1.680	0.50	2.01
8.99	0.727	1.455	0.43	2.32
13.48	0.594	1.188	0.35	2.84
22.47	0.460	0.920	0.27	3.66

Table 14.2 – The estimated timescales and fragmentation wavelength of gravitational instability as a function of collision Mach number. The transition time $t_{\rm transistion}$ is the earliest time at which the gravitational instability dominates over layer accretion, and the non-linear time $t_{\rm nl}$ is the time at which growth become non-linear. The wavelength $\lambda_{\rm frag}$ is the wavelength of the mode which first becomes non-linear.

dominate the fragmentation of the layer, has a wavenumber

$$k_{\rm frag} = 2\frac{\sqrt{G\rho_{\rm ext}\mathscr{M}}}{c_{\rm s}\sqrt{2\pi}}\,, \tag{14.8}$$

as described in Section 6.3. The equivalent wavelength is then

$$\lambda_{\rm frag} = \frac{c_{\rm s}\sqrt{2\pi}}{2\sqrt{G\rho_{\rm ext}\mathscr{M}}}\,, \tag{14.9}$$

which for our simulations is

$$\lambda_{\rm frag} = \frac{1.3\,{\rm pc}}{\sqrt{\mathscr{M}}}\,. \tag{14.10}$$

Table 14.2 shows this expected fragmentation wavelength for each of our simulations.

14.3 Results

We perform our simulations using the SEREN code described in Section 3.1. We include both hydrodynamics and self-gravity. We solve the energy equation using the method described in Section 3.4. Dense, bound objects are replaced with sink particles as described in Section 3.7; we use a sink density of $10^{-11}\,{\rm g\,cm^{-3}}$.

Due to computational limitations, we are only able to follow the evolution of the collisions for a limited time. We follow the four lowest-velocity collisions, collisions I to IV, until 2% of the mass of the clouds has been converted into sink particles. The highest-velocity collision, collision VI, does not produce sinks, and we are able to follow the entire collision and subsequent expansion of the layer. The remaining simulation, collision V, is only followed until shortly after the formation of the first sink particle.

Each collision produces a dense shock-compressed layer as expected. We define the accretion fraction as the fraction of the original spheres that has been accreted onto the

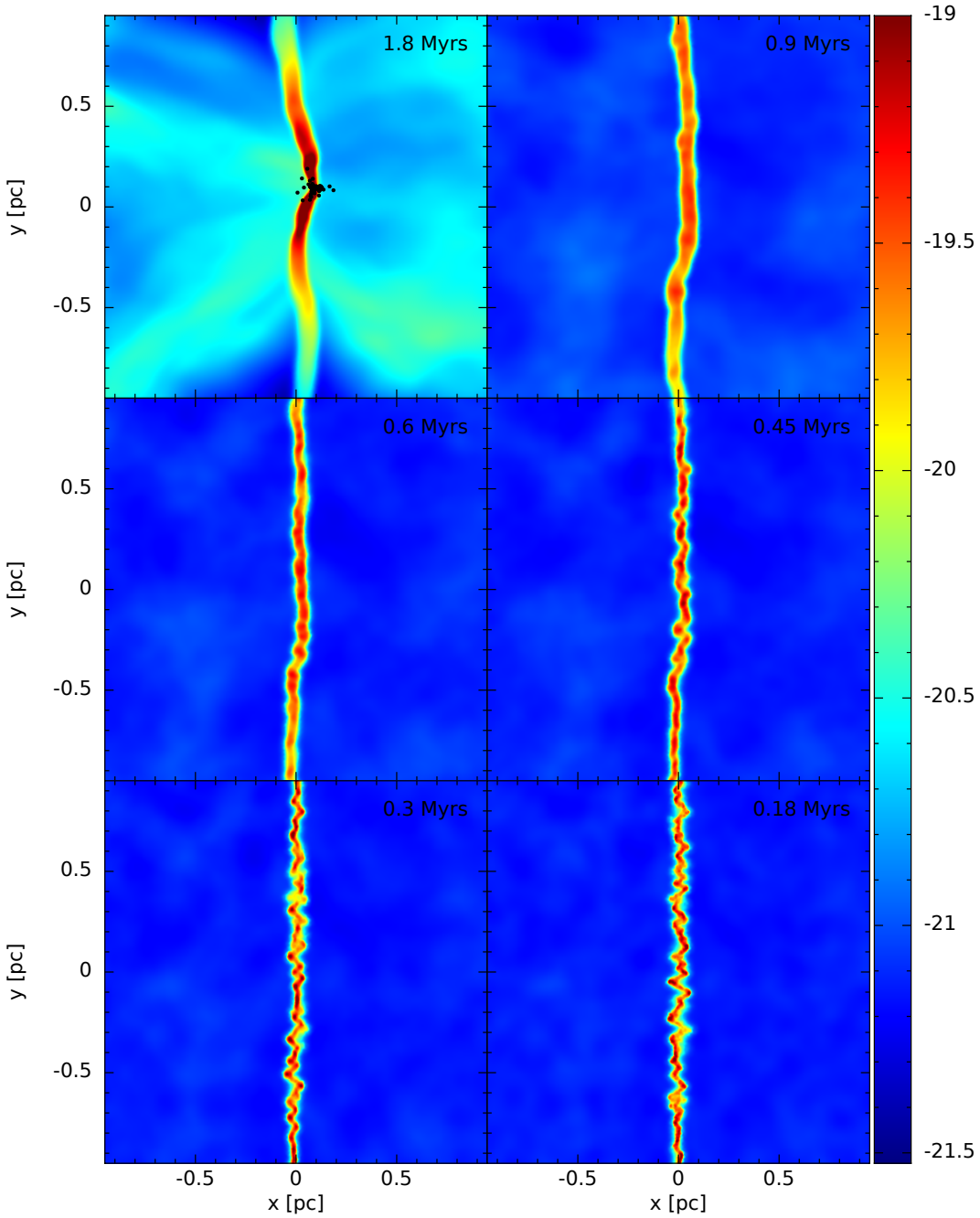


Figure 14.1 – Cross-sections of density $\log_{10}(\mathrm{g\,cm^{-3}})$ in the xy plane at equivalent collision times. From top left to bottom right: Mach 2.3, 4.5, 6.7, 9.0, 13.5 and 22.5.

layer. We then define 'equivalent times' as the time it takes for each collision to reach a specified accretion fraction.

Figure 14.1 shows a cross-section for each collision at such equivalent times. As the lowest-velocity collision takes longer to reach the specified accretion fraction, the spheres have had more time to collapse globally. For the higher-velocity collisions the global collapse of the spheres is not so significant.

For the lower collision velocities the layer is wider, as expected. At equivalent times

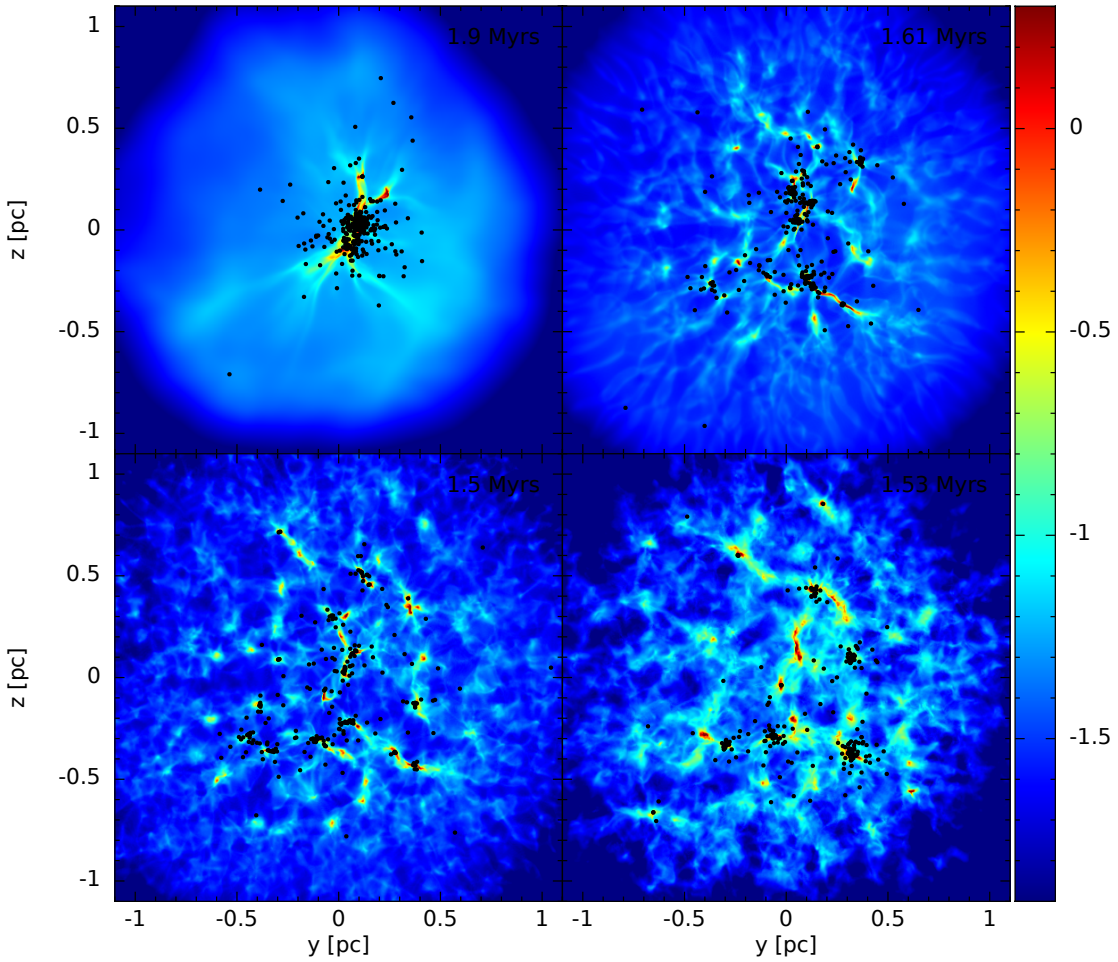


Figure 14.2 – Column density in $\log_{10}(\mathrm{g\,cm^{-2}})$ in the yz plane at approximately the time where 2% of the mass of the clouds has been converted into sink particles. From top left to bottom right: Mach 2.3, 4.5, 6.7 and 9.0.

the layer will have the same surface density, but as described in Section 6.2.1 the volume density of the layer is proportional to the square of the Mach number. The thickness of the layer will therefore be smaller for higher Mach numbers.

As the collision velocity increases and the layer thickness decreases, the layer becomes more perturbed; for the higher-velocity collisions strong bending modes can be seen. This is due to the non-linear thin shell instability, which as predicted in Chapter 7 is more easily excited in a thin layer.

14.3.1 Low-velocity collisions

We now consider the fragmentation of the layer for the four lowest-velocity collisions, collisions I to IV. Figure 14.2 shows the column density through the layer at the time where 2% of the gas has been accreted onto sink particles. As the background gas is approximately uniform in density, the column density traces the surface density of the layer. Gravitational instability in the layer causes fragmentation, and the subsequent

formation of sink particles.

For the lowest-velocity collision, collision I, the global collapse of the spheres somewhat distorts the results. For this collision fragmentation and sink formation only occur in the innermost regions of the layer where the density is highest. A number of dense filaments can be seen which radiate from the centre. A dense cluster of sinks is seen at the centre of the layer.

This can be compared to collision II, where the fragmentation is not as centrally concentrated but is spread out over the layer. A network of filaments covers the central regions of layer. Sinks form along these filaments and at the nodes between filaments. These sinks then form a loose network of subclusters.

For collisions III and IV we see similar networks of filaments, but we also see a more clumpy background structure. These clumps are a result of the increasing influence of the NTSI. Sinks form both along the filaments and in the clumps, creating a more distributed pattern of star formation than collisions I or II.

Figure 14.3 shows column density as a function of time for collision II, with a Mach number of 4.5. This simulation appears unaffected by the NTSI, and is not significantly corrupted by the global collapse of the spheres. It is therefore a good demonstration of gravitational instability leading to fragmentation in the layer.

At early times weak density perturbations can be seen, probably created by the initial turbulence. Gravitational instability causes some of these perturbations to collapse, forming dense clumps which then collapse to form sink particles. Later the formation of a dense filament can be seen. Gas flows towards the dense clumps and their sink particles, fuelling the accretion of the sink particles and triggering the formation of new sink particles which are subsequently ejected. This results in a loose network of subclusters distributed over the central regions of the layer.

14.3.2 High-velocity collisions

The highest-velocity collision, collision VI, does not undergo gravitational instability and does not produce sink particles during the collision. Figure 14.4 shows density cross-sections at different times to show the destructive nature of the collision. At the earliest times, a thin dense layer forms. This begins to undergo bending-mode instabilities, which are a characteristic of the NTSI. The NTSI drives these bending modes to high amplitude until the layer is effectively destroyed. The interface between the two clouds becomes bloated and turbulent, and allows efficient mixing of the two clouds.

Partly as a result of the NTSI, gravitational instability is unable to operate before the collision has ended. Once the confining ram pressure of the collision is removed, the bloated turbulent layer initially expands. As the original clouds were gravitationally unstable, the expanding layer is also still gravitationally unstable. Following the collision and the initial expansion of the bloated layer, the layer collapses gravitationally as

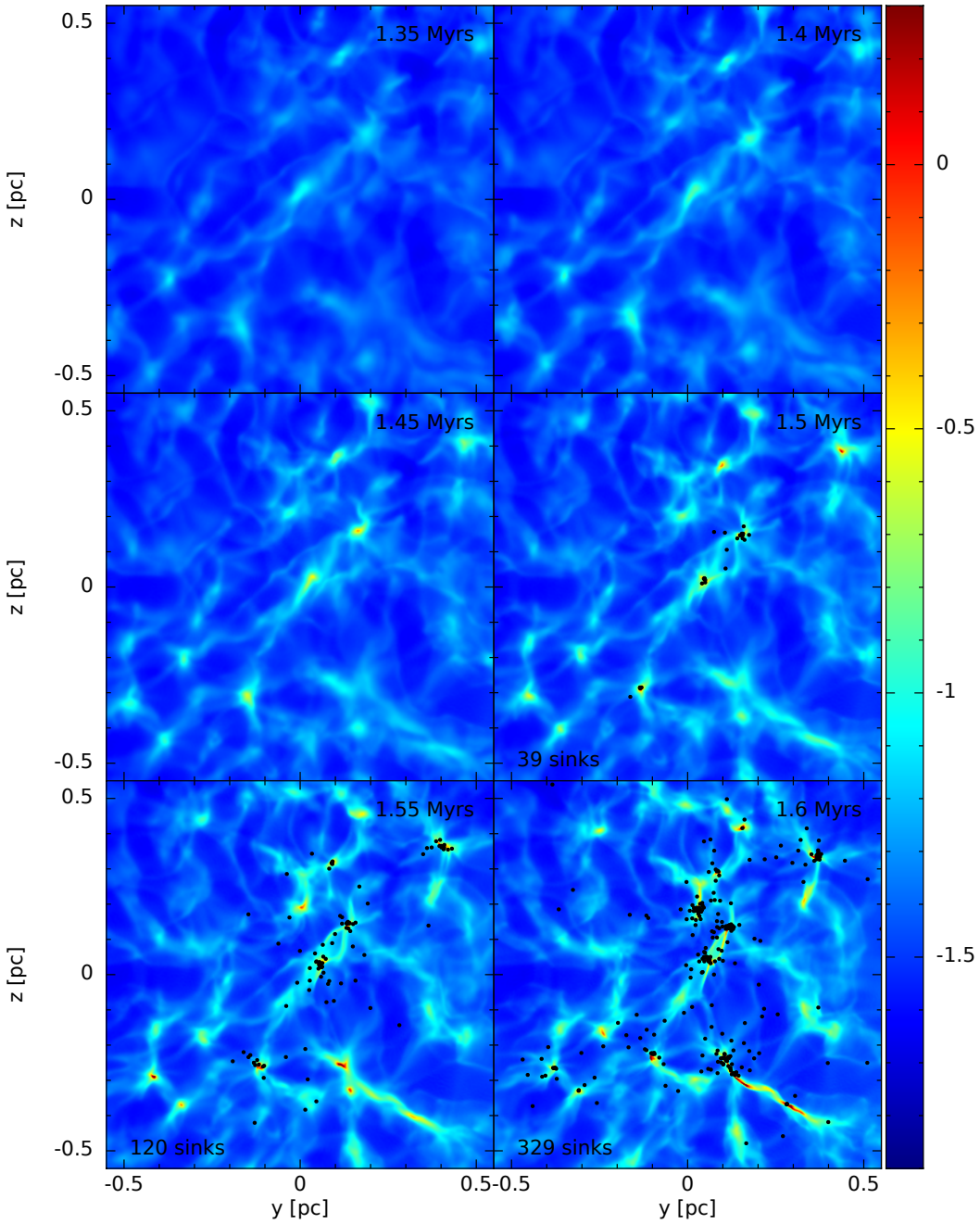


Figure 14.3 – Column density in $\log_{10}(\mathrm{g\,cm^{-2}})$ in the yz plane for the Mach 4.5 collision.

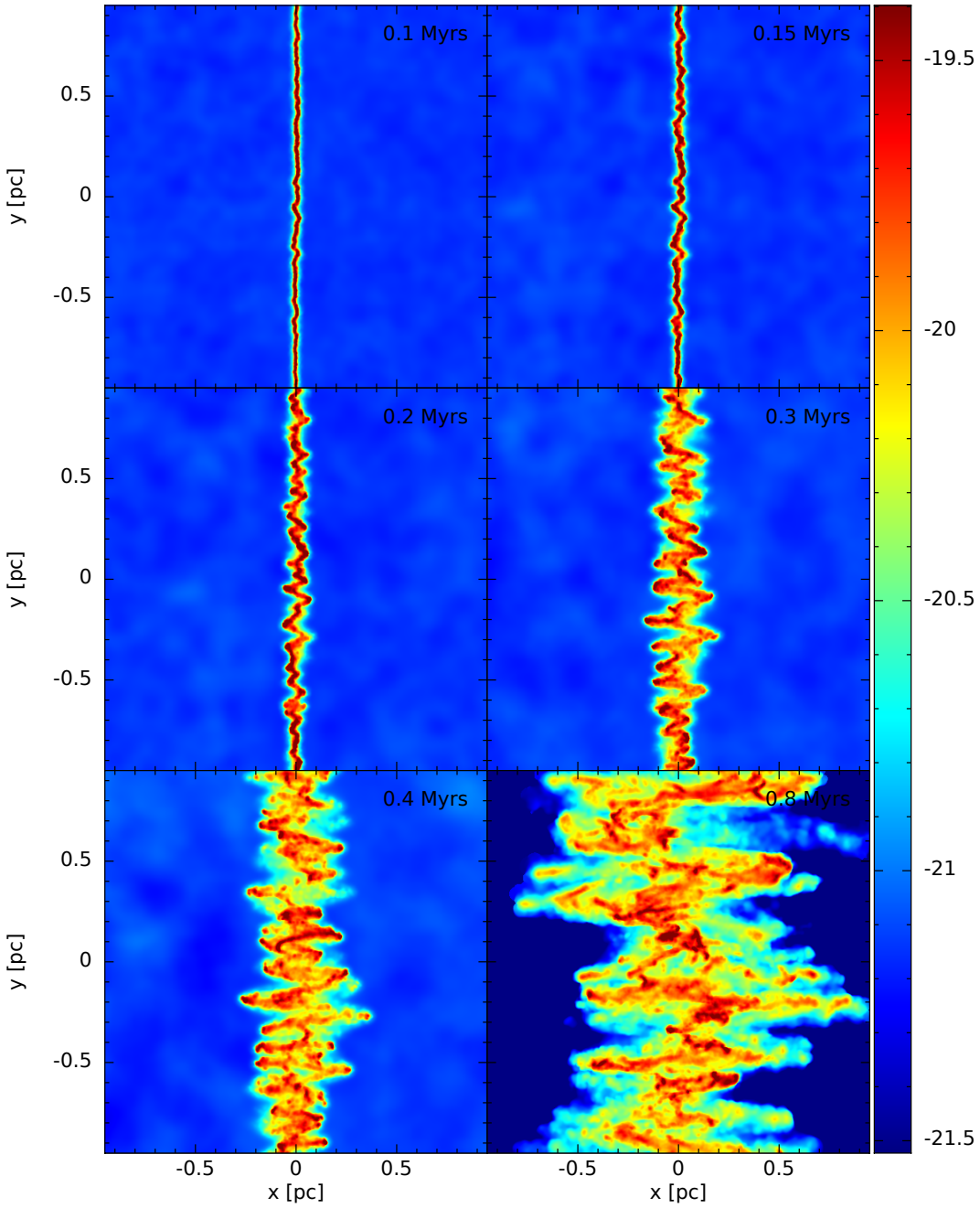


Figure 14.4 – Cross-sections of density in $\log_{10}(\mathrm{g\,cm^{-3}})$ in the xy plane for the Mach 22.5 collision.

expected. However, we are not interested in this portion of the evolution.

14.4 Instabilities

In order to demonstrate the nature of the instabilities in our collisions, we need to analyse the perturbations present in the layer in more detail. We use the same analysis techniques as are used in Chapter 9.

We use a snapshot output from each simulation every 0.05 Myr from 0 Myr to 1.5 Myr. We use only the central regions of each simulations, defined as a cubic box centred on the origin with sides of length 4 pc. We convert the regions within this box for each snapshot to a 256^3 grid of density using the SPLASH visualization tool, as described in Section 5.1.3. We filter this grid, as described in Section 5.2, to identify which grid cells are part of the layer.

Once we have identified the regions of the layer, we calculate the centre-of-mass positions of the layer along the collision axis, as described in Section 5.3.1, and the surface density of the layer, as described in Section 5.3.2. We take a Fourier transform of these quantities, as described in Section 5.3.3, to produce a Fourier spectrum for each. As these are two-dimensional perturbations this includes taking the circular average of the two-dimensional Fourier transform.

For each simulation, we construct time series consisting of the Fourier spectrum of the centre-of-mass positions and the Fourier spectrum of the surface density for each snapshot. We then calculate rates of growth as a function of wavenumber at all but the first and last timestep, as described in Section 5.3.4.

14.4.1 Layer width and thickness

We calculate the layer width and thickness as described in Section 5.3.5. The thickness is the distance across the layer, including only regions identified as part of the layer and not including voids within the layer structure, while the width is simply the distance from the top to the bottom of the layer. Both are subsequently corrected for the bending angle of the layer.

Figure 14.5 shows the layer thicknesses and widths for our collisions. For our simulations of the NTSI in Chapters 10, 11 and 12 we did not find any trends between different types of perturbations and the rate at which the layer grew in size. Here we see a trend of increasing layer thickness with increasing collision velocity.

This does not match the predictions for an idealized isothermal shock, where the thickness of the layer should be inversely proportional to the Mach number. Even for our lowest-velocity collision the thickness of the layer exceeds the predicted thickness by at least a factor of 2; the higher-velocity collisions exceed the prediction by much larger factors. The anomalous expansions of the layer is due to the additional perturbations in the layer. At low collision velocities these perturbations will be due to the continuous

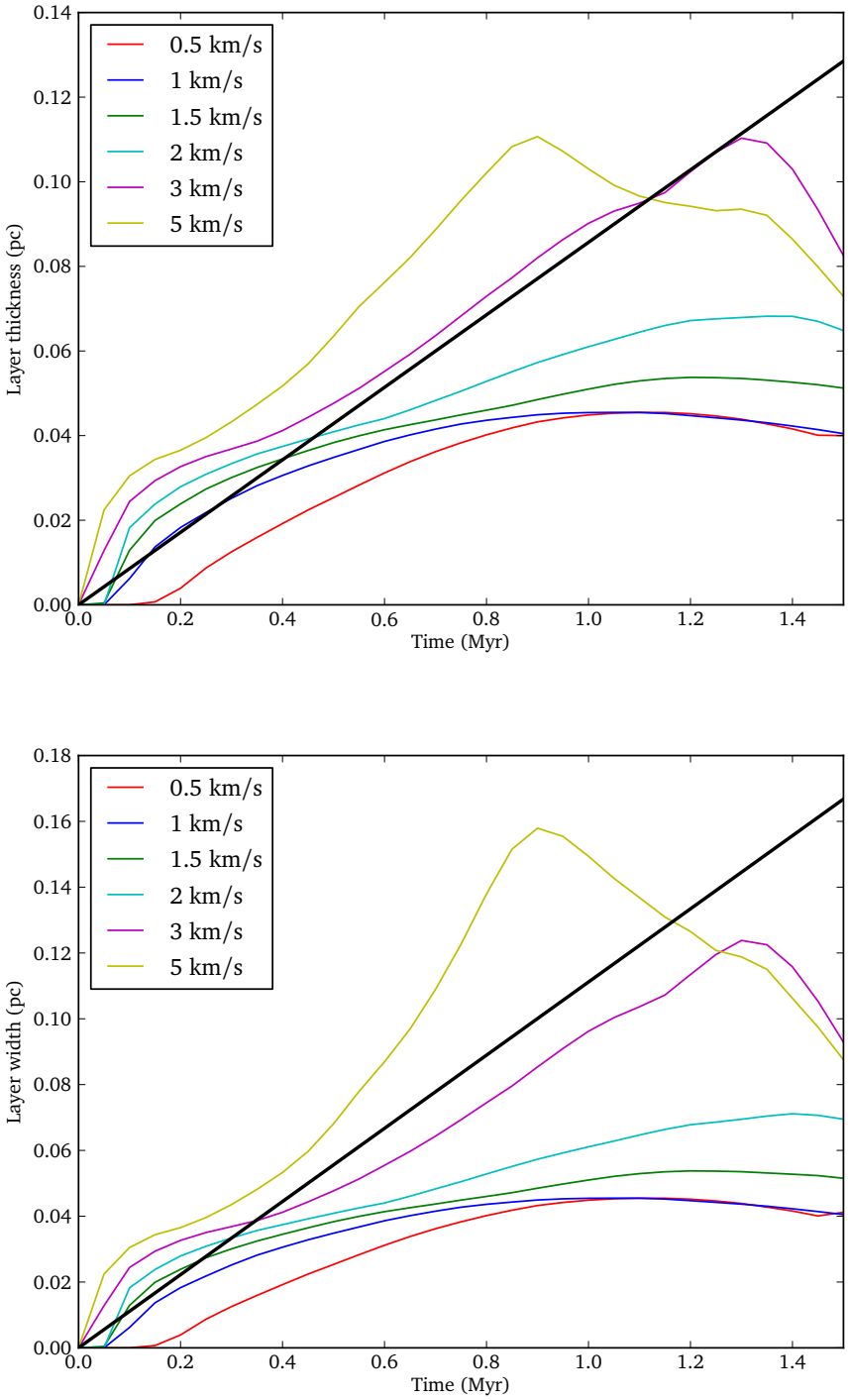


Figure 14.5 – Thickness and width, as defined in Section 5.3.5, with respect to time. Each coloured line represents the averaged results from the labelled collision. Each set of realizations is fitted with a straight line; the solid black line indicates the average of those fits. The fits were constrained to pass through the origin.

accretion of subsonic turbulence onto the layer, while at higher collision velocities the NTSI will dominate.

For the lowest-velocity collisions, the layer thickness increases initially, before decreasing at later times. This is due to the initial expansion of the layer due to the weak turbulent perturbations, and the subsequent contraction of the layer under self-gravity. This contraction may be concentrated at dense clumps and filaments.

Similar trends can be seen for intermediate-velocity collisions, although the layer grows more rapidly in these collisions. The layer grows fastest in the highest-velocity collision until the collision ends. The end of the collision causes the peak in collision V and collision VI; once the original spheres have been completed accreted into the layer the confining ram pressure is removed and the layer expands.

We calculate a maximum resolvable wavenumber as described in Section 9.3.1, which we include on plots of rates of growth but otherwise do not use. As the layer width in each simulation only agrees with the averaged fit to within a factor of a few, these estimates are often inaccurate. Furthermore, we are only interested in intermediate wavenumbers, as the NTSI is not triggered at very low wavenumbers. We therefore use fixed limits of $6\,\mathrm{pc}^{-1}$ to $20\,\mathrm{pc}^{-1}$ when fitting power law slopes to rates of growth.

14.4.2 Rates of growth

Figure 14.6 shows the rates of growth of centre-of-mass position and column density for collision I, the lowest-velocity collision. The index of power law fits to each timestep are also shown. Both the centre-of-mass position and column density rates of growth appear decrease with increasing wavenumber, giving a negative index to power law fits.

These rates of growth show no evidence of the NTSI, where rates of growth would be expected to increase with increasing wavenumber. Instead, the rates of growth can be compared with those obtained in Section 12.4 for simulations with supersonic turbulence. The turbulence in our cloud collisions is not supersonic, but due to the low collision velocity the NTSI is not excited, and the continuous input of turbulence remains dominant.

Figure 14.7 shows the same results as Figure 14.6 for collision II. The rates of growth of the centre-of-mass position are very similar to those for collision I. However, the rates of growth of column density show a slight rise in rates of growth at intermediate wavenumbers. This rise may be due to the NTSI growing at these smaller wavenumbers. We note in Section 12.4 there is some suggestion of a similar rise in the rates of growth of column density; we suggest that in such simulations the NTSI may appear first in the column density rather than the centre-of-mass position.

Figure 14.8 shows that the rates of growth for collision III are similar to those in collision II. There is still little evidence for the appearance of the NTSI in the rates of growth of the centre-of-mass. The column density shows the same increases at intermediate wavenumbers as before, although stronger than seen for collision II.

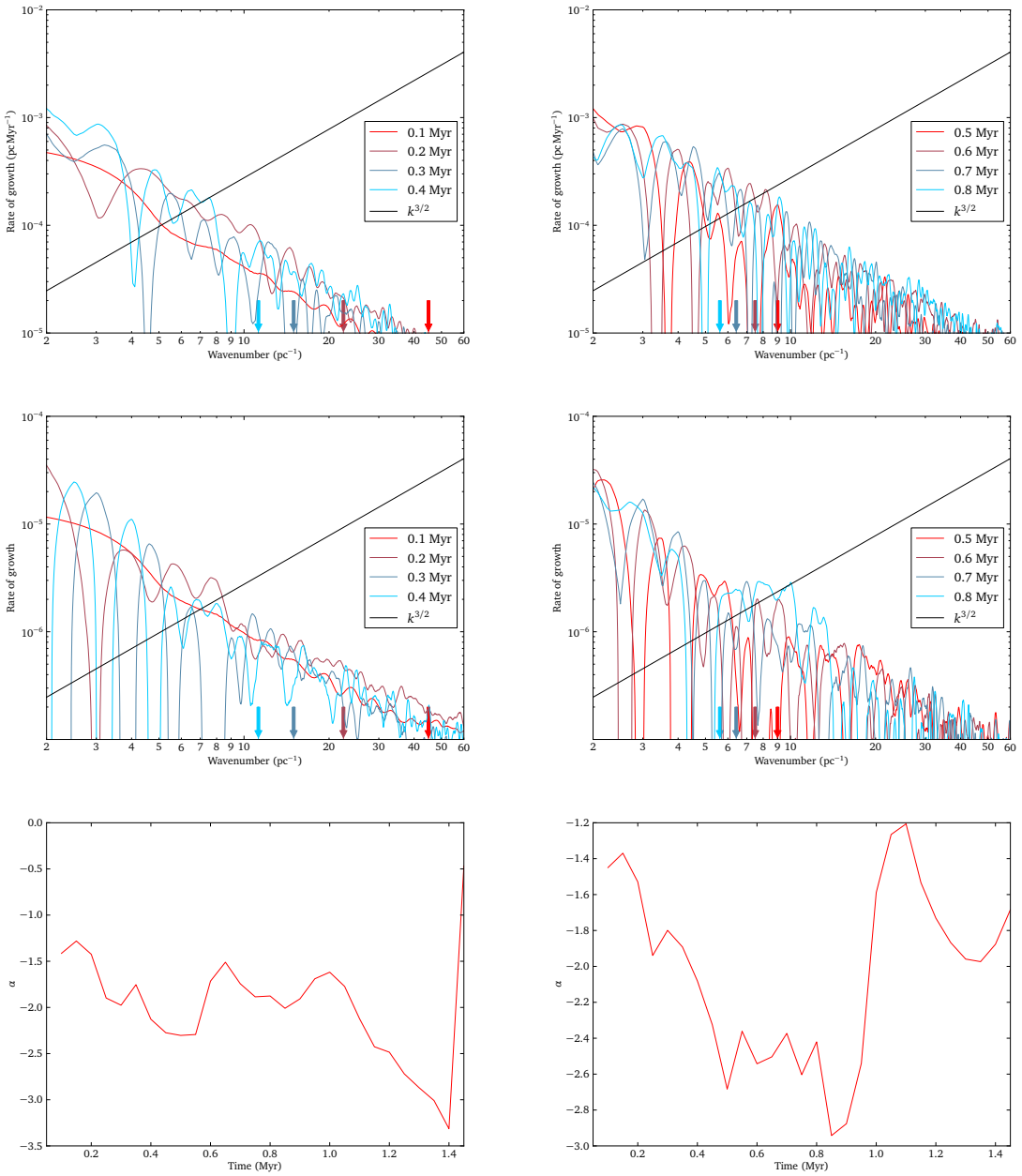


Figure 14.6 – The rates of growth of centre-of-mass position for collision I as a function of wavenumber. Each line represents results from a different timestep. The upper plots are for rates of growth of centre-of-mass position; the centre plots are rates of growth of column density. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep. The lower left plot and lower right plot show the index of power law fits to the rates of growth of centre-of-mass position and column density respectively.

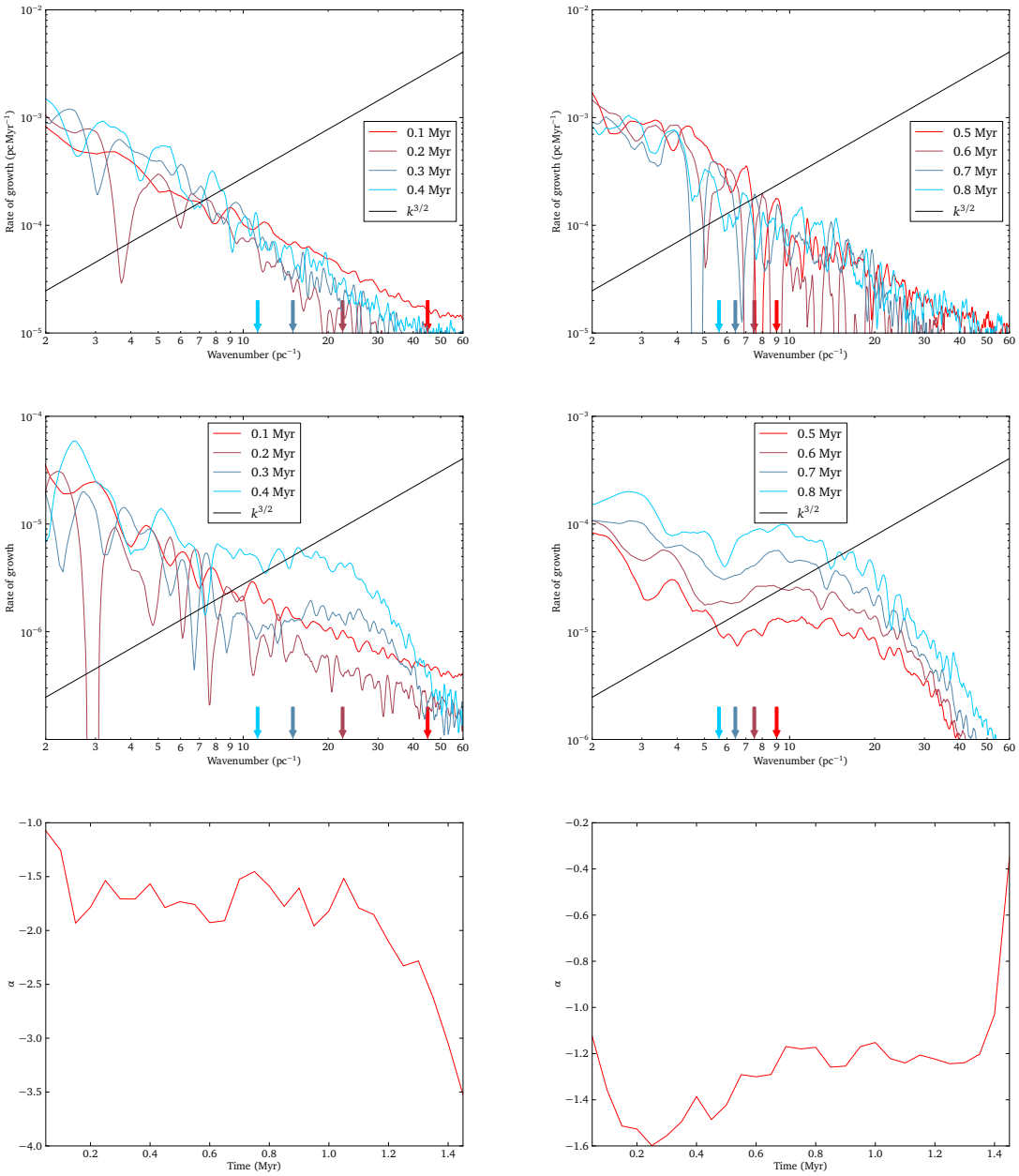


Figure 14.7 – The rates of growth of centre-of-mass position for collision II as a function of wavenumber. Each line represents results from a different timestep. The upper plots are for rates of growth of centre-of-mass position; the centre plots are rates of growth of column density. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep. The lower left plot and lower right plot show the index of power law fits to the rates of growth of centre-of-mass position and column density respectively.

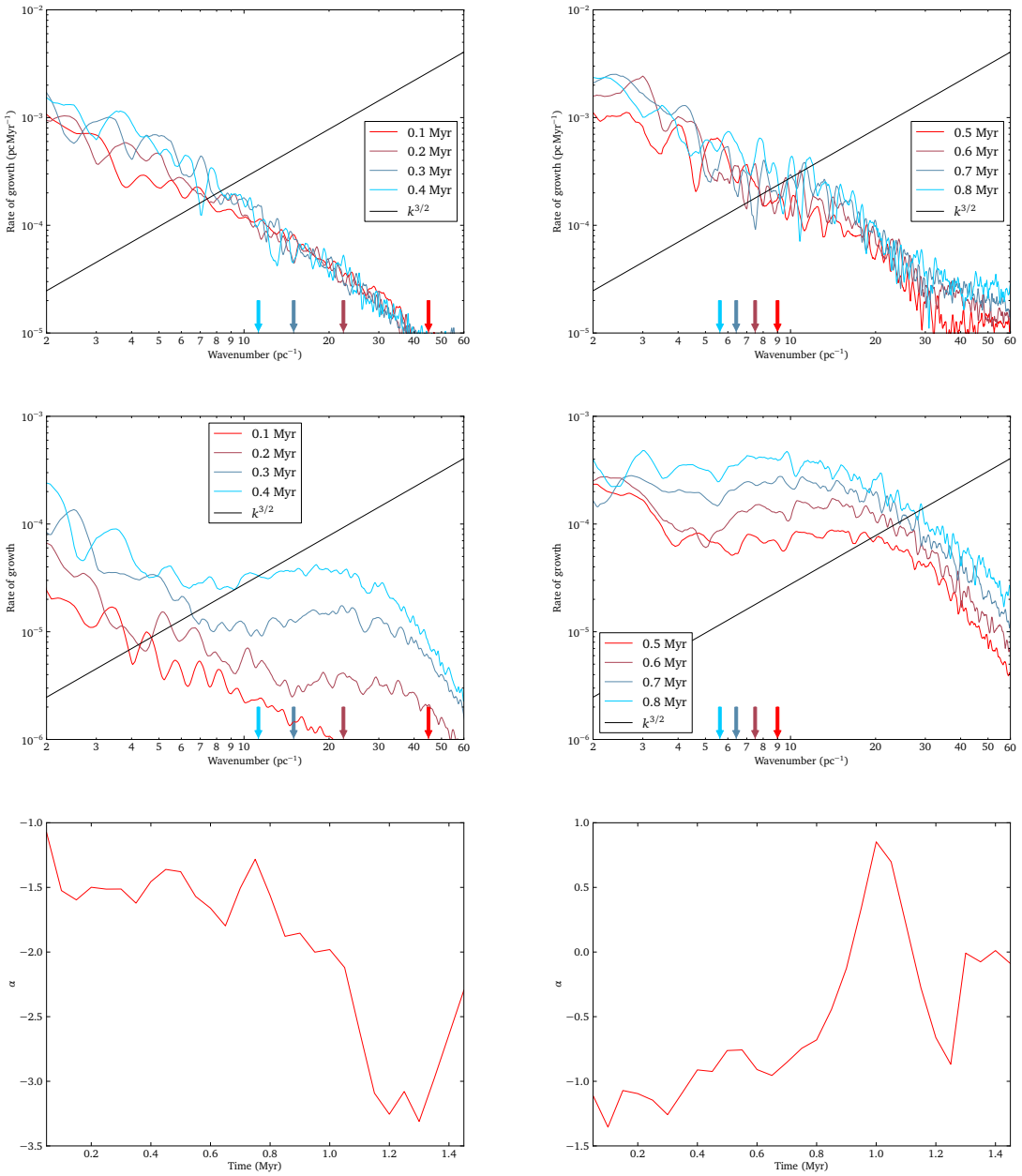


Figure 14.8 – The rates of growth of centre-of-mass position for collision III as a function of wavenumber. Each line represents results from a different timestep. The upper plots are for rates of growth of centre-of-mass position; the centre plots are rates of growth of column density. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep. The lower left plot and lower right plot show the index of power law fits to the rates of growth of centre-of-mass position and column density respectively.

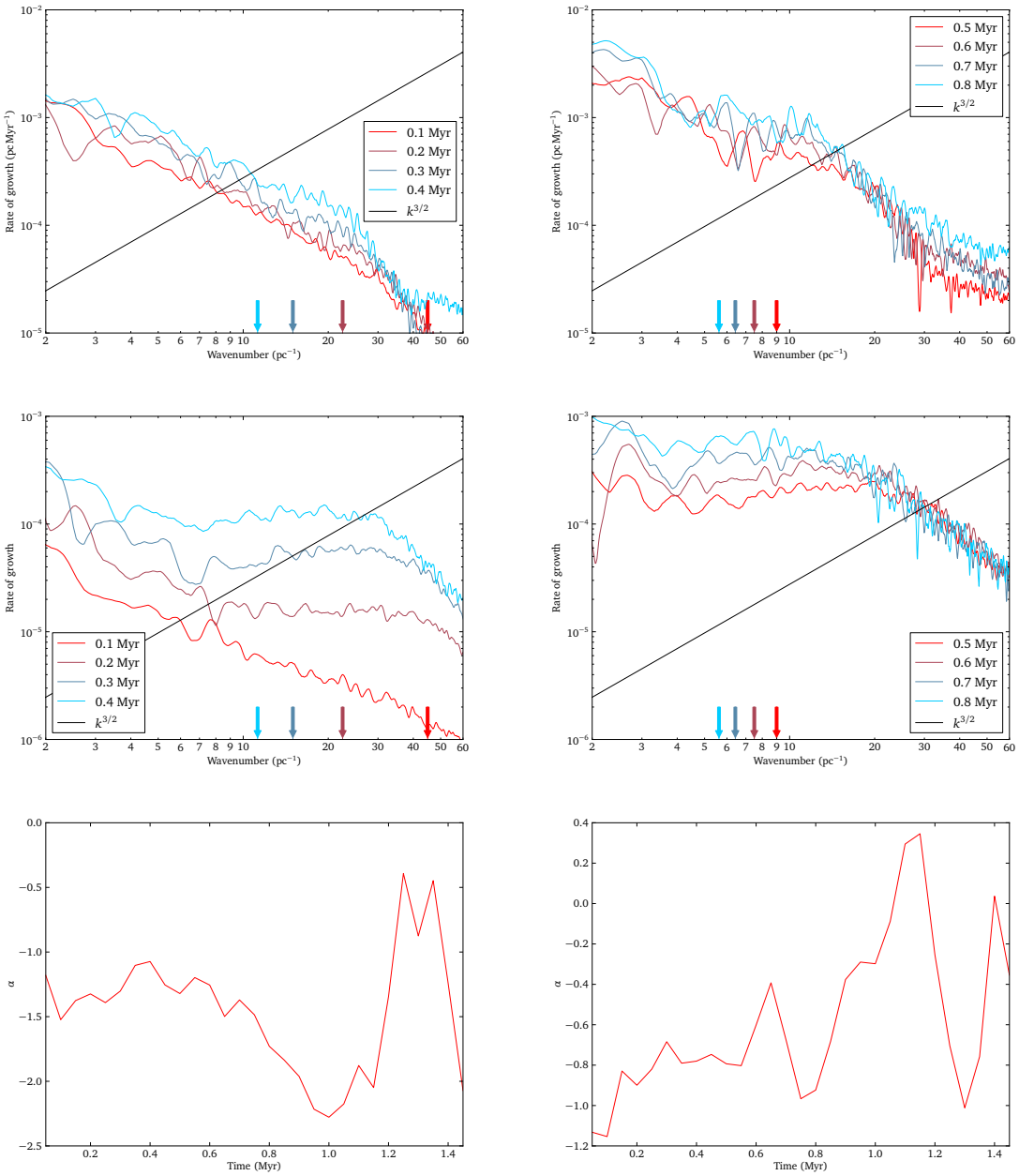


Figure 14.9 – The rates of growth of centre-of-mass position for collision IV as a function of wavenumber. Each line represents results from a different timestep. The upper plots are for rates of growth of centre-of-mass position; the centre plots are rates of growth of column density. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep. The lower left plot and lower right plot show the index of power law fits to the rates of growth of centre-of-mass position and column density respectively.

Figure 14.9 shows the results for collision IV. This is the first collision to show a small excess in the rates of growth of centre-of-mass position at intermediate wavenumbers. This excess is likely due to the NTSI. In the rates of growth of column density, a large excess can be seen at intermediate wavenumbers which is due to the growth of the NTSI, although the indices of the power law fits still remains negative for both centre-of-mass position and column density rates of growth.

Figure 14.10 shows the results for collision V. There is now a clear excess in the rates of growth of centre-of-mass position for wavenumbers between approximately $6\,\mathrm{pc}^{-1}$ and $20\,\mathrm{pc}^{-1}$, which is the range of wavenumbers we choose to perform power law fits over. The rates of growth of the column density is more confused, and is similar to the flat distributions seen for subsonic turbulence in Section 12.2. As before, we attribute the excess of rates of growth to the development of the NTSI at higher collision velocities. The sharp drop in the power law indices after approximately 1.4 Myr is due to the rarefaction shock at the rear of the clouds reaching the layer at the end of the collision.

Figure 14.11 shows the results for collision VI, the highest-velocity collision. Even at this velocity, the NTSI is still not completely dominant over the input of turbulent perturbations at low wavenumber. The index of power law fits for the centre-of-mass position approaches and briefly exceeds zero for the first time. The rates of growth of centre-of-mass position begin to show steepening positive gradients, albeit over a very small range of wavenumbers. At around 0.8 Myr, the rates of growth decrease dramatically as the rarefaction shock at the rear of the clouds reaches the layer at the end of the collision. The sharp drop in the power law indices at the same time is also due to this effect.

We conclude that the NTSI has no impact on the low-velocity collisions. Examination of the rates of growth, together with the pattern of bending modes show in cross-section views, shows that the NTSI has a strong impact on the layer at higher collision velocities. However, this is modulated by the turbulence which excites it, which is stronger at low wavenumbers, and so the combined rates of growth do not reproduce the relationship of Vishniac (1994).

If we consider the absolute values of the rates of growth of centre-of-mass position over all our simulations, we find that the maximum rates of growth are approximately $0.01\,\mathrm{pc\,Myr}^{-1}$. This is only about 5% of the sound speed, making the motion of the layer subsonic. This is in contrast to our simulations of the NTSI in Chapters 10, 11 and 12, where the motion of the layer is often supersonic.

14.5 Star formation properties

In this section, we describe results obtained from our sink particles. Each sink represents a protostar which accretes mass through its gravitational attraction. We have followed the four lowest-velocity collisions until 2% of the original mass of the clouds has been

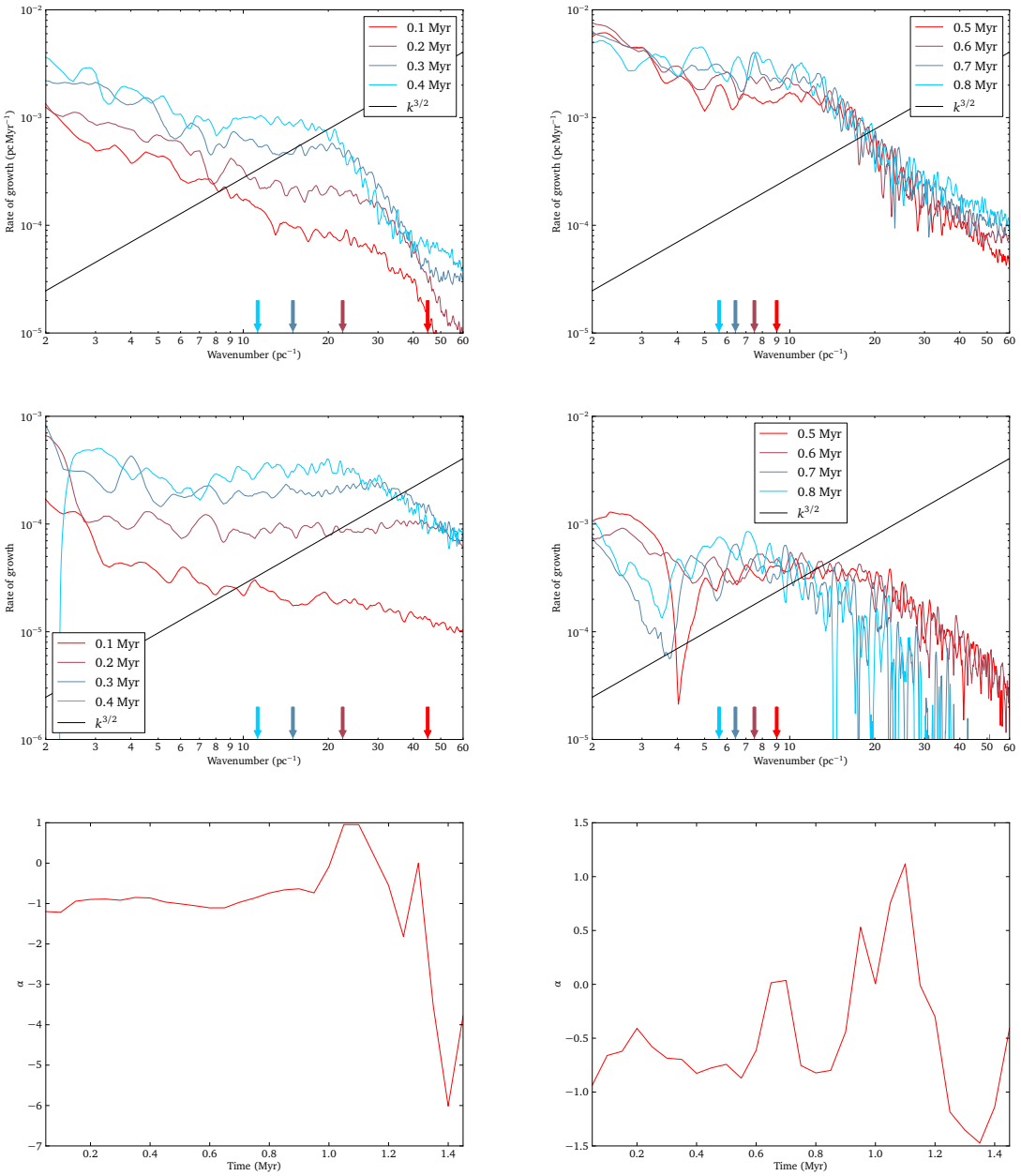


Figure 14.10 – The rates of growth of centre-of-mass position for collision V as a function of wavenumber. Each line represents results from a different timestep. The upper plots are for rates of growth of centre-of-mass position; the centre plots are rates of growth of column density. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep. The lower left plot and lower right plot show the index of power law fits to the rates of growth of centre-of-mass position and column density respectively.

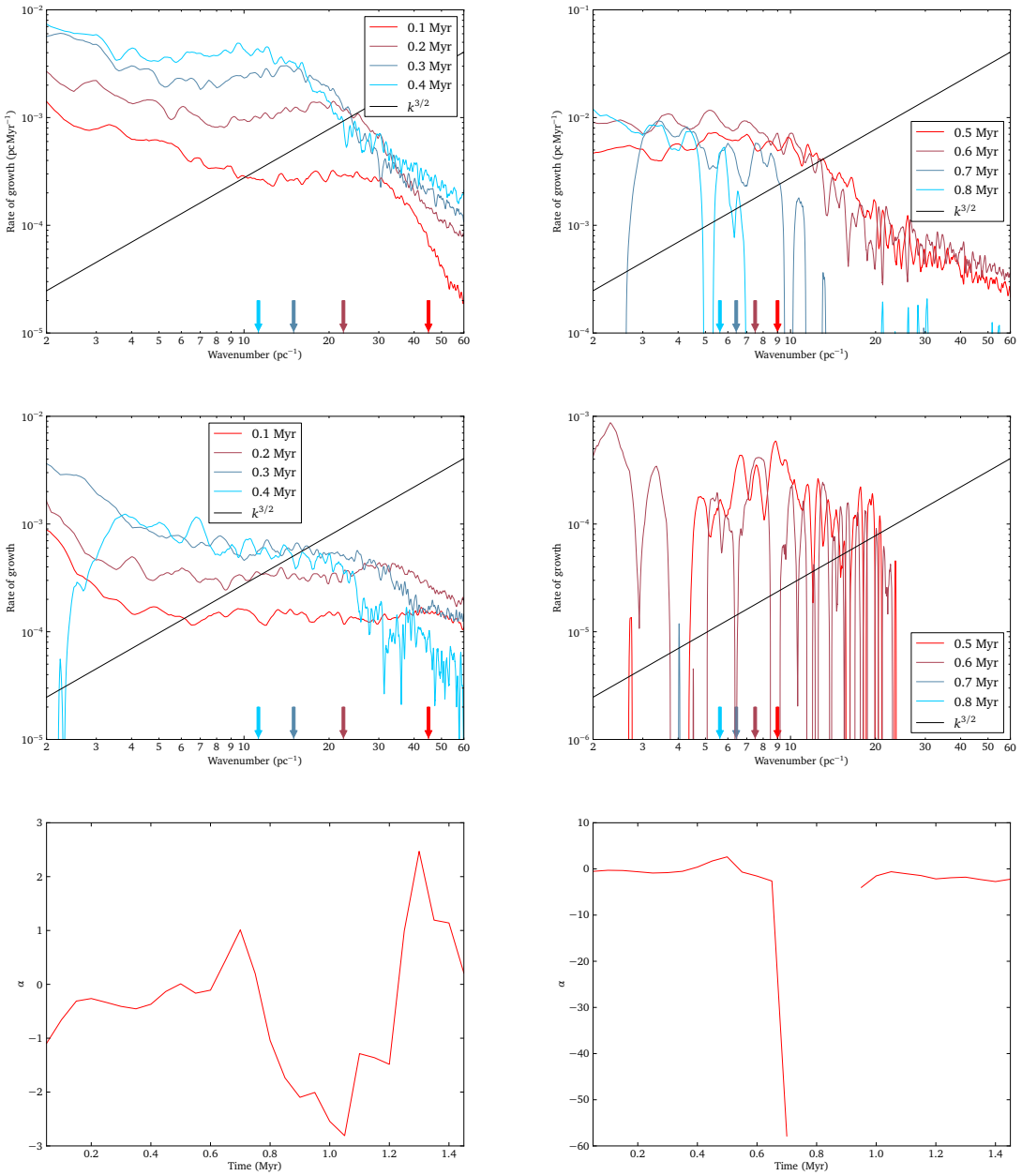


Figure 14.11 – The rates of growth of centre-of-mass position for collision VI as a function of wavenumber. Each line represents results from a different timestep. The upper plots are for rates of growth of centre-of-mass position; the centre plot shows rates of growth of column density. The gradient of the solid black line indicates the relation predicted by Vishniac (1994). The arrows indicate the maximum resolvable wavenumber for the correspondingly coloured timestep. The lower left plot and lower right plot show the index of power law fits to the rates of growth of centre-of-mass position and column density respectively.

Collision	Time of first sink formation (Myr)	Time of 2% star formation efficiency (Myr)
I	1.74	1.90
II	1.48	1.61
III	1.35	1.50
IV	1.44	1.53
V	1.48	No data
VI	—	—

Table 14.3 – The times of first sink formation, and the time where 2% of the original clouds has been accreted onto sinks.

accreted onto the sinks. This is the time of 2% star formation efficiency (SFE). Although collision V also produces sinks, we did not have the computational resources to follow this collision beyond the early stages. Collision VI does not produce any sinks during the original collision, although it will form sinks if the expanding post-collision layer is allowed to collapse.

Section 8.3 describes resolution requirements for self-gravitating SPH simulations. As described in that section, we would fully resolve the collapse of a dense, spherical core of one solar mass. However we might not resolve the collapse of smaller cores. The star formation properties we obtain from our simulations should be treated with caution.

We are interested in differences between the stars formed in low-velocity collisions, where the gravitational instability is dominant, and stars formed in higher-velocity collisions, where the NTSI is dominant. Even though our simulations may not be resolved for low-mass star formation, we can still make comparative studies.

The most obvious difference between low-velocity and high-velocity collisions is that the highest-velocity collision does not form stars at all, while the lower-velocity collisions form a large number of stars. In collision VI the NTSI completely disrupts the layer, allowing material to penetrate the layer and hindering the formation of dense gas in the shocked layer. A bloated, turbulent layer is formed which does not undergo gravitational collapse on the timescale of the collision.

14.5.1 Time of first sink formation

Table 14.3 shows the times of first sink formation and the times of 2% star formation efficiency for each collision. The time of first sink formation is higher at both low and high collision velocities, reaching a minimum for collision III. As the collision velocity increases, the density of the layer should increase, decreases the time to the first gravitational collapse. However, at higher collision velocities the NTSI appears and disrupts the layer, delaying gravitational fragmentation. Since collision VI takes approximately only 0.8 Myr, no star formation can occur for this collision. The time taken to reach 2% star formation

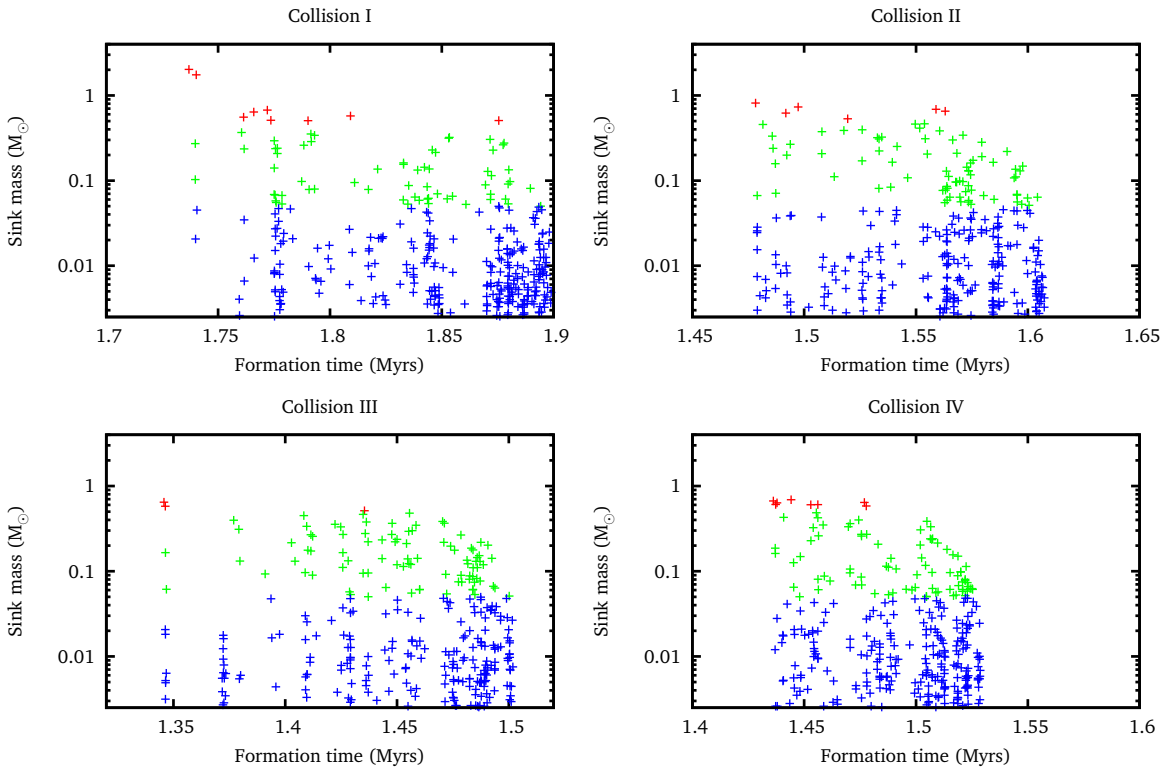


Figure 14.12 – Sink formation time and mass at the time of 2% SFE. Sinks with masses greater than 0.5 $M_\odot$ are indicated by red points. Sinks with masses between 0.05 $M_\odot$ and 0.5 $M_\odot$ are indicated with green points. All other sinks are indicated with blue points.

efficiency appears to follow a similar trend for the first four collisions.

14.5.2 Formation time and final sink mass and speed

Figure 14.12 shows the sink mass at the time of 2% SFE with respect to the original formation time of that sink. Sinks further to the left of the figure are therefore older. We note that at all collision velocities there are a large number of very low-mass sinks. These are expected as sinks form at approximately 0.0025 $M_\odot$, corresponding to 50 SPH particles, and have to accrete material in order to become more massive.

Examining the results for collision I shows that the most massive stars form first in an early burst of star formation that produces only six sinks. There is then a pause before another burst of star formation occurs. Careful examination of the simulations suggests that this first generation of stars forms as a result of the pattern of fragmentation in the layer. For collision II the first generation of stars is not as well separated in from later star formation, although there are still bursts of star formation at later times.

Fragmentation produces dense filaments, and the first generation of stars forms at the intersection of these filaments. As these stars do not have many competitors, they are able to rapidly accrete mass and so two of the first generation stars become the most massive stars formed. This can be seen in 14.3 for collision II, where the first stars form

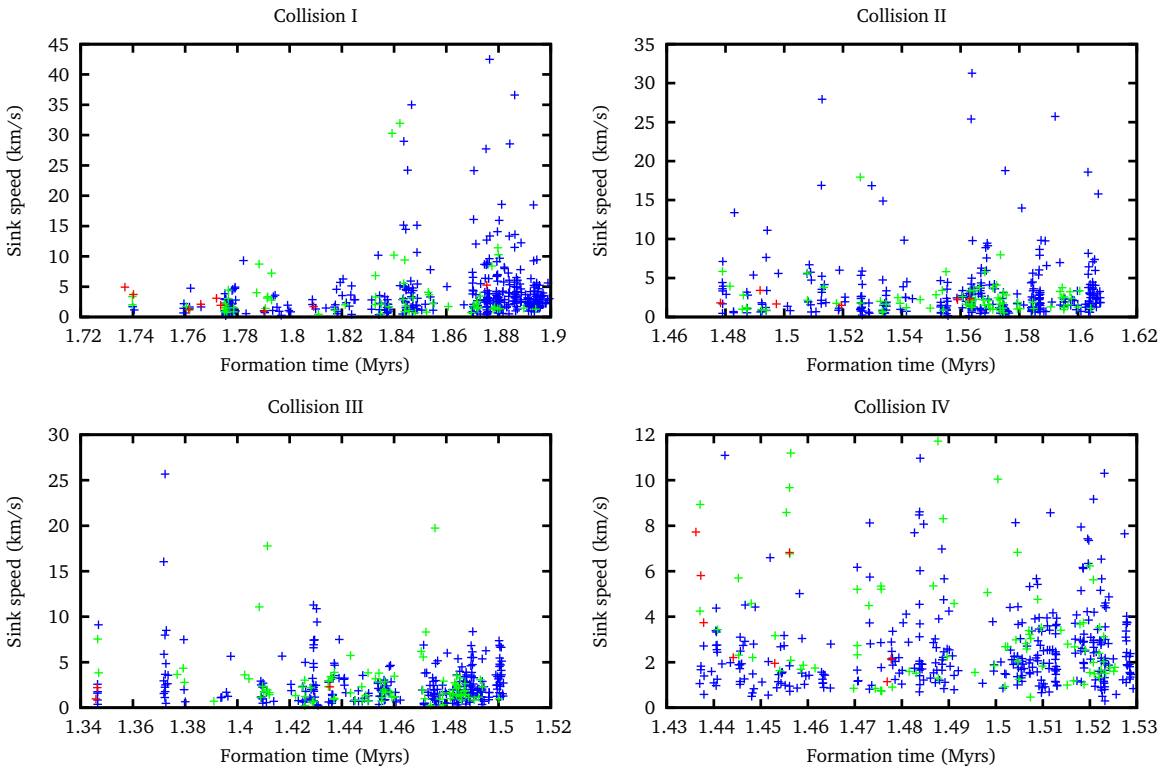


Figure 14.13 – Sink formation time and speed at the time of 2% SFE. Sinks with masses greater than 0.5 $M_\odot$ are indicated by red points. Sinks with masses between 0.05 $M_\odot$ and 0.5 $M_\odot$ are indicated with green points. All other sinks are indicated with blue points.

in dense regions such as filaments, and later stars form around them and are ejected.

Later stars tend to form around or otherwise in association with other stars; as material falls towards existing stars it fragments and unstable multiple systems are formed. This is shown in the data as a burst of star formation; the production of sinks is clearly episodic.

For collision III, we again see a first generation of stars, some of which are the most massive stars produced. A clearly defined second generation is also seen. collision IV is more similar to collision II, with a more continual production of stars.

In all cases the most massive stars tend to form early in the simulation. These massive stars have had the most time to accrete matter and thus increase their mass. Although there are some differences between the collisions, there is not a clear trend with respect to collision velocity; any differences may be due to the chaotic nature of the systems.

Figure 14.13 shows the sink speed at the time of 2% SFE with respect to the original formation time of that sink. Sinks are colour-coded by mass to indicate trends of speed with respect to mass.

The maximum speed reached by sinks decreases with increasing collision velocity. Ten sinks in collision I reach over 20 $\mathrm{km\,s^{-1}}$, with one reaching over 40 $\mathrm{km\,s^{-1}}$, while no sink in collision IV reaches over 12 $\mathrm{km\,s^{-1}}$.

This may be due to the density of star formation. In the lower-velocity collision,

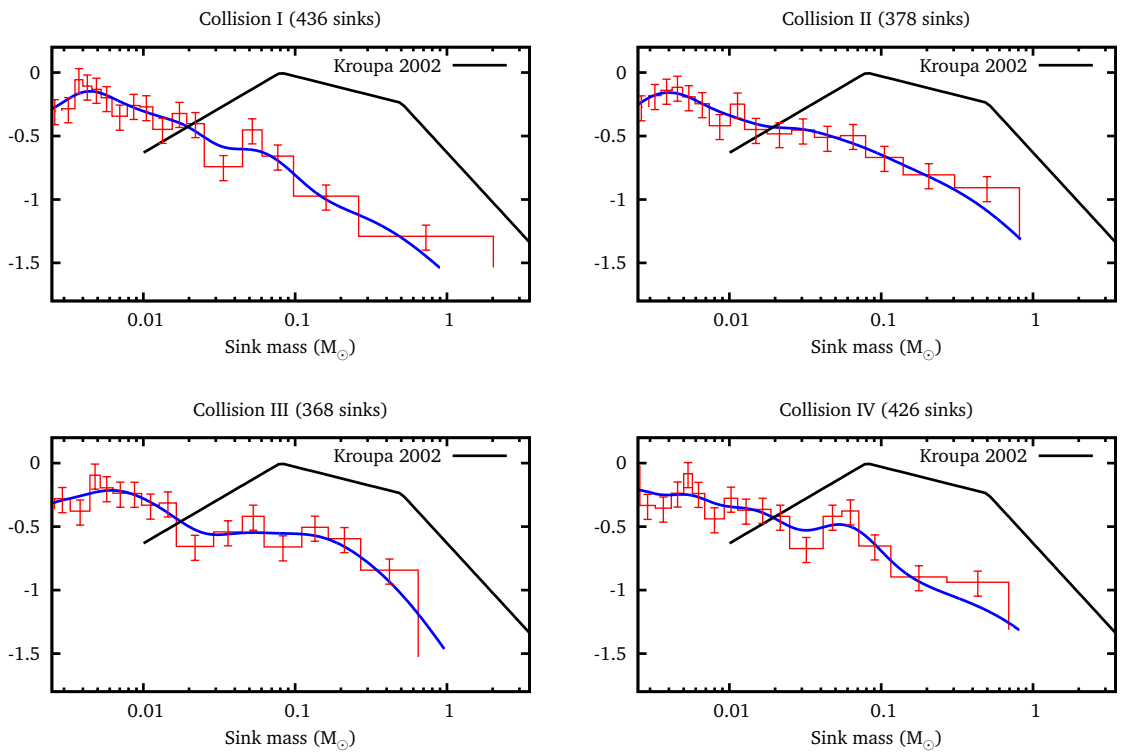


Figure 14.14 – Sink mass function at the time of 2% SFE. The stellar IMF found by Kroupa (2002) is shown for comparison. This figure is produced as described in Section 5.4.1.

only the most central regions of the dense layer undergo gravitational fragmentation. In higher-velocity collisions the pattern of star formation is more distributed, with multiple subclusters. This can be seen in Figure 14.2. At the time of 2% SFE, the low-velocity collisions have a small dense cluster which from which stars are ejected. The higher density may allow more violent interactions, while the more distributed star formation of higher-velocity collisions mitigates the strength of stellar interactions.

We also note that the most massive stars tend to have low speeds. We do not attempt to follow the details of stellar interactions, but we assume it is easier to eject a low-mass star from a cluster at a high velocity than to eject a high-mass star. We therefore expect that for low-velocity collisions, the resulting star cluster may be denser, and there may be ejections at greater velocity, while for high-velocity collisions, the star cluster may be more distributed, may have subclusters, and will not have as many high-velocity ejections.

14.5.3 Sink mass function

Figure 14.14 shows the sink mass function from our simulations. All the mass functions for our collisions are similar, indicating that the collision velocity does not have a significant effect on the masses of stars produced. Our mass function peaks at a much smaller mass than the Kroupa 2002 stellar IMF; however, our sinks are still accreting. The mass function can be expected to shift to higher masses once the production of sinks has stopped.

14.6 Conclusions

We have simulated the head-on collision of identical uniform-density gas clouds at a range of supersonic velocities. The collision creates a dense shock-confined layer, which depending on collision velocity may fragment and form stars. At low collision velocities, the layer is of lower density and is relatively unperturbed by hydrodynamic instabilities. At higher collision velocities, the layer is of higher density and is strongly perturbed by the non-linear thin shell instability. All simulations form stars except for the highest-velocity collision.

The layer grows more rapidly than expected for an idealized isothermal layer. We find that the layer grows slightly faster at higher collision velocities. Iwasaki & Tsuribe (2008) predicted two timescales for the gravitational instability in the layer: the time at which the gravitational instability becomes dominant, and the time at which the gravitational instability becomes non-linear. These timescales decrease with increasing collision velocity. We find that the time to the formation of the first sink does decrease with increasing collision velocity for the three lowest-velocity collisions, but increases for the higher-velocity collisions.

This is due to the increasing influence of the NTSI, which inhibits the formation of stars by disrupting the layer and preventing gravitational fragmentation. For the lower-velocity collisions, we find the non-linear timescales of Iwasaki & Tsuribe (2008) are an overestimate but are correct to within a factor of two. Iwasaki & Tsuribe (2008) also predict the wavelength of gravitational fragmentation, but we do not examine this further as the predicted wavelengths are quite large compared to the size of our box.

Examination of the rates of growth shows the result of the continuous input of turbulence as material is accreted to the layer. For low-velocity collisions, the rates of growth of both centre-of-mass position and column density decrease with increasing wavenumber, as the turbulence is preferentially excited on large wavenumbers. As the collision velocity increases, there is a rise in the rates of growth of column density at intermediate wavenumbers. At higher collision velocities this increase at intermediate wavenumbers is also seen in the rates of growth of centre-of-mass position. These increases can be compared with the results obtained in Chapter 12, and shows the presence of the NTSI at higher collision velocities.

Low-velocity collisions only trigger fragmentation in the parts of the layer closest to the centre, as these have the highest density. At 2% SFE, the result is a single dense cluster. Higher-velocity collisions have a more distributed pattern of star formation, with subclusters spread over a larger fraction of layer. The highest-velocity collision does not form stars; instead the layer become bloated and turbulent as it is strongly excited by the NTSI.

In some simulations a first generation of stars is formed. These stars are formed in the

densest regions of the layer along filaments and where filaments meet. This may also be followed by a second generation of stars, before more continuous star formation occurs. These later stars tend to form around existing stars as these stars attract material towards them.

As low-velocity collisions produce a denser cluster, ejections can be more violent, and higher sink speeds are observed than are seen for the higher-velocity collisions. This is the only significant difference in stellar properties between stars formed in low-velocity and high-velocity collisions. The sink mass functions are similar at all collision velocities. Although the sink mass functions are peaked at a much lower mass than the stellar IMF, our sinks are still accreting and are still being produced. Once sink formation has slowed, the peak mass will increase.

Chapter 15

Conclusions

In Section 1.1 we have shown the potential importance of cloud–cloud collisions to star formation in the Galaxy. Section 1.2 reviews some of the observational evidence for cloud–cloud collisions within our galaxy. Sections 1.3 and 7.1 review previous numerical work relating to cloud–cloud collisions, some of which has shown that instabilities play a significant role in such collisions. We have therefore performed a series of simulations to examine the effects of instabilities in cloud–cloud collisions.

In Chapter 2 we describe the method, smoothed particle hydrodynamics, that we have used in our simulations. Chapter 3 describes the algorithms used in our numerical code, and Chapter 4 describes the parallelization techniques we have used or implemented. Chapter 5 describes the analysis techniques we have used to obtain our results from our simulation data, and Chapter 8 describes the practicalities of creating initial conditions for our simulations.

In Chapter 7 we review the non-linear thin shell instability. The NTSI is a bending-mode instability that occurs in a shock-confined layer. It is caused by shear between ram pressure and thermal pressure. A cloud–cloud collision produces a shock-confined layer which may be susceptible to the NTSI. Vishniac (1994) provides a theoretical description of the NTSI, together with a time-independent growth rate as a function of wavenumber, and a minimum and maximum wavenumber susceptible to the NTSI. He considered only monochromatic one-dimensional perturbations. We test these predictions, and empirically extend them to more general scenarios, in our first series of simulations.

In Chapter 6 we review gravitational instability, focussing on the instability in a thin dense layer, such as that formed in a supersonic cloud–cloud collision. We review estimates of the timescales of gravitational collapse and the conditions for such a layer to gravitationally fragment.

15.1 Simulations of the NTSI

To simulate the growth of the NTSI, we perform purely hydrodynamic simulations as described in Chapter 9. We impose a number of different initial velocity perturbations on

the simulations, but use a single supersonic collision velocity.

The layers in our simulations grow faster than predicted for an idealized isothermal shock. The rate at which the layer grows appears to be independent of the type of initial velocity perturbations used.

We obtain the rates of growth of our simulations for two quantities: the centre-of-mass position of the layer and the column density. As a bending-mode instability, the NTSI will create changes in the centre-of-mass position of the layer. The NTSI tends to move material to the peaks of the bending modes, and so the NTSI also creates changes in the column density. These changes can be used to measure the rates of growth of the NTSI or other instabilities as a function of time and wavenumber.

When using monochromatic perturbations we recover the qualitative result predicted by Vishniac (1994) that larger wavenumbers grow more rapidly. We do not recover the predicted growth rate for our simulations with one-dimensional perturbations, probably due to the large amplitude of our initial perturbations rapidly saturating the instability. However, we do recover the predicted growth rate for simulations with two-dimensional perturbations.

We also find that only some wavenumbers are susceptible to the NTSI. Vishniac (1994) predicts a minimum wavenumber, but we excite the NTSI at much lower wavenumbers. The maximum wavenumber is predicted to be set by the thickness of the layer; here our simulations confirm the predicted limit.

We then perform simulations with white noise initial velocity perturbations. For one-dimensional white noise perturbations, we again recover the predicted relation of Vishniac (1994). The two-dimensional white noise perturbations produce less clear results, but agree qualitatively that higher wavenumbers grow more rapidly up to the maximum wavenumber set by the width of the layer.

Finally we use either subsonic or supersonic turbulence as an initial velocity perturbation. Our turbulence is strongest at small wavenumbers, and the turbulence has a corresponding effect on the rates of growth for these simulations. Unlike for the NTSI, rates of growth decrease with increasing wavenumber. For supersonic turbulence, the strong turbulence prevents the NTSI from becoming significant, while the NTSI is able to coexist with the turbulence if the turbulence is only subsonic.

For the simulations with subsonic turbulence, the rates of growth as a function of wavenumber therefore have several components. At all wavenumbers there is a background level with a negative gradient corresponding to the turbulence, and at higher wavenumbers the NTSI grows above this, creating a range of wavenumbers where the gradient is positive. Wavenumbers above the maximum resolvable wavenumber, set by the layer width, decrease as expected.

These simulations extend the work of Vishniac (1994) to provide some empirical

estimates of the behaviour of the NTSI in these more complex situations. Such situations are more likely than the collision of perfectly smooth unperturbed clouds or clouds with a one-dimensional perturbation.

15.2 Simulations of cloud–cloud collision

A realistic collision may be subject to both initial turbulence, the NTSI and gravitational instability; we therefore conduct a series of more realistic simulations. These include self-gravity, and model the supersonic collision of a pair of identical uniform-density spherical clouds. We are limited by computational resources, and therefore only explore the effect of varying collision velocity.

We collide the spheres at six different supersonic velocities. At low velocities the NTSI does not appear; there is no evidence for increased rates of growth at higher wavenumbers. Gravitational instability combined with the initial turbulent perturbations creates a pattern of filaments in the central regions of the layer. A first generation of stars form in the densest parts of these filaments and at their intersections. Further stars then form in the dense gas surrounding these stars, forming a dense cluster from which low-mass stars are ejected.

At higher velocities there is increased rates of growth at higher wavenumbers as the NTSI appears, first in the rates of growth of column densities and then in the rates of growth of centre-of-mass position. These increased rates of growth are very similar to those identified in the purely hydrodynamic studies of the NTSI with initial turbulence, demonstrating the significance of the NTSI at higher collision velocities.

The pattern of fragmentation becomes clumpier as the NTSI appears at these higher velocities. As more of the layer become gravitationally unstable, star formation is more distributed, forming subclusters. Stars are ejected from these subclusters but typically at lower speeds than for the lower-velocity collisions.

Iwasaki & Tsuribe (2008) estimate timescales for the gravitational instability. For the low-velocity collisions, these estimates are correct to within a factor of two, and the timescale to gravitational instability decreases with increasing collision velocity as expected. This trend reverses at higher collision velocity due to the influence of the NTSI. The NTSI disrupts the layer, preventing the formation and gravitational collapse of dense regions. Although the NTSI also produces dense regions at the peaks of bending modes, we find the primary effect of the NTSI is to suppress star formation. At very high collision velocities, a combination of the NTSI and the limited time available for collapse prevents the clouds from forming stars. Instead, the dense layer becomes bloated and turbulent, allowing efficient mixing between the clouds.

Although the collisions look quite different at low and high collision velocities, the properties of the resulting stars are quite similar. We produce very similar sink mass

functions for each collision velocity. Our sink mass functions peak at a much lower mass than typically expected; however, our stars have not stopped accreting at the end of our simulations. As stars accrete more material, the peak of the mass function will tend to move to higher masses.

15.3 Summary

We have conducted hydrodynamic simulations of the NTSI and partially confirmed the predicted results of Vishniac (1994). We have empirically explored a wider range of initial perturbations than considered by Vishniac (1994). We have conducted simulations of supersonic cloud–cloud collision, varying only the collision velocity. We have compared these simulations to our hydrodynamic simulations of the NTSI to determine the effect of collision velocity on the growth of the NTSI and the resulting star formation. We have shown that the NTSI appears at higher collision velocities and suppresses star formation at very high velocities, while gravitational instability dominates at lower collision velocities. Finally we show that the properties of the stars themselves do not depend significantly on the collision velocity.

References

Abramowski A. *et al.* 2011; **Revisiting the Westerlund 2 field with the HESS telescope array**; *A&A* **525**, A46

Alves J., Lada C. & Lada E. 2001a; **Internal structure of a cold dark molecular cloud inferred from the extinction of background starlight**; *Nature* **409**, 159

Alves J., Lada C. & Lada E. 2001b; **Seeing the light through the dark**; *The Messenger* **103**, 1

Arthur S. 2008; **The Evolution of the Circumstellar and Interstellar Medium Around Massive Stars**; in *IAU Symposium*; ed. by F. Bresolin, P. A. Crowther, & J. Puls; vol. **250**, 355

Ascenso J., Alves J., Beletsky Y. & Lago M. 2007; **Near-IR imaging of Galactic massive clusters: Westerlund 2**; *A&A* **466**, 137

Attwood R., Goodwin S. & Whitworth A. 2007; **Adaptive smoothing lengths in SPH**; *A&A* **464**, 447

Bally J., Stark A., Wilson R. & Henkel C. 1987; **Galactic center molecular clouds. I – Spatial and spatial-velocity maps**; *ApJS* **65**, 13

Bally J., Stark A., Wilson R. & Henkel C. 1988; **Galactic center molecular clouds. II – Distribution and kinematics**; *ApJ* **324**, 223

Barnes J. & Hut P. 1986; **A hierarchical O(N log N) force-calculation algorithm**; *Nature* **324**, 446

Bate M., Bonnell I. & Bromm V. 2002; **The formation of close binary systems by dynamical interactions and orbital decay**; *MNRAS* **336**, 705

Bate M., Bonnell I. & Bromm V. 2003; **The formation of a star cluster: predicting the properties of stars and brown dwarfs**; *MNRAS* **339**, 577

Bate M., Bonnell I. & Price N. 1995; **Modelling accretion in protobinary systems**; *MNRAS* **277**, 362

Bate M. & Burkert A. 1997; **Resolution requirements for smoothed particle hydrodynamics calculations with self-gravity**; *MNRAS* **288**, 1060

Bell K. & Lin D. 1994; **Using FU Orionis outbursts to constrain self-regulated protostellar disk models**; *ApJ* **427**, 987

Benson J. & Johnston K. 1984; **Arc second resolution maps of the compact sources in Sagittarius B2 and G34.3+0.2**; *ApJ* **277**, 181

Benz W., Cameron A., Press W. & Bowers R. 1990; **Dynamic mass exchange in doubly degenerate binaries. I – 0.9 and 1.2 solar mass stars**; *ApJ* **348**, 647

Bhattal A., Francis N., Watkins S. & Whitworth A. 1998; **Dynamically triggered star formation in giant molecular clouds**; *MNRAS* **297**, 435

Blondin J. & Marks B. 1996; **Evolution of cold shock-bounded slabs**; *New Astronomy* **1**, 235

Bok B. & Reilly E. 1947; **Small Dark Nebulae**; *ApJ* **105**, 255

Bonnor W. 1956; **Boyle's Law and gravitational instability**; *MNRAS* **116**, 351

Brand J. & Blitz L. 1993; **The Velocity Field of the Outer Galaxy**; *A&A* **275**, 67

Bronfman L., Casassus S., May J. & Nyman L.-Å. 2000; **The radial distribution of OB star formation in the Galaxy**; *A&A* **358**, 521

Bronfman L., Cohen R., Alvarez H., May J. & Thaddeus P. 1988; **A CO survey of the southern Milky Way - The mean radial distribution of molecular clouds within the solar circle**; *ApJ* **324**, 248

Burkert A. & Alves J. 2009; **The Inevitable Future of the Starless Core Barnard 68**; *ApJ* **695**, 1308

Burton W., Gordon M., Bania T. & Lockman F. 1975; **The overall distribution of carbon monoxide in the plane of the Galaxy**; *ApJ* **202**, 30

Byleveld S. & Pongracic H. 1996; **The influence of magnetic fields on star formation**; *Publications of the Astronomical Society of Australia* **13**, 71

Cambrésy L., Rho J., Marshall D. & Reach W. 2011; **Variation of the extinction law in the Trifid nebula**; *A&A* **527**, A141

Carraro G. & Munari U. 2004; **A multicolour CCD photometric study of the open clusters NGC 2866, Pismis 19, Westerlund 2, ESO96-SC04, NGC 5617 and NGC 6204**; *MNRAS* **347**, 625

Caswell J. & Haynes R. 1987; **Southern HII regions – an extensive study of radio recombination line emission**; *A&A* **171**, 261

Cernicharo J., Lefloch B., Cox P., Cesarsky D., Esteban C., Yusef-Zadeh F., Mendez D., Acosta-Pulido J., Garcia Lopez R. & Heras A. 1998; **Induced Massive Star Formation in the Trifid Nebula?**; *Science* **282**, 462

Chapman S., Pongracic H., Disney M., Nelson A., Turner J. & Whitworth A. 1992; **The formation of binary and multiple star systems**; *Nature* **359**, 207

Clemens D., Sanders D. & Scoville N. 1988; **The large-scale distribution of molecular gas in the first Galactic quadrant**; *ApJ* **327**, 139

Colella P. & Woodward P. 1984; **The Piecewise Parabolic Method (PPM) for Gas-Dynamical Simulations**; *Journal of Computational Physics* **54**, 174

Combes F. 1991; **Distribution of CO in the Milky Way**; *Annual Reviews Astronomy and Astrophysics* **29**, 195

Dame T. 2007; **On the Distance and Molecular Environment of Westerlund 2 and HESS J1023-575**; *ApJL* **665**, L163

Dame T., Hartmann D. & Thaddeus P. 2001; **The Milky Way in Molecular Clouds: A New Complete CO Survey**; *ApJ* **547**, 792

De Vicente P., Martin-Pintado J. & Wilson T. 1997; **A hot ring in the Sagittarius B2 molecular cloud**; *A&A* **320**, 957

Dickel J., Dickel H. & Wilson W. 1978; **The detailed structure of CO in molecular cloud complexes. II – The W75-DR 21 region**; *ApJ* **223**, 840

Dowell C. 1997; **Far-Infrared Polarization by Absorption in the Molecular Cloud Sagittarius B2**; *ApJ* **487**, 237

Dowell C., Hildebrand R., Schleuning D., Vaillancourt J., Dotson J., Novak G., Renbarger T. & Houde M. 1998; **Submillimeter Array Polarimetry with Hertz**; *ApJ* **504**, 588

Duarte-Cabral A., Dobbs C., Peretto N. & Fuller G. 2011; **Was a cloud-cloud collision the trigger of the recent star formation in Serpens?**; *A&A* **528**, A50

Eckart C. 1960; **Variation Principles of Hydrodynamics**; *Physics of Fluids* **3**, 421

Elmegreen B. 1998; **Observations and Theory of Dynamical Triggers for Star Formation**; in *Origins*; ed. by C. Woodward, J. Shull & J. H. Thronson; vol. **148**, 150

Frigo M. & Johnson S. 2005; **The Design and Implementation of FFTW3**; *Proceedings of the IEEE* **93**.2; Special issue on "Program Generation, Optimization, and Platform Adaptation", 216

Fryxell B., Olson K., Ricker P., Timmes F., Zingale M., Lamb D., MacNeice P., Rosner R., Truran J. & Tufo H. 2000; **FLASH: An Adaptive Mesh Hydrodynamics Code for Modeling Astrophysical Thermonuclear Flashes**; *ApJS* **131**, 273

Fukui Y., Furukawa N., Dame T., Dawson J., Yamamoto H., Rowell G., Aharonian F., Hofmann W., Wilhelmi E. d. O., Minamidani T., Kawamura A., Mizuno N., Onishi T., Mizuno A. & Nagataki S. 2009; **A Peculiar Jet and Arc of Molecular Gas toward the Rich and Young Stellar Cluster Westerlund 2 and a TeV Gamma Ray Source**; *PASJ* **61**, L23+

Furukawa N., Dawson J., Ohama A., Kawamura A., Mizuno N., Onishi T. & Fukui Y. 2009; **Molecular Clouds Toward RCW49 and Westerlund 2: Evidence for Cluster Formation Triggered by Cloud–Cloud Collision**; *ApJL* **696**, L115

Gardner F., Whiteoak J., Forster J. & Pankonin V. 1986; **Characteristics of the H2CO maser emission in SGR B2**; *MNRAS* **218**, 385

Gaume R. & Claussen M. 1990; **The Sagittarius B2 star-forming region – Subarcsecond radio spectral line and continuum observations**; *ApJ* **351**, 538

Gaume R., Claussen M., de Pree C., Goss W. & Mehringer D. 1995; **The Sagittarius B2 Star-forming Region. I. Sensitive 1.3 Centimeter Continuum Observations**; *ApJ* **449**, 663

Gaume R. & Mutel R. 1987; **A study of the ground-state hydroxyl maser emission associated with 11 regions of star formation**; *ApJS* **65**, 193

Gilden D. 1982; **Dynamical evolution of substucture in molecular clouds**; PhD thesis; Texas Univ., Austin.

Gilden D. 1984; **Clump collisions in molecular clouds – Gravitational instability and coalescence**; *ApJ* **279**, 335

Gingold R. & Monaghan J. 1977; **Smoothed particle hydrodynamics – Theory and application to non-spherical stars**; *MNRAS* **181**, 375

Goodwin S., Whitworth A. & Ward-Thompson D. 2004; **Simulating star formation in molecular cloud cores. I. The influence of low levels of turbulence on fragmentation and multiplicity**; *A&A* **414**, 633

Grabelsky D., Cohen R., Bronfman L. & Thaddeus P. 1988; **Molecular clouds in the Carina arm – The largest objects, associated regions of star formation, and the Carina arm in the Galaxy**; *ApJ* **331**, 181

Greaves J. & White G. 1991; **The influence of shocks on star formation in the OMC1 Ridge**; *A&A* **248**, L27

Green A., Cram L., Large M. & Ye T. 1999; **The Molonglo Galactic Plane Survey. I. Overview and Images**; *ApJS* **122**, 207

Güsten R. 1989; **Gas and Dust in the Inner Few Degrees of the Galaxy (review)**; in *The Center of the Galaxy*; ed. by M. Morris; vol. **136**, 89

Habe A. & Ohta K. 1992; **Gravitational instability induced by a cloud-cloud collision – The case of head-on collisions between clouds with different sizes and densities**; *PASJ* **44**, 203

Hasegawa T., Arai T., Yamaguchi N. & Sato F. 2008; **ASTE observations of the massive-star forming region Sgr B2: a giant impact scenario**; *Astrophysics and Space Science* **313**, 91

Hasegawa T., Morita K., Okumura S., Kaifu N., Suzuki H., Ohishi M., Hayashi M. & Ukita N. 1986; **New SIO Masers in Star Forming Regions W:51 IRS:2 and SAGITTARIUS-B2 MD:5**; in *Masers, Molecules, and Mass Outflows in Star Formation Regions*; ed. by A. D. Haschick & J. M. Moran, 275

Hasegawa T., Sato F., Whiteoak J. & Miyawaki R. 1994; **A large-scale cloud collision in the galactic center molecular cloud near Sagittarius B21**; *ApJL* **429**, L77

Hausman M. 1979; **Hydrodynamic calculations of interstellar cloud collisions**; PhD thesis; Princeton Univ., NJ.

Hausman M. 1981; **Collisional mergers and fragmentation of interstellar clouds**; *ApJ* **245**, 72

Hayes J., Norman M., Fiedler R., Bordner J., Li P., Clark S., ud-Doula A. & Mac Low M. 2006; **Simulating Radiating and Magnetized Flows in Multiple Dimensions with ZEUS-MP**; *ApJS* **165**, 188

Heitsch F., Hartmann L. & Burkert A. 2008; **Fragmentation of Shocked Flows: Gravity, Turbulence, and Cooling**; *ApJ* **683**, 786

Hennebelle P., Banerjee R., Vázquez-Semadeni E., Klessen R. & Audit E. 2008; **From the warm magnetized atomic medium to molecular clouds**; *A&A* **486**, L43

Higuchi A., Kurono Y., Saito M. & Kawabe R. 2010; **A Mapping Survey of Dense Clumps Associated with Embedded Clusters. II. Can Clump-Clump Collisions Induce Stellar Clusters?**; *ApJ* **719**, 1813

Houghton S. & Whiteoak J. 1995; **The small-scale distribution of emission from the 6.7-GHz transition of methanol in SGR B2**; *MNRAS* **273**, 1033

Hubber D., Batty C., McLeod A. & Whitworth A. 2011; **SEREN – a new SPH code for star and planet formation simulations. Algorithms and tests**; *A&A* **529**, A27

Hubber D., Goodwin S. & Whitworth A. 2006; **Resolution requirements for simulating gravitational fragmentation using SPH**; *A&A* **450**, 881

Hueckstaedt R. 2003; **Nonlinear thin shell instabilities in molecular clouds**; *New Astronomy* **8**, 295

Hunter J. J., Sandford I. M., Whitaker R. & Klein R. 1986; **Star formation in colliding gas flows**; *ApJ* **305**, 309

Hüttemeister S., Wilson T., Mauersberger R., Lemme C., Dahmen G. & Henkel C. 1995; **A multilevel study of ammonia in star-forming regions. 6: The envelope of Sagittarius B2**; *A&A* **294**, 667

Iwasaki K. & Tsuribe T. 2008; **Gravitational Instability of Shocked Interstellar Gas Layers**; *PASJ* **60**, 125

Jeans J. 1902; **The Stability of a Spherical Nebula**; *Royal Society of London Philosophical Transactions Series A* **199**, 1

Kahn F. 1954; **The acceleration of interstellar clouds**; *B.A.I.N.* **12**, 187

Kashiwa B., Padial N., Rauenzahn R. & Vanderheyden W. 1994; **A cell-centered ICE method for multiphase flow simulations**; in *Numerical Methods in Multiphase Flows*; ed. by C. Crowe & J. Katz; New-York: ASME

Kaspi V., Lyne A., Manchester R., Johnston S., D'Amico N. & Shemar S. 1993; **A young, glitching pulsar near the direction of W28**; *ApJL* **409**, L57

Kenney J. & Lord S. 1991; **Orbit crowding of molecular gas at a bar-spiral arm transition zone in M83**; *ApJ* **381**, 118

Kiessling M. K.-H. 2003; **The "Jeans swindle": A true story – mathematically speaking**; *Adv. Appl. Math.* **31**, 132

Kimura T. & Tosa M. 1996; **Collision of clumpy molecular clouds**; *A&A* **308**, 979

Kitsionas S. 2000; **Particle Splitting: A New Method for SPH Star Formation Simulations**; PhD thesis; School of Physics & Astronomy, Cardiff University, The Parade, Cardiff, CF24 3AA, UK

Kitsionas S. & Whitworth A. 2007; **High-resolution simulations of clump-clump collisions using SPH with particle splitting**; *MNRAS* **378**, 507

Klein R. & Woods D. 1998; **Bending Mode Instabilities and Fragmentation in Interstellar Cloud Collisions: A Mechanism for Complex Structure**; *ApJ* **497**, 777

Kohoutek L., Mayer P. & Lorenz R. 1999; **Photometry and spectroscopy of the central star of the Trifid nebula**; *A&AS* **134**, 129

Koo B.-C., Lee Y., Fuller G., Lee M., Kwon S.-M. & Jung J.-H. 1994; **A multiwavelength study of IRAS 19550+3248: A protostar possibly formed by a cloud-cloud collision**; *ApJ* **429**, 233

Kroupa P. 2002; **The Initial Mass Function of Stars: Evidence for Uniformity in Variable Systems**; *Science* **295**, 82

Kuan Y.-J. & Snyder L. 1996; **Three Millimeter Molecular Line Observations of Sagittarius B2. II. High-Resolution Studies of $^{1}8$ CO, HNCO, NH_2 CHO, and HCOOCH_3** ; *ApJ* **470**, 981

Kwan J. & Valdes F. 1983; **Spiral gravitational potentials and the mass growth of molecular clouds**; *ApJ* **271**, 604

Kwan J. & Valdes F. 1987; **The spatial and mass distributions of molecular clouds and spiral structure**; *ApJ* **315**, 92

Larson R. 1978; **A finite-particle scheme for three-dimensional gas dynamics**; *Journal of Computational Physics* **27**, 397

Larson R. 1981; **Turbulence and star formation in molecular clouds**; *MNRAS* **194**, 809

Lattanzio J., Monaghan J., Pongracic H. & Schwarz M. 1985; **Interstellar Cloud Collisions**; *MNRAS* **215**, 125

Lattanzio J., Monaghan J., Pongracic H. & Schwarz M. 1986; **Controlling Penetration**; *SIAM Journal on Scientific and Statistical Computing* **7**.2, 591

Lefloch B. & Cernicharo J. 2000; **Pre-Orion Cores in the Trifid Nebula**; *ApJ* **545**, 340

Lefloch B., Cernicharo J. & Pardo J. 2008; **Star formation in the Trifid Nebula. Cores and filaments**; *A&A* **489**, 157

Lepine J. & Duvert G. 1994; **Star formation by infall of high velocity clouds on the galactic disk**; *A&A* **286**, 60

Lis D. & Goldsmith P. 1989; **CO isotope studies and mass of the Sagittarius B2 molecular cloud**; *ApJ* **337**, 704

Lis D. & Goldsmith P. 1990; **Modeling of the continuum and molecular line emission from the Sagittarius B2 molecular cloud**; *ApJ* **356**, 195

Lis D., Goldsmith P., Carlstrom J. & Scoville N. 1993; **Millimeter-wavelength aperture synthesis observations of massive star-forming regions in Sagittarius B2**; *ApJ* **402**, 238

Liszt H. 1992; **Radio images of Sagittarius. I – Overview, Sagittarius D and Sagittarius E**; *ApJS* **82**, 495

Liszt H., Burton W. & Xiang D.-L. 1984; **CO-13 in the inner galactic plane**; *A&A* **140**, 303

Looney L., Wang S., Hamidouche M., Safier P. & Klein R. 2006; **Colliding Clouds: The Star Formation Trigger of the Stellar Cluster around BD +40 4124**; *ApJ* **642**, 330

Loren R. 1976; **Colliding clouds and star formation in NGC 1333**; *ApJ* **209**, 466

Loren R. 1977; **The star-formation process in molecular clouds associated with Herbig Be/Ae stars. I - LK H-alpha 198, BD + 40 deg 4124, and NGC 7129**; *ApJ* **218**, 716

Lucy L. 1977; **A numerical approach to the testing of the fission hypothesis**; *AJ* **82**, 1013

Lynds B., Canzian B. & O'Neil J. E. 1985; **Optical measurements of the Trifid dust**; *ApJ* **288**, 164

Mac Low M., Klessen R., Burkert A. & Smith M. 1998; **Kinetic Energy Decay Rates of Supersonic and Super-Alfvénic Turbulence in Star-Forming Clouds**; *Physical Review Letters* **80**, 2754

Marinho E. & Lépine J. 2000; **SPH simulations of clumps formation by dissipative collision of molecular clouds. I. Non magnetic case**; *A&AS* **142**, 165

Martin-Pintado J., Bachiller R. & Fuente A. 1992; **SiO Emission as a Tracer of Shocked Gas in Molecular Outflows**; *A&A* **254**, 315

Martin-Pintado J., de Vicente P., Fuente A. & Planesas P. 1997; **SiO Emission from the Galactic Center Molecular Clouds**; *ApJL* **482**, L45+

Masunaga H. & Inutsuka S. 2000; **A Radiation Hydrodynamic Model for Protostellar Collapse. II. The Second Collapse and the Birth of a Protostar**; *ApJ* **531**, 350

McMillan P. & Binney J. 2010; **The uncertainty in Galactic parameters**; *MNRAS* **402**, 934

Mehringer D. 1995; **BIMA Array CS J = 2 1 Observations of Sagittarius B2**; *ApJ* **454**, 782

Mehringer D., Goss W. & Palmer P. 1994; **New formaldehyde masers in Sagittarius B2**; *ApJ* **434**, 237

Mehringer D. & Menten K. 1997; **44 GHz Methanol Masers and Quasi-thermal Emission in Sagittarius B2**; *ApJ* **474**, 346

Mehringer D., Palmer P., Goss W. & Yusef-Zadeh F. 1993; **Radio continuum and radio recombination line observations of Sagittarius B2**; *ApJ* **412**, 684

Mikami H., Umemoto T., Yamamoto S. & Saito S. 1992; **Detection of SiO emission in the L1157 dark cloud**; *ApJL* **392**, L87

Minh Y., Haikala L., Hjalmarson A. & Irvine W. 1998; **Are Clouds Collapsing at the 2 North Position of Sagittarius B2?**; *ApJ* **498**, 261

Minier V., Ellingsen S., Norris R. & Booth R. 2003; **The protostellar mass limit for 6.7 GHz methanol masers. I. A low-mass YSO survey**; *A&A* **403**, 1095

Miyazaki A. & Tsuboi M. 2000; **Dense Molecular Clouds in the Galactic Center Region. II. Statistical Properties of the Galactic Center Molecular Clouds**; *ApJ* **536**, 357

Moffat A., Shara M. & Potter M. 1991; **New Wolf–Rayet stars in Galactic open clusters – Sher 1 and the giant HII region core Westerlund 2**; *AJ* **102**, 642

Monaghan J. 1992; **Smoothed particle hydrodynamics**; *Annual Reviews Astronomy and Astrophysics* **30**, 543

Monaghan J. 1997; **SPH and Riemann Solvers**; *Journal of Computational Physics* **136**, 298

Monaghan J. 2002; **SPH compressible turbulence**; *MNRAS* **335**, 843

Monaghan J. & Lattanzio J. 1985; **A refined particle method for astrophysical problems**; *A&A* **149**, 135

Morgan W., Whitford A. & Code A. 1953; **Studies in Galactic Structure. I. a Preliminary Determination of the Space Distribution of the Blue Giants**; *ApJ* **118**, 318

Morita K.-I., Hasegawa T., Ukita N., Okumura S. & Ishiguro M. 1992; **Accurate positions of SiO masers in active star-forming regions – Orion-KL, W51-IRS2, and Sagittarius-B2 MD5**; *PASJ* **44**, 373

Morris J. & Monaghan J. 1997; **A Switch to Reduce SPH Viscosity**; *Journal of Computational Physics* **136**, 41

Nagasawa M. & Miyama S. 1987; **Three-Dimensional Numerical Simulation of Interstellar Cloud-Cloud Collisions and Triggered Star Formation. I —Head-On Collisions—**; *Progress of Theoretical Physics* **78**, 1250

Nelson R. & Papaloizou J. 1994; **Variable Smoothing Lengths and Energy Conservation in Smoothed Particle Hydrodynamics**; *MNRAS* **270**, 1

Neufeld D., Lepp S. & Melnick G. 1995; **Thermal Balance in Dense Molecular Clouds: Radiative Cooling Rates and Emission-Line Luminosities**; *ApJS* **100**, 132

Novak G., Dotson J., Dowell C., Goldsmith P., Hildebrand R., Platt S. & Schleuning D. 1997; **Polarized Far-Infrared Emission from the Core and Envelope of the Sagittarius B2 Molecular Cloud**; *ApJ* **487**, 320

Odenwald S., Campbell M., Shivanandan K., Schwartz P., Fazio G. & Moseley H. 1990; **Multiwavelength observations of two B-star nurseries – DR 15 and DR 20**; *AJ* **99**, 288

Odenwald S., Fischer J., Lockman F. & Stemwedel S. 1992; **The unusual cometary star-forming region G110-13**; *ApJ* **397**, 174

Ogura K. & Ishida K. 1975; **UBV photometry of the stars in the fields of emission nebulae. I. M20**; *PASJ* **27**, 119

Ohama A., Dawson J., Furukawa N., Kawamura A., Moribe N., Yamamoto H., Okuda T., Mizuno N., Onishi T., Maezawa H., Minamidani T., Mizuno A. & Fukui Y. 2010; **Temperature and Density Distribution in the Molecular Gas Toward Westerlund 2: Further Evidence for Physical Association**; *ApJ* **709**, 975

Oort J. & Spitzer J. L. 1955; **Acceleration of Interstellar Clouds by O-Type Stars**; *ApJ* **121**, 6

O'Shea B., Bryan G., Bordner J., Norman M., Abel T., Harkness R. & Kritsuk A. 2004; **Introducing Enzo, an AMR Cosmology Application**; *ArXiv Astrophysics e-prints*; arXiv:astro-ph/0403044

Piatti A., Bica E. & Clariá J. 1998; **Fundamental parameters of the highly reddened young open clusters Westerlund 1 and 2**; *A&AS* **127**, 423

Pierce-Price D., Richer J., Greaves J., Holland W., Jenness T., Lasenby A., White G., Matthews H., Ward-Thompson D., Dent W., Zylka R., Mezger P., Hasegawa T., Oka T., Omont A. & Gilmore G. 2000; **A Deep Submillimeter Survey of the Galactic Center**; *ApJL* **545**, L121

Pongracic H., Chapman S., Davies J., Disney M., Nelson A. & Whitworth A. 1992; **Computer simulations of the formation of massive protostars**; *MNRAS* **256**, 291

Porter D., Pouquet A. & Woodward P. 1992; **Three-dimensional supersonic homogeneous turbulence – A numerical study**; *Physical Review Letters* **68**, 3156

Price D. 2004; **Magnetic fields in Astrophysics**; PhD thesis; Institute of Astronomy, Madingley Rd, Cambridge, CB2 0HA, UK

Price D. 2007; **SPLASH: An Interactive Visualisation Tool for Smoothed Particle Hydrodynamics Simulations**; *Publications of the Astronomical Society of Australia* **24**, 159

Price D. 2008; **Modelling discontinuities and Kelvin Helmholtz instabilities in SPH**; *Journal of Computational Physics* **227**, 10040

Price D. 2012; **Smoothed particle hydrodynamics and magnetohydrodynamics**; *Journal of Computational Physics* **231**, 759

Price D. & Monaghan J. 2007; **An energy-conserving formalism for adaptive gravitational force softening in smoothed particle hydrodynamics and N-body codes**; *MNRAS* **374**, 1347

Rauw G., Crowther P., de Becker M., Gosset E., Nazé Y., Sana H., van der Hucht K., Vreux J.-M. & Williams P. 2005; **The spectrum of the very massive binary system WR 20a (WN6ha + WN6ha): Fundamental parameters and wind interactions**; *A&A* **432**, 985

Rauw G., Manfroid J., Gosset E., Nazé Y., Sana H., de Becker M., Foellmi C. & Moffat A. 2007; **Early-type stars in the core of the young open cluster Westerlund 2**; *A&A* **463**, 981

Redman M., Keto E. & Rawlings J. 2006; **Oscillations in the stable starless core Barnard 68**; *MNRAS* **370**, L1

Rho J., Corcoran M., Chu Y.-H. & Reach W. 2001; **X-Rays and Protostars in the Trifid Nebula**; *ApJ* **562**, 446

Rho J., Reach W., Lefloch B. & Fazio G. 2006; **Spectacular Spitzer Images of the Trifid Nebula: Protostars in a Young, Massive-Star-forming Region**; *ApJ* **643**, 965

Roberts J. W. & Stewart G. 1987; **The role of orbital dynamics and cloud-cloud collisions in the formation of giant molecular clouds in global spiral structures**; *ApJ* **314**, 10

Sabano Y. & Tosa M. 1985; **Star formation induced by supersonic turbulence in interstellar molecular cloud**; *Astrophysics and Space Science* **115**, 85

Sato F., Hasegawa T., Whiteoak J. & Miyawaki R. 2000; **Cloud Collision-induced Star Formation in Sagittarius B2. I. Large-Scale Kinematics**; *ApJ* **535**, 857

Scoville N. & Solomon P. 1975; **Molecular clouds in the Galaxy**; *ApJL* **199**, L105

Scoville N., Solomon P. & Penzias A. 1975; **The molecular cloud Sagittarius B2**; *ApJ* **201**, 352

Serabyn E., Güsten R. & Schulz A. 1993; **Fragmentation and kinematics of the W49N cloud core**; *ApJ* **413**, 571

Shara M., Smith L., Potter M. & Moffat A. 1991; **A deep survey for Galactic Wolf–Rayet stars. I – Motivation, search technique, and first results**; *AJ* **102**, 716

Shiki S., Ohishi M. & Deguchi S. 1997; **SiO Maser Sources toward the Sagittarius B2 Molecular Cloud**; *ApJ* **478**, 206

Smith J. 1980; **Cloud–cloud collisions in the interstellar medium**; *ApJ* **238**, 842

Sodroski T., Odegard N., Dwek E., Hauser M., Franz B., Freedman I., Kelsall T., Wall W., Berriman G., Odenwald S., Bennett C., Reach W. & Weiland J. 1995; **The Ratio of H 2 Column Density to 12CO Intensity in the Vicinity of the Galactic Center**; *ApJ* **452**, 262

Solomon P., Rivolo A., Barrett J. & Yahil A. 1987; **Mass, luminosity, and line width relations of Galactic molecular clouds**; *ApJ* **319**, 730

Springel V. 2005; **The cosmological simulation code GADGET-2**; *MNRAS* **364**, 1105

Springel V. & Hernquist L. 2002; **Cosmological smoothed particle hydrodynamics simulations: the entropy equation**; *MNRAS* **333**, 649

Springel V., Yoshida N. & White S. 2001; **GADGET: a code for collisionless and gasdynamical cosmological simulations**; *New Astronomy* **6**, 79

Stamatellos D., Whitworth A., Bisbas T. & Goodwin S. 2007; **Radiative transfer and the energy equation in SPH simulations of star formation**; *A&A* **475**, 37

Stark A. & Brand J. 1989; **Kinematics of molecular clouds. II - New data on nearby giant molecular clouds**; *ApJ* **339**, 763

Stark A. & Lee Y. 2005; **The Scale Height of Giant Molecular Clouds Is Less than That of Smaller Clouds**; *ApJL* **619**, L159

Stone M. 1970a; **Collisions Between HI Clouds. II. Two-Dimensional Model**; *ApJ* **159**, 293

Stone M. 1970b; **Collisions Between HI Clouds. I. One-Dimensional Model**; *ApJ* **159**, 277

Stone R. 1978; **Mean secular parallax at low galactic latitude**; *AJ* **83**, 393

Tasker E. & Tan J. 2009; **Star Formation in Disk Galaxies. I. Formation and Evolution of Giant Molecular Clouds via Gravitational Instability and Cloud Collisions**; *ApJ* **700**, 358

Teyssier R. 2002; **Cosmological hydrodynamics with adaptive mesh refinement. A new high resolution code called RAMSES**; *A&A* **385**, 337

Thomas P. & Couchman H. 1992; **Simulating the formation of a cluster of galaxies**; *MNRAS* **257**, 11

Torii K., Enokiya R., Sano H., Yoshiike S., Hanaoka N., Ohama A., Furukawa N., Dawson J., Moribe N., Oishi K., Nakashima Y., Okuda T., Yamamoto H., Kawamura A., Mizuno N., Maezawa H., Onishi T., Mizuno A. & Fukui Y. 2011; **Molecular Clouds in the Trifid Nebula M20: Possible Evidence for a Cloud–Cloud Collision in Triggering the Formation of the First Generation Stars**; *ApJ* **738**, A46

Tsuboi M., Handa T. & Ukita N. 1999; **Dense Molecular Clouds in the Galactic Center Region. I. Observations and Data**; *ApJS* **120**, 1

Turner J., Chapman S., Bhattal A., Disney M., Pongracic H. & Whitworth A. 1995; **Binary star formation: gravitational fragmentation followed by capture**; *MNRAS* **277**, 705

Vallee J. 1995; **IRAS 2306+1451: Evidence for a Possible Cloud-Cloud Collision**; *AJ* **110**, 2256

Vallee J. & Avery L. 1990; **New CO observations of the high-latitude molecular cloud MBM 55, near IRAS 230604 + 145055**; *A&A* **233**, 553

Van der Hucht K. A. 2001; **The VIIth catalogue of galactic Wolf–Rayet stars**; *New Astronomy Reviews* **45**, 135

Vázquez-Semadeni E., Banerjee R., Klessen R. & Ballesteros-Paredes J. 2007a; **Molecular Cloud Formation III. Influence of the Magnetic Field**; in *Bulletin of the American Astronomical Society*; vol. **38**, 984

Vázquez-Semadeni E., Gómez G., Jappsen A., Ballesteros-Paredes J., González R. & Klessen R. 2007b; **Molecular Cloud Evolution. II. From Cloud Formation to the Early Stages of Star Formation in Decaying Conditions**; *ApJ* **657**, 870

Vázquez-Semadeni E., Ryu D., Passot T., González R. & Gazol A. 2006; **Molecular Cloud Evolution. I. Molecular Cloud and Thin Cold Neutral Medium Sheet Formation**; *ApJ* **643**, 245

Vishniac E. 1983; **The dynamic and gravitational instabilities of spherical shocks**; *ApJ* **274**, 152

Vishniac E. 1994; **Nonlinear instabilities in shock-bounded slabs**; *ApJ* **428**, 186

Vogel S., Genzel R. & Palmer P. 1987; **The dispersal of dense protostellar material – NH3 hot cores and outflows in Sagittarius B2**; *ApJ* **316**, 243

Wadsley J., Stadel J. & Quinn T. 2004; **Gasoline: a flexible, parallel implementation of TreeSPH**; *New Astronomy* **9**, 137

Wang J., Chen W., Miller M., Qin S. & Wu Y. 2004; **Massive Star Formation Triggered by Collision between Galactic and Accreted Intergalactic Clouds**; *ApJL* **614**, L105

Westerlund B. 1961; **A Heavily Reddened Cluster in ARA**; *PASP* **73**, 51

Wetzstein M., Nelson A., Naab T. & Burkert A. 2009; **Vine – A Numerical Code for Simulating Astrophysical Systems Using Particles. I. Description of the Physics and the Numerical Methods**; *ApJS* **184**, 298

Whitney B. *et al.* 2004; **A GLIMPSE of Star Formation in the Giant HII Region RCW 49**; *ApJS* **154**, 315

Whitworth A. 1998; **The Jeans instability in smoothed particle hydrodynamics**; *MNRAS* **296**, 442

Whitworth A., Bhattal A., Chapman S., Disney M. & Turner J. 1994; **Fragmentation of shocked interstellar gas layers**; *A&A* **290**, 421

Whitworth A., Bhattal A., Turner J. & Watkins S. 1995a; **Estimating density in smoothed particle hydrodynamics**; *A&A* **301**, 929

Whitworth A., Chapman S., Bhattal A., Disney M., Pongracic H. & Turner J. 1995b; **Binary star formation: accretion-induced rotational fragmentation**; *MNRAS* **277**, 727

Williams J. & McKee C. 1997; **The Galactic Distribution of OB Associations in Molecular Clouds**; *ApJ* **476**, 166

Wilson T., Mezger P., Gardner F. & Milne D. 1970; **A Survey of H 109 α Recombination Line Emission in Galactic HII Regions of the Southern Sky**; *A&A* **6**, 364

Womack M., Ziurys L. & Sage L. 1993; **N_2 H^+ in the Orion ambient ridge – Cloud clumping versus rotation**; *ApJL* **406**, L29

Ziurys L., Friberg P. & Irvine W. 1989; **Interstellar SiO as a tracer of high-temperature chemistry**; *ApJ* **343**, 201

List of Permissions

Fig. 1.1 Reprinted from the Astrophysical Journal Letters, Vol. 665, Aug 2007, Dame, T. M., On the Distance and Molecular Environment of Westerlund 2 and HESS J1023-575, pp L163–L166, Copyright (2007), reproduced by permission of the AAS 6

Fig. 1.2 Reprinted from the Astrophysical Journal Letters, Vol. 696, May 2009, Furukawa, N., Dawson, J. R., Ohama, A., Kawamura, A., Mizuno, N., Onishi, T. & Fukui, Y., Molecular Clouds Toward RCW49 and Westerlund 2: Evidence for Cluster Formation Triggered by Cloud–Cloud Collision, pp L115–L119, Copyright (2009), reproduced by permission of the AAS 7

Fig. 1.3 Reprinted from the Astrophysical Journal Letters, Vol. 696, May 2009, Furukawa, N., Dawson, J. R., Ohama, A., Kawamura, A., Mizuno, N., Onishi, T. & Fukui, Y., Molecular Clouds Toward RCW49 and Westerlund 2: Evidence for Cluster Formation Triggered by Cloud–Cloud Collision, pp L115–L119, Copyright (2009), reproduced by permission of the AAS 8

Fig. 1.4 Reprinted from the Astrophysical Journal, Vol. 709, February 2010, Ohama, A., Dawson, J. R., Furukawa, N., Kawamura, A., Moribe, N., Yamamoto, H., Okuda, T., Mizuno, N., Onishi, T., Maezawa, H., Minamidani, T., Mizuno, A., and Fukui, Y., Temperature and Density Distribution in the Molecular Gas Toward Westerlund 2: Further Evidence for Physical Association, pp 975–982, Copyright (2010), reproduced by permission of the AAS 9

Fig. 1.5 Reprinted from the Astrophysical Journal, Vol. 709, February 2010, Ohama, A., Dawson, J. R., Furukawa, N., Kawamura, A., Moribe, N., Yamamoto, H., Okuda, T., Mizuno, N., Onishi, T., Maezawa, H., Minamidani, T., Mizuno, A., and Fukui, Y., Temperature and Density Distribution in the Molecular Gas Toward Westerlund 2: Further Evidence for Physical Association, pp 975–982, Copyright (2010), reproduced by permission of the AAS 10

Fig. 1.6 Reprinted from the Astrophysical Journal, Vol. 643, June 2006, Rho, J., Reach, W., Lefloch, B., Fazio, G., Spectacular Spitzer Images of the Trifid Nebula: Protostars in a Young, Massive-Star-forming Region, pp 965–977, Copyright (2006), reproduced by permission of the AAS 14

Fig. 1.7 Reprinted from the Astrophysical Journal, Vol. 738, September 2011, Torii, K., Enokiya, R., Sano, H., Yoshiike, S., Hanaoka, N., Ohama, A., Furukawa, N., Dawson, J. R., Moribe, N., Oishi, K., Nakashima, Y., Okuda, T., Yamamoto, H., Kawamura, A., Mizuno, N., Maezawa, H., Onishi, T., Mizuno, A., Fukui, Y., Molecular Clouds in the Trifid Nebula M20: Possible Evidence for a Cloud–Cloud Collision in Triggering the Formation of the First Generation Stars, Article 46, Copyright (2011), reproduced by permission of the AAS 15

Fig. 1.8 Reprinted from the Astrophysical Journal, Vol. 738, September 2011, Torii, K., Enokiya, R., Sano, H., Yoshiike, S., Hanaoka, N., Ohama, A., Furukawa, N., Dawson, J. R., Moribe, N., Oishi, K., Nakashima, Y., Okuda, T., Yamamoto, H., Kawamura, A., Mizuno, N., Maezawa, H., Onishi, T., Mizuno, A., Fukui, Y., Molecular Clouds in the Trifid Nebula M20: Possible Evidence for a Cloud–Cloud Collision in Triggering the Formation of the First Generation Stars, Article 46, Copyright (2011) 16

Fig. 1.9 Reprinted from the Astrophysical Journal, Vol. 738, September 2011, Torii, K., Enokiya, R., Sano, H., Yoshiike, S., Hanaoka, N., Ohama, A., Furukawa, N., Dawson, J. R., Moribe, N., Oishi, K., Nakashima, Y., Okuda, T., Yamamoto, H., Kawamura, A., Mizuno, N., Maezawa, H., Onishi, T., Mizuno, A., Fukui, Y., Molecular Clouds in the Trifid Nebula M20: Possible Evidence for a Cloud–Cloud Collision in Triggering the Formation of the First Generation Stars, Article 46, Copyright (2011), reproduced by permission of the AAS 17

Fig. 1.10 Reprinted from the Astrophysical Journal, Vol. 738, September 2011, Torii, K., Enokiya, R., Sano, H., Yoshiike, S., Hanaoka, N., Ohama, A., Furukawa, N., Dawson, J. R., Moribe, N., Oishi, K., Nakashima, Y., Okuda, T., Yamamoto, H., Kawamura, A., Mizuno, N., Maezawa, H., Onishi, T., Mizuno, A., Fukui, Y., Molecular Clouds in the Trifid Nebula M20: Possible Evidence for a Cloud–Cloud Collision in Triggering the Formation of the First Generation Stars, Article 46, Copyright (2011), reproduced by permission of the AAS 18

Fig. 1.11 Reprinted from the Astrophysical Journal, Vol. 738, September 2011, Torii, K., Enokiya, R., Sano, H., Yoshiike, S., Hanaoka, N., Ohama, A., Furukawa, N., Dawson, J. R., Moribe, N., Oishi, K., Nakashima, Y., Okuda, T., Yamamoto, H., Kawamura, A., Mizuno, N., Maezawa, H., Onishi, T., Mizuno, A., Fukui, Y., Molecular Clouds in the Trifid Nebula M20: Possible Evidence for a Cloud–Cloud Collision in Triggering the Formation of the First Generation Stars, Article 46, Copyright (2011), reproduced by permission of the AAS 19

Fig. 1.12 Reprinted from Astrophysics and Space Sciences, Vol. 313, Hasegawa, T., Arai, T., Yamaguchi, N., Sato, F. and the ASTE team, ASTE observations of the massive-star forming region Sgr B2: a giant impact scenario, pp 91–94, Copyright (2008), with permission from Springer Science+Business Media 22

Fig. 1.13 Reprinted from the Astrophysical Journal, Vol. 535, June 2000, Sato, F., Hasegawa, T., Whiteoak, J., Miyawaki, R., Cloud Collision-induced Star Formation in Sagittarius B2. I. Large-Scale Kinematics, pp 857–868, Copyright (2000), reproduced by permission of the AAS 24

Fig. 1.14 Reprinted with permission from the Publications of the Astronomical Society of Japan, Vol. 44, June 1992, Habe, A., Ohta, K., Gravitational Instability Induced by a Cloud–Cloud Collision: The Case of Head-on Collisions between Clouds with Different Sizes and Densities, pp 203–226, Copyright (1992) 25

Fig. 1.15 Reprinted from Astrophysics and Space Sciences, Vol. 313, Hasegawa, T., Arai, T., Yamaguchi, N., Sato, F. and the ASTE team, ASTE observations of the massive-star forming region Sgr B2: a giant impact scenario, pp 91–94, Copyright (2008), with permission from Springer Science+Business Media 26

Fig. 1.16 Reprinted from the Astrophysical Journal, Vol. 695, April 2009, Burkert, A., Alves, J., The inevitable future of the starless core Barnard 68, pp 1308–1314, Copyright (2009), reproduced by permission of the AAS 29

Fig. 1.17 Reprinted from the Astrophysical Journal, Vol. 695, April 2009, Burkert, A., Alves, J., The inevitable future of the starless core Barnard 68, pp 1308–1314, Copyright (2009), reproduced by permission of the AAS 30

Fig. 1.18 Reprinted from the Astrophysical Journal, Vol. 695, April 2009, Burkert, A., Alves, J., The inevitable future of the starless core Barnard 68, pp 1308–1314, Copyright (2009), reproduced by permission of the AAS 31

Fig. 6.3 Reprinted with permission from the Publications of the Astronomical Society of Japan, Vol. 60, Feb 2008, Kazunari Iwasaki & Toru Tsuribe, Gravitational Instability of Shocked Interstellar Gas Layers, pp 125–136, Copyright (2008) 115

Fig. 7.2 Reprinted from New Astronomy, Vol. 1, Issue 3, Nov 1996, John M. Blondin & Brian S. Marks, Evolution of cold shock-bounded slabs, pp 235–244, Copyright (1996), with permission from Elsevier. 120

Fig. 7.3 Reprinted from New Astronomy, Vol. 1, Issue 3, Nov 1996, John M. Blondin & Brian S. Marks, Evolution of cold shock-bounded slabs, pp 235–244, Copyright (1996), with permission from Elsevier. 120

Fig. 7.4 Reprinted from the Astrophysical Journal, Vol. 498, Apr 1998, Richard I. Klein & D. Tod Woods, Bending mode instabilities and fragmentation in interstellar cloud collisions: a mechanism for complex structure, pp 777–799, Copyright (1998), reproduced by permission of the AAS 121

Fig. 7.5 Reprinted from the Astrophysical Journal, Vol. 498, Apr 1998, Richard I. Klein & D. Tod Woods, Bending mode instabilities and fragmentation in interstellar cloud collisions: a mechanism for complex structure, pp 777–799, Copyright (1998), reproduced by permission of the AAS 121

Fig. 7.6 Reprinted from New Astronomy, Vol. 8, Issue 4, May 2003, R. M. Hueckstaedt, Nonlinear thin shell instabilities in molecular clouds, pp 295–311, Copyright (2003), with permission from Elsevier. 122

Fig. 7.7 Reprinted from New Astronomy, Vol. 8, Issue 4, May 2003, R. M. Hueckstaedt, Nonlinear thin shell instabilities in molecular clouds, pp 295–311, Copyright (2003), with permission from Elsevier. 123

Fig. 8.1 Reproduced with permission from Astronomy & Astrophysics, Vol. 475, D. Stamatellos, A. P. Whitworth, T. Bisbas & S. Goodwin, Radiative Transfer and the energy equation in SPH simulations of star formation, pp 37–49, Copyright (2007) ESO 131

Appendix A

Derivation of the Jeans length and mass

The Jeans mass, as described in Section 6.1, is the critical mass scale of the gravitational instability. The Jeans mass is derived from the Jeans length, which is the critical length scale of the gravitational instability. The analysis of the Jeans mass begins by considering a small density perturbation in an otherwise infinite, homogeneous uniform density gas. Perturbations larger than the Jeans length will collapse gravitationally, while smaller perturbations oscillate. This problem was first studied by Jeans (1902).

Section 6.1 describes a simple, but relatively inaccurate, estimate of the Jeans mass by comparing the crossing time of sound waves in the gas with the freefall time. In this appendix we first discuss *Jeans' swindle*, and then use three different methods to derive a Jeans length. First we derive a Jeans length from considering stable and unstable wavelengths of spherical waves, similar to Jeans' original analysis. Secondly we derive a Jeans length using the virial theorem. Finally we derive the Jeans length by considering the hydrodynamic forces acting on a spherically symmetric cloud of gas.

In all cases we assume that the gas remains isothermal. At low densities interstellar gas is optically thin and can radiate away excess heat, keeping the gas approximately isothermal. At higher densities, typically above $10 \times 10^{-12}\,\mathrm{g\,cm^{-3}}$, interstellar gas becomes optically thick and begins to rise in temperature. This heating can significantly slow the gravitational contraction.

A.1 Jeans' swindle

Historically the Jeans length has been derived by ignoring the gravitation of the infinite background medium and considering only the gravity of the density perturbation. This has been described as *Jeans' swindle*. The gravitational potential is not well-defined in the case of an infinite medium. However Kiessling (2003) shows that although the potential is not well-defined, the gravitational forces cancel to zero and give no force. Following the method of Kiessling, we show it is mathematically valid to consider only the self-gravity due to the perturbation.

We set the initial state of the background medium such that it has infinite extent,

constant density and pressure, and zero velocity. The potential Φ is not well-defined in this limit, and so we introduce the 'Einsteinian potential' defined by Kiessling (2003). This includes a cosmological constant κ^2, and reduces to the standard gravitational potential in the limit $\kappa \to 0$. The Einsteinian potential is defined by

$$\nabla^2\Psi - \kappa^2\Psi = 4\pi G\rho \; . \tag{A.1}$$

This is the inhomogeneous Helmholtz equation

$$\nabla^2 A + k^2 A = C \; , \tag{A.2}$$

and if ρ is locally well-defined everywhere has a solution

$$\Psi(x) = -G \int \frac{\mathrm{e}^{-\kappa|x-y|}}{|\boldsymbol{x}-\boldsymbol{y}|}\rho(\boldsymbol{y})\,\mathrm{d}^3 y \; , \tag{A.3}$$

where the integral is over all space. For a uniform density this converges to a constant value (Kiessling 2003) given by

$$\Psi_0 = -4\pi G\rho_0 \frac{1}{\kappa^2} \; . \tag{A.4}$$

We now add a small density perturbation $\sigma(\boldsymbol{x})$ such that

$$\rho(\boldsymbol{x}) = \rho_0 + \sigma(\boldsymbol{x}) \, , \tag{A.5}$$

where ρ_0 is the background density. This perturbation must always be greater than $-\rho_0$ so that the total density remains positive, and for simplicity we only consider perturbations that are smooth and decay rapidly to zero at infinity. We impose this perturbation without adding or removing mass, and so require

$$\int \sigma(\boldsymbol{x})\,\mathrm{d}^3 x = 0 \, . \tag{A.6}$$

Since equation A.3 is linear, we can define the Einsteinian potential in terms of the background value Ψ_0 and a component due to the perturbation $\psi(\boldsymbol{x})$, giving

$$\Psi(\boldsymbol{x}) = \Psi_0 + \psi(\boldsymbol{x}) \, . \tag{A.7}$$

The function $\psi(\boldsymbol{x})$ is given by

$$\psi(\boldsymbol{x}) = -G \int \frac{\mathrm{e}^{-\kappa|x-y|}}{|\boldsymbol{x}-\boldsymbol{y}|}\sigma(\boldsymbol{y})\,\mathrm{d}^3 y \; , \tag{A.8}$$

which is a solution of the inhomogeneous Helmholtz equation

$$\nabla^2\psi - \kappa^2\psi = 4\pi G\sigma \,. \tag{A.9}$$

This is similar to equation A.1, but considers only the perturbation density σ rather than the total density ρ. The force is given by

$$\nabla\Psi(\boldsymbol{x}) = \nabla\Psi_0 + \nabla\psi(\boldsymbol{x}) \,, \tag{A.10}$$

but as Ψ_0 is constant $\nabla\Psi_0$ is zero, giving

$$\nabla\Psi(\boldsymbol{x}) = \nabla\psi(\boldsymbol{x}) \,. \tag{A.11}$$

Taking the limit $\kappa \to 0$, we obtain

$$\lim_{\kappa\to 0} \nabla\psi(\boldsymbol{x}) = \nabla\phi(\boldsymbol{x}) \,, \tag{A.12}$$

and so

$$\phi(\boldsymbol{x}) = -G\int \frac{1}{|\boldsymbol{x}-\boldsymbol{y}|}\sigma(\boldsymbol{y})\,\mathrm{d}^3y \,. \tag{A.13}$$

This is a solution to the Poisson equation

$$\nabla^2\phi = 4\pi G\sigma \,, \tag{A.14}$$

and so we have shown that, in the case of an infinite homogeneous density with a small perturbation, it is valid to ignore the background density and potential in the Poisson equation. Instead we can consider only the density and the gravitational potential of perturbation density σ. In doing so we have validated the *Jeans' swindle* as shown by Kiessling (2003).

A.2 Wavelengths of stable and unstable oscillations

In this section, we consider the effect of small perturbations on the medium, and find the wavelengths corresponding to stable and unstable oscillations. The equations governing the evolution of a fluid are the continuity equation,

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\boldsymbol{u}) = 0 \,, \tag{A.15}$$

the Euler momentum equation,

$$\frac{\partial\boldsymbol{u}}{\partial t} + \boldsymbol{u}\cdot\nabla\boldsymbol{u} = -\frac{\nabla P}{\rho} - \nabla\Phi \,, \tag{A.16}$$

the Poisson equation,

$$\nabla^2\Phi = 4\pi G\rho \,, \tag{A.17}$$

and an equation of state,

$$P = c_s^2\rho \,. \tag{A.18}$$

First we replace the potential Φ with the Einsteinian potential Ψ defined in Section A.1. The continuity equation and the equation of state are not changed, while the equation A.1 replaces the Poisson equation with

$$\nabla^2\Psi - \kappa^2\Psi = 4\pi G\rho \,. \tag{A.19}$$

In the Euler momentum equation we simply replace the potential Φ with the Einsteinian potential Ψ, giving

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u}\cdot\nabla\boldsymbol{u} = -\frac{\nabla P}{\rho} - \nabla\Psi \,, \tag{A.20}$$

We now separate the total density $\rho(\boldsymbol{x},t)$, the total pressure $P(\boldsymbol{x},t)$, the total velocity $u(\boldsymbol{x},t)$ and the total Einsteinian potential $\Phi(\boldsymbol{x},t)$ into background and perturbation components. The density $\rho(\boldsymbol{x},t)$ is given by

$$\rho(\boldsymbol{x},t) = \rho_0 + \sigma(\boldsymbol{x},t) \,, \tag{A.21}$$

where ρ_0 is the constant background density and $\sigma(\boldsymbol{x},t)$ is the perturbation density, while the pressure (using the equation of state) is given by

$$P(\boldsymbol{x},t) = c_s^2\left(\rho_0 + \sigma(\boldsymbol{x},t)\right) \,. \tag{A.22}$$

Similarly the velocity is given by

$$\boldsymbol{u}(\boldsymbol{x},t) = \boldsymbol{u}_0 + \boldsymbol{v}(\boldsymbol{x},t) \,, \tag{A.23}$$

but since the system is initially at rest the background velocity $\boldsymbol{u}_0 = 0$, the total velocity $\boldsymbol{u}(\boldsymbol{x},t)$ is equal to the perturbation velocity $\boldsymbol{v}(\boldsymbol{x},t)$. Finally the Einsteinian potential is given by

$$\Psi(\boldsymbol{x},t) = \Psi_0 + \psi(\boldsymbol{x},t) \,, \tag{A.24}$$

where Ψ_0 is the background potential and $\psi(\boldsymbol{x},t)$ is the potential due to the perturbation.

Substituting these into A.15, the continuity equation becomes

$$\frac{\partial\rho_0}{\partial t} + \frac{\partial\sigma}{\partial t} + \nabla\cdot\left(\rho_0\boldsymbol{v}\right) + \nabla\cdot(\sigma\boldsymbol{v}) = 0 \,. \tag{A.25}$$

ρ_0 is a constant term and so its derivative is zero, finally giving

$$\frac{\partial \sigma}{\partial t} + \rho_0 \nabla \cdot \boldsymbol{v} + \nabla \cdot (\sigma \boldsymbol{v}) = 0 \,. \tag{A.26}$$

Substituting the background and perturbation components into equation A.20, the Euler motion equation becomes

$$\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v} \cdot \nabla \boldsymbol{v} = \frac{-1}{\rho_0 + \sigma} \nabla \left(c_s^2 \rho_0 + c_s^2 \sigma \right) - \nabla \left(\Psi_0 + \psi \right) \,. \tag{A.27}$$

Again ρ_0 and Ψ_0 are constants, and so their spatial derivatives are zero, giving

$$\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v} \cdot \nabla \boldsymbol{v} = \frac{-c_s^2}{\rho_0 + \sigma} \nabla \sigma - \nabla \psi \,. \tag{A.28}$$

We can now use the limit that the cosmological constant κ^2 tends to zero. The Poisson equation is then

$$\nabla^2 \Psi = 4\pi G \rho \,. \tag{A.29}$$

Substituting in the background and perturbation components gives

$$\nabla^2 \Psi_0 + \nabla^2 \psi = 4\pi G \rho_0 + 4\pi G \sigma \,. \tag{A.30}$$

In the absence of a perturbation the Poisson equation is $\nabla^2 \Psi_0 = 4\pi G \rho_0$, giving

$$\nabla^2 \psi = 4\pi G \rho \,. \tag{A.31}$$

To linearize the equations, as suggested by Kiessling (2003), we replace the perturbation variables with sums of the form

$$\sigma = \sigma_1 + \sigma_2 + \sigma_3 + \dots \,, \tag{A.32}$$

$$\boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2 + \boldsymbol{v}_3 + \dots \,, \tag{A.33}$$

and

$$\psi = \psi_1 + \psi_2 + \psi_3 + \dots \,. \tag{A.34}$$

We define the *total index* of a term as the sum of the indices of all the terms, such that $\rho_0 \sigma_2$ and $\sigma_1 \boldsymbol{v}_1$ both have an index of two. Substituting these variables into the continuity equation, and retaining only terms with a total index of one or less yields

$$\frac{\partial \sigma_1}{\partial t} + \rho_0 \nabla \cdot \boldsymbol{v}_1 = 0 \,. \tag{A.35}$$

Similarly the Euler momentum equation becomes

$$\frac{\partial \boldsymbol{v}_1}{\partial t} = \frac{-c_s^2}{\rho_0}\nabla\sigma_1 - \nabla\psi_1 \,, \tag{A.36}$$

and the Poisson equation becomes

$$\nabla^2\psi_1 = 4\pi G\sigma_1 \,. \tag{A.37}$$

Dropping the indices, as all remaining terms are of index one except ρ_0, we finally obtain our equations of motion

$$\frac{\partial \sigma}{\partial t} + \rho_0\nabla\cdot\boldsymbol{v} = 0 \,, \tag{A.38}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} = \frac{-c_s^2}{\rho_0}\nabla\sigma - \nabla\psi \,, \tag{A.39}$$

and

$$\nabla^2\psi = 4\pi G\sigma \,. \tag{A.40}$$

Taking the time derivative of equation A.38 gives

$$\frac{\partial^2\sigma}{\partial t^2} = -\rho_0\nabla\cdot\frac{\partial \boldsymbol{v}}{\partial t} \,. \tag{A.41}$$

We can then take the divergence of equation A.39 to obtain

$$\rho_0\nabla\cdot\frac{\partial \boldsymbol{v}}{\partial t} = -c_s^2\nabla^2\sigma - \rho_0\nabla^2\psi \,. \tag{A.42}$$

Using equations A.41 and A.40 we obtain the wave equation

$$\frac{\partial^2\sigma}{\partial t^2} - c_s^2\nabla^2\sigma = -4\pi G\rho_0\sigma \,. \tag{A.43}$$

We now introduce a spherical wave of the form

$$\delta\rho = \frac{A}{r}e^{i(\omega t - kr)} \,. \tag{A.44}$$

Substituting this into the first term of equation A.43, we find

$$\frac{\partial^2\sigma}{\partial t^2} = \omega^2\sigma \,. \tag{A.45}$$

The Laplacian ∇^2 in spherical coordinates for a spherically symmetric system is

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) . \tag{A.46}$$

The second term of equation A.43 is then

$$c_s^2\nabla^2\sigma = k^2\sigma . \tag{A.47}$$

Combining equations A.43, A.45 and A.47 we obtain the dispersion relation

$$\omega^2 = c_s^2k^2 - 4\pi G\rho_0 . \tag{A.48}$$

Positive values of ω^2 correspond to oscillating (stable) solutions, while negative values give imploding or exploding (unstable) solutions. The condition for collapse is therefore $\omega^2 \leq 0$. When $\omega^2 = 0$ we obtain the critical *Jeans wavenumber*

$$k_J = \frac{\sqrt{4\pi G\rho}}{c_s} . \tag{A.49}$$

Converting this to a wavelength by the relation $\lambda = \dfrac{2\pi}{k}$, we obtain the *Jeans wavelength*

$$\lambda_J = \sqrt{\frac{\pi}{G\rho}}\,c_s . \tag{A.50}$$

This wavelength corresponds to the maximum stable wavelength of a spherical wave, and therefore to the maximum stable radius of a spherical cloud. It is therefore related to the *Jeans radius* calculated in the following section.

A.3 Jeans length from the virial theorem

We now consider a spherical cloud of particles in static equilibrium with no internal degrees of freedom. The cloud has radius R, and within this radius all particles have the same temperature T and density ρ. The virial theorem tells us that for a collection of particles in equilibrium

$$\langle E_{KIN}\rangle = -\frac{1}{2}\langle E_{POT}\rangle , \tag{A.51}$$

where the angled brackets indicate the average of the kinetic and gravitational potential energies respectively. For a set of N particles an equivalent relation holds for the total kinetic and gravitational potential energies. The average kinetic energy of a particle with

only translational degrees of freedom is

$$\langle E_{\mathrm{KIN}} \rangle = \frac{3}{2} k_{\mathrm{B}} T \,, \tag{A.52}$$

where k_{B} is the Boltzmann constant. Vibrational modes are not excited at the typical temperature of cold molecular hydrogen. The total kinetic energy of particles is given by the average kinetic energy multiplied by the number of particles, and is given by

$$E_{\mathrm{KIN}} = \frac{4\pi R^3 \rho}{3\mu} \frac{3 k_{\mathrm{B}} T}{2} \,, \tag{A.53}$$

where μ is the average mass of each particle. Using the isothermal sound speed $c_{\mathrm{s}} = \sqrt{\dfrac{k_{\mathrm{B}} T}{\mu}}$, this simplifies to

$$E_{\mathrm{KIN}} = 2\pi\rho \, c_{\mathrm{s}}^2 R^3 \,. \tag{A.54}$$

The total gravitational potential energy is found by integrating the gravitational potential energy of a series of shells, each of mass $4\pi r^2 \rho \, \mathrm{d}r$, which feel gravity only from the mass within it. The gravitational potential of a shell is then

$$\Phi(r) = -\frac{G M(r) 4\pi r^2 \rho \, \mathrm{d}r}{r} \,, \tag{A.55}$$

where $M(r)$ is the mass within the shell, given by

$$M(r) = \frac{4\pi r^3 \rho}{3} \,. \tag{A.56}$$

The total gravitational potential energy is then

$$E_{\mathrm{POT}} = -\int G \frac{4\pi r^3 \rho}{3} \frac{4\pi r^2 \rho \, \mathrm{d}r}{r} \,. \tag{A.57}$$

This gives

$$E_{\mathrm{POT}} = -\frac{16}{15} \pi^2 G \rho^2 R^5 \,, \tag{A.58}$$

which is the standard result for a uniform density sphere. Combining equations A.51, A.54 and A.58 we obtain

$$2\pi\rho \, c_{\mathrm{s}}^2 R^3 = \frac{16}{30} \pi^2 G \rho^2 R^5 \,, \tag{A.59}$$

which rearranges to give the *Jeans radius*

$$R_{\mathrm{J}} = \sqrt{\frac{15}{4\pi G \rho}} c_{\mathrm{s}} \,. \tag{A.60}$$

This value of the Jeans radius is only 1.09 times larger than the Jeans wavelength derived in Section A.2. Clouds with a larger radius cannot support themselves against their own gravitational potential and must collapse.

A.4 Jeans length from hydrodynamic forces

We now consider the balance of forces on a hydrostatically supported spherically symmetric distribution of gas. We assume the temperature is constant throughout the sphere, and that there is a distribution of density and pressure forces supporting the cloud against gravitational collapse.

The pressure force acting on a thin shell of thickness $\mathrm{d}r$ is given by the pressure gradient across the shell, giving

$$F_{\mathrm{P}} = -4\pi r^2 \frac{\mathrm{d}P}{\mathrm{d}r}\,\mathrm{d}r\ . \tag{A.61}$$

The gravitational force acting on the shell is then given by

$$F_{\mathrm{G}} = -\frac{GM(r)4\pi r^2\,\mathrm{d}r}{r}\ , \tag{A.62}$$

where $M(r)$ is the mass interior to the shell. The total force at all points in the cloud must be zero such that $F_{\mathrm{P}} + F_{\mathrm{G}} = 0$. Substituting in equations A.61 and A.62, multiplying both sides by r and integrating over the cloud radius R gives

$$-\int_0^R 4\pi r^3 \frac{\mathrm{d}P}{\mathrm{d}r}\,\mathrm{d}r = \int_0^R GM(r)4\pi r^2\,\mathrm{d}r\ . \tag{A.63}$$

We note that

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(4\pi r^3 P\right) = 4\pi r^3 \frac{\mathrm{d}P}{\mathrm{d}r} + 3\left(4\pi r^2 P\right)\ , \tag{A.64}$$

and integrating between $r = 0$ and $r = R$ and rearranging yields

$$-\int_0^R 4\pi r^3 \frac{\mathrm{d}P}{\mathrm{d}r}\,\mathrm{d}r = 4\pi r^3 P\Big|_{r=0}^{r=R} + 3\int_0^R 4\pi r^2 P\,\mathrm{d}r\ . \tag{A.65}$$

If we assume that at the edge of the cloud the pressure is zero such that $P(R) = 0$, then the first term on the right hand side is equal to zero. The integral in the second term is the volume-averaged pressure multiplied by the total volume, giving

$$-\int_0^R 4\pi r^3 \frac{\mathrm{d}P}{\mathrm{d}r}\,\mathrm{d}r = 3\left\langle P\right\rangle \frac{4\pi R^3}{3}\ . \tag{A.66}$$

Substituting this into equation A.63 gives

$$3\left\langle P\right\rangle \frac{4\pi R^3}{3} = \int_0^R GM(r)4\pi r^2\,\mathrm{d}r\ . \tag{A.67}$$

Using the ideal gas law $P = c_s^2\rho$, we find $\langle P\rangle = c_s^2\langle\rho\rangle$, and so

$$3\left\langle\rho\right\rangle c_s^2\frac{4\pi R^3}{3} = \int_0^R GM(r)4\pi r^2\,\mathrm{d}r\ . \tag{A.68}$$

The right hand side is the total gravitational potential energy, giving

$$3\left\langle\rho\right\rangle c_s^2\frac{4\pi R^3}{3} = E_{\mathrm{POT}}\ . \tag{A.69}$$

The total potential energy is equal to $-\frac{16}{15}\pi^2 G\rho^2 R^5$ for a uniform density sphere as derived in Section A.3. We can find this more generally by assuming the density has the form $\rho = k\rho_0 r^\alpha$, where k and α are constants and ρ_0 is the volume-averaged density. We can find k by explicitly calculating the volume-averaged density, given by the

$$\left\langle\rho\right\rangle = \frac{3}{4\pi R^3}\int_0^R k\rho_0 r^\alpha 4\pi r^2\,\mathrm{d}r\ . \tag{A.70}$$

This gives

$$\left\langle\rho\right\rangle = \frac{3k\rho_0}{R^3}\frac{R^{3+\alpha}}{3+\alpha}\ . \tag{A.71}$$

Since we defined $\langle\rho\rangle = \rho_0$, this gives

$$k = \frac{3+\alpha}{3R^\alpha}\ , \tag{A.72}$$

and so

$$\rho(r) = \frac{3+\alpha}{3R^\alpha}\rho_0 r^\alpha\ . \tag{A.73}$$

We can now find the total mass contained interior to a radius r, which is given by

$$M(r) = \int_0^r 4\pi r^2\frac{3+\alpha}{3R^\alpha}\rho_0 r^\alpha\,\mathrm{d}r\ , \tag{A.74}$$

and simplifies to

$$M(r) = 4\pi\frac{r^{3+\alpha}}{3R^\alpha}\rho_0\ . \tag{A.75}$$

Finally we can calculate the total gravitational potential energy,

$$E_{\mathrm{POT}} = \int \frac{GM(r)4\pi r^2\rho(r)}{r}\,\mathrm{d}r\ . \tag{A.76}$$

This gives us

$$E_{\mathrm{POT}} = \int_0^R G\,4\pi\frac{r^{3+\alpha}}{3R^\alpha}\rho_0 4\pi r\frac{3+\alpha}{3R^\alpha}\rho_0 r^\alpha\,\mathrm{d}r\ , \tag{A.77}$$

which simplifies to

$$E_{\mathrm{POT}} = \frac{16\pi^2 G}{9(R^\alpha)^2}(3+\alpha)\rho_0^2\int_0^R r^{4+2\alpha}\,\mathrm{d}r\ , \tag{A.78}$$

and finally

$$E_{\mathrm{POT}} = \frac{16\pi^2 G}{9}\frac{3+\alpha}{5+2\alpha}\rho_0^2 R^5\ . \tag{A.79}$$

We define

$$f = \frac{3+\alpha}{5+2\alpha}\ , \tag{A.80}$$

and note that $f = {}^3/_5$ for a cloud of uniform density ($\alpha = 0$). We can therefore rewrite equation A.79 in terms of the average density and the purely numerical factor f which accounts for the density structure within the cloud. This gives

$$E_{\mathrm{POT}} = \frac{16\pi^2 G}{9} f\,\langle\rho\rangle^2 R^5\ . \tag{A.81}$$

Substituting this into equation A.69, we obtain

$$3c_{\mathrm{s}}^2 = \frac{4\pi G}{3} f\,\langle\rho\rangle R^2\ . \tag{A.82}$$

and thus we obtain the *Jeans radius*

$$R_{\mathrm{J}} = \sqrt{\frac{9}{4\pi G f\,\langle\rho\rangle}}\,c_{\mathrm{s}}\ . \tag{A.83}$$

In the case of constant density where $f = {}^3/_5$, this reduces to the form calculated in Section A.3.

A.5 Jeans mass

From Section A.2 we have the Jeans wavelength,

$$\lambda_{\mathrm{J}} = \sqrt{\frac{\pi}{G\rho}}c_{\mathrm{s}}\ , \tag{A.84}$$

and from Sections A.3 and A.4, we have the Jeans radius

$$R_{\mathrm{J}} = \sqrt{\frac{15}{4\pi G\rho}} c_{\mathrm{s}} \,. \tag{A.85}$$

We can now calculate the Jeans mass, which is the mass of a spherical cloud of uniform density with a radius equal to the Jeans radius. The mass of the cloud is

$$M_{\mathrm{J}} = \frac{4\pi R_{\mathrm{J}}^3}{3}\rho \,, \tag{A.86}$$

where R_{J} is the Jeans radius and ρ is the density of the cloud.

Using the Jeans wavelength from section A.2, and noting $\lambda_{\mathrm{J}} = 2R_{\mathrm{J}}$, we obtain a Jeans mass

$$M_{\mathrm{J}} = \frac{1}{6}\sqrt{\frac{\pi^5}{G^3\rho}} c_{\mathrm{s}}^3 \,. \tag{A.87}$$

The numerical factors simplify to ~ 2.92 in this case, which reflects the maximum mass of stable oscillations of small spherical perturbations in density in an infinite, approximately uniform density medium.

Using the Jeans radius from sections A.3 and A.4, we obtain a Jeans mass

$$M_{\mathrm{J}} = \sqrt{\frac{375}{4\pi G^3\rho}} c_{\mathrm{s}}^3 \,. \tag{A.88}$$

This Jeans mass represents the maximum mass that a uniform density isothermal cloud can have while remaining virialized and supported by hydrostatic pressure gradients against gravitational collapse. Here the numerical factors simplify to ~ 5.46, which makes the Jeans mass calculated from the Jeans wavelength ~ 1.87 times smaller than the Jeans mass calculated from the consideration of the Virial theorem or of the hydrodynamic forces supporting a cloud. We choose to use the Jeans radius from the consideration of hydrodynamic forces described in Section A.4.

Appendix B

Plots of simulations of the NTSI

In this appendix we provide example plots for most of the simulations described in Chapter 9. We only include plots from a single realization for each set of simulation parameters, and we only include plots for a limited set of monochromatic wavenumbers. Plots are either cross-section plots or column density plots produced using SPLASH (Price 2007) as described in Sections 5.1.1 and 5.1.2 respectively. Table B.1 shows which plots are included in this appendix. Each figure includes plots at six timesteps: 0.005, 0.02, 0.035, 0.05, 0.065 and 0.095 Myr.

Cross-sections are taken through the $z = 0$ plane, with the exception of the simulations of two-dimensional monochromatic perturbations. These cross-sections are taken through the $z = 0.063\,\mathrm{pc}$, $z = 0.016\,\mathrm{pc}$ and $z = 0.0078\,\mathrm{pc}$ planes for the $k = 4\,\mathrm{pc}^{-1}$, $k = 16\,\mathrm{pc}^{-1}$ and $k = 32\,\mathrm{pc}^{-1}$ wavenumbers respectively. These planes correspond to the first peaks of the two-dimensional perturbations in the z axis, as the perturbation are of zero amplitude at $z = 0$.

Column density plots are taken across the yz axis to illustrate the nature of the two-dimensional perturbation. The initial velocity perturbation does not produce any perturbation in column density, as described in Section 10.1. Any perturbations observed in column density are therefore purely the result of the growth of the NTSI.

Figure	Suite[1]	Type of plot	Type of perturbations
B.1	1DLR	cross-section	Monochromatic, $k = 4\,\mathrm{pc}^{-1}$
B.2	1DLR	cross-section	Monochromatic, $k = 16\,\mathrm{pc}^{-1}$
B.3	1DLR	cross-section	Monochromatic, $k = 32\,\mathrm{pc}^{-1}$
B.4	1DHR	cross-section	Monochromatic, $k = 4\,\mathrm{pc}^{-1}$
B.5	1DHR	cross-section	Monochromatic, $k = 16\,\mathrm{pc}^{-1}$
B.6	1DHR	cross-section	Monochromatic, $k = 32\,\mathrm{pc}^{-1}$
B.7	1DHR	cross-section	White noise
B.8	1DHR	cross-section	Subsonic turbulence
B.9	1DHR	cross-section	Supersonic turbulence
B.10	2D	cross-section	Monochromatic, $k_y = k_z = 4\,\mathrm{pc}^{-1}$
B.11	2D	cross-section	Monochromatic, $k_y = k_z = 16\,\mathrm{pc}^{-1}$
B.12	2D	cross-section	Monochromatic, $k_y = k_z = 32\,\mathrm{pc}^{-1}$
B.13	2D	cross-section	White noise
B.14	2D	cross-section	Subsonic turbulence
B.15	2D	cross-section	Supersonic turbulence
B.16	2D	column density	Monochromatic, $k_y = k_z = 4\,\mathrm{pc}^{-1}$
B.17	2D	column density	Monochromatic, $k_y = k_z = 16\,\mathrm{pc}^{-1}$
B.18	2D	column density	Monochromatic, $k_y = k_z = 32\,\mathrm{pc}^{-1}$
B.19	2D	column density	White noise
B.20	2D	column density	Subsonic turbulence
B.21	2D	column density	Supersonic turbulence

[1] For a description of the simulation suites, see Section 9.1.

Table B.1 – The range of plots included in this appendix.

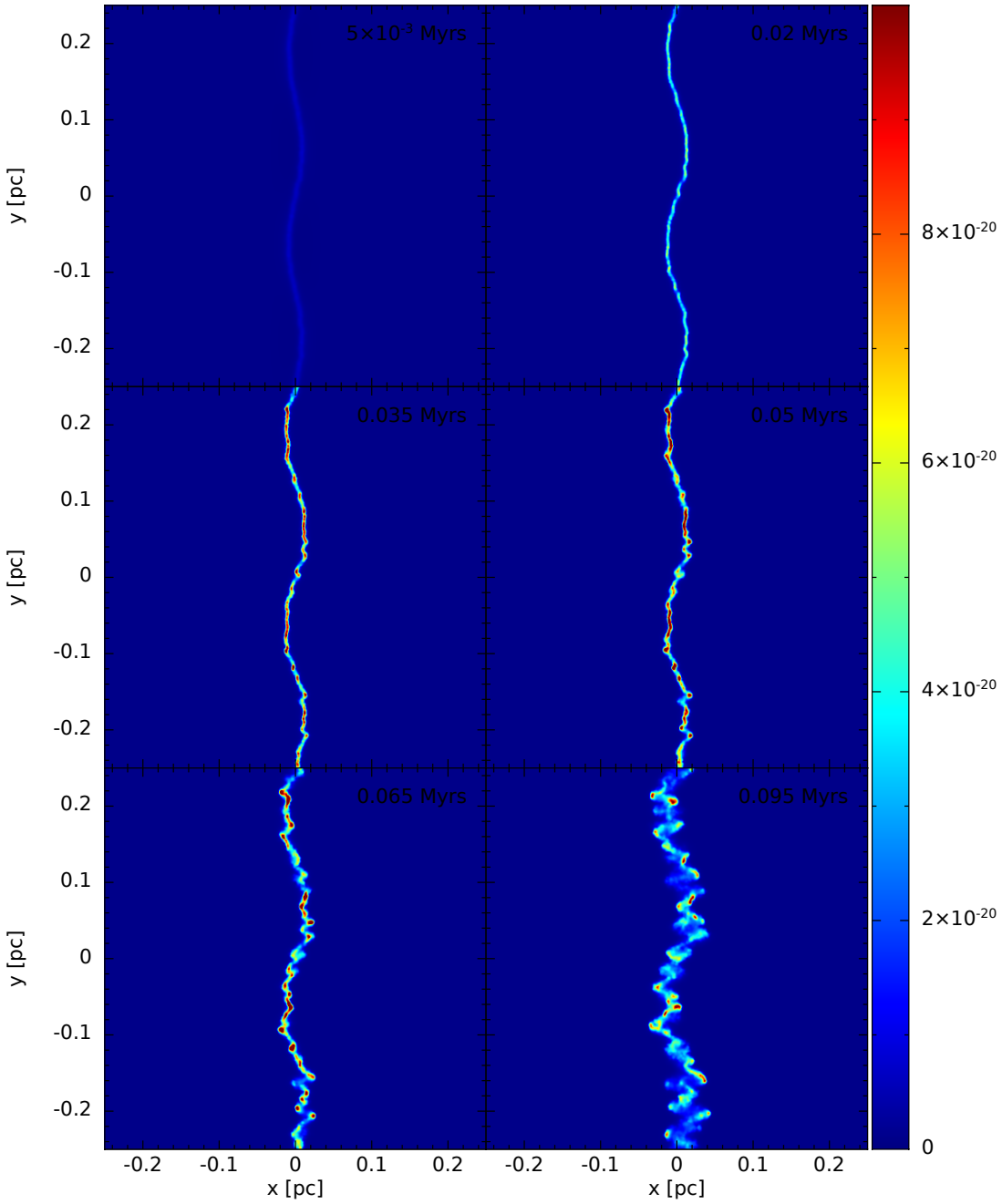


Figure B.1 – Cross-sections of density ($\mathrm{g\,cm^{-3}}$) in the xy plane for 1DLR simulation of monochromatic perturbations of wavenumber $k = 4\,\mathrm{pc}^{-1}$.

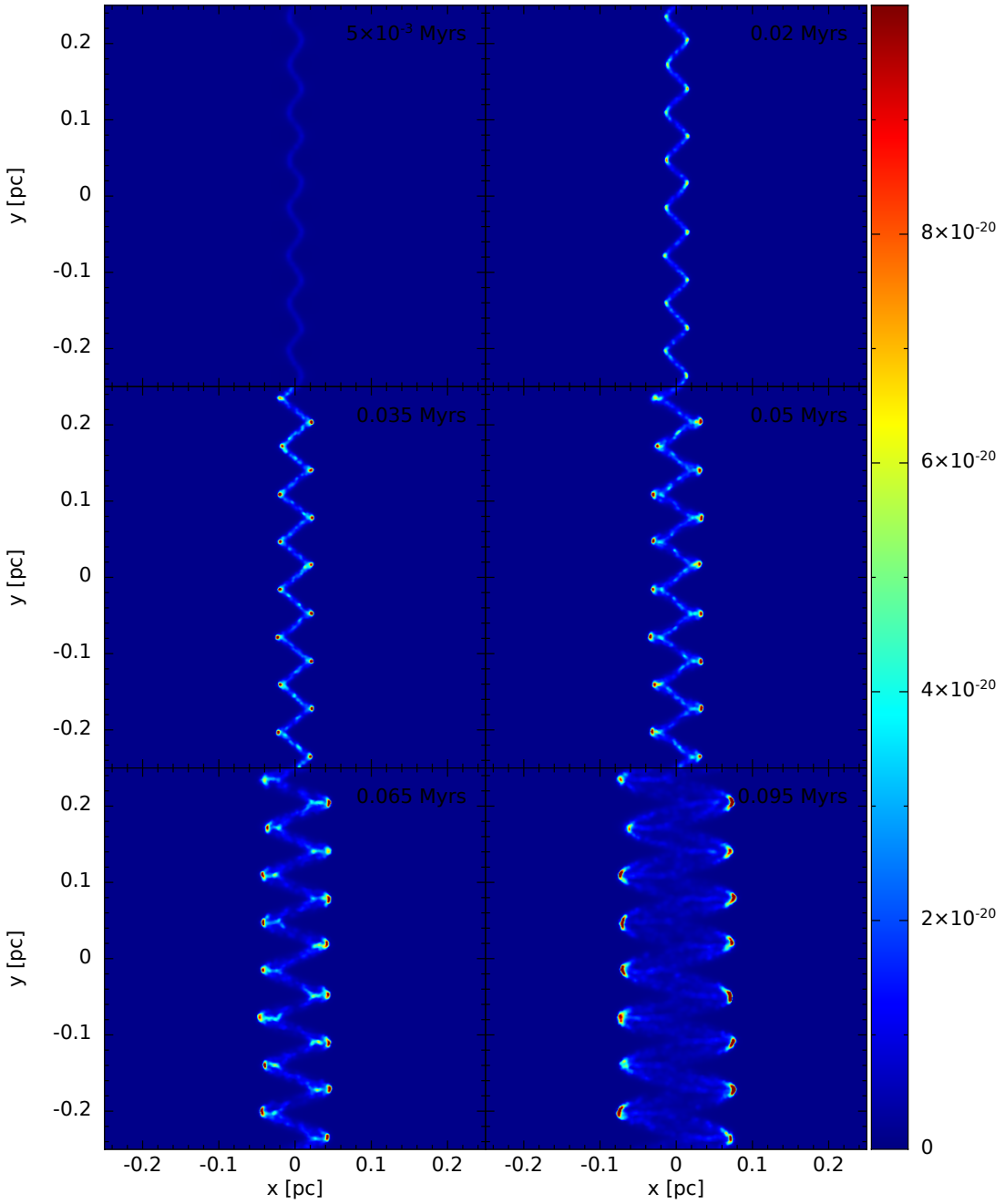


Figure B.2 – Cross-sections of density ($\mathrm{g\,cm^{-3}}$) in the xy plane for 1DLR simulation of monochromatic perturbations of wavenumber $k = 16\,\mathrm{pc}^{-1}$.

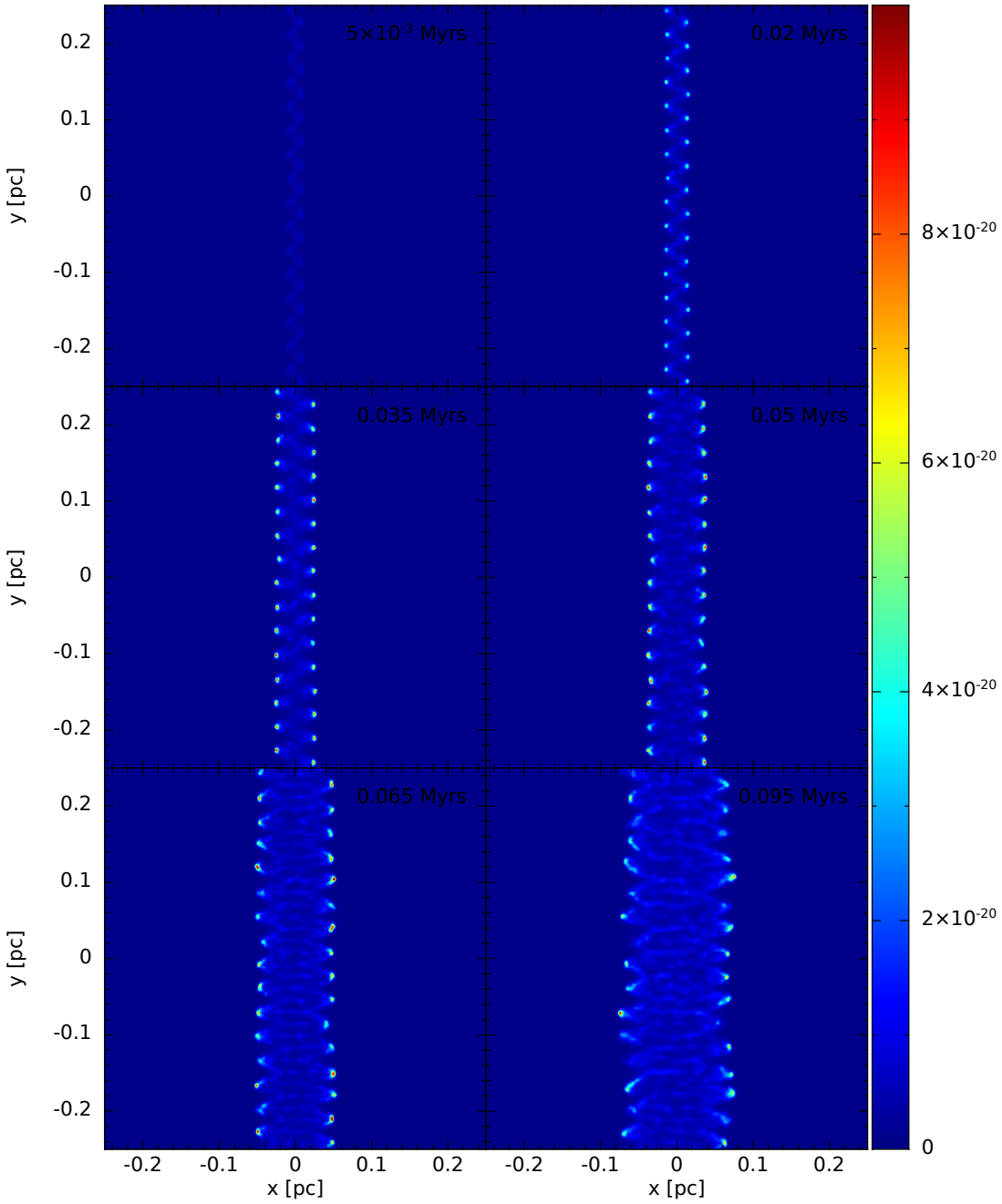


Figure B.3 – Cross-sections of density ($\mathrm{g\,cm^{-3}}$) in the xy plane for 1DLR simulation of monochromatic perturbations of wavenumber $k = 32\,\mathrm{pc}^{-1}$.

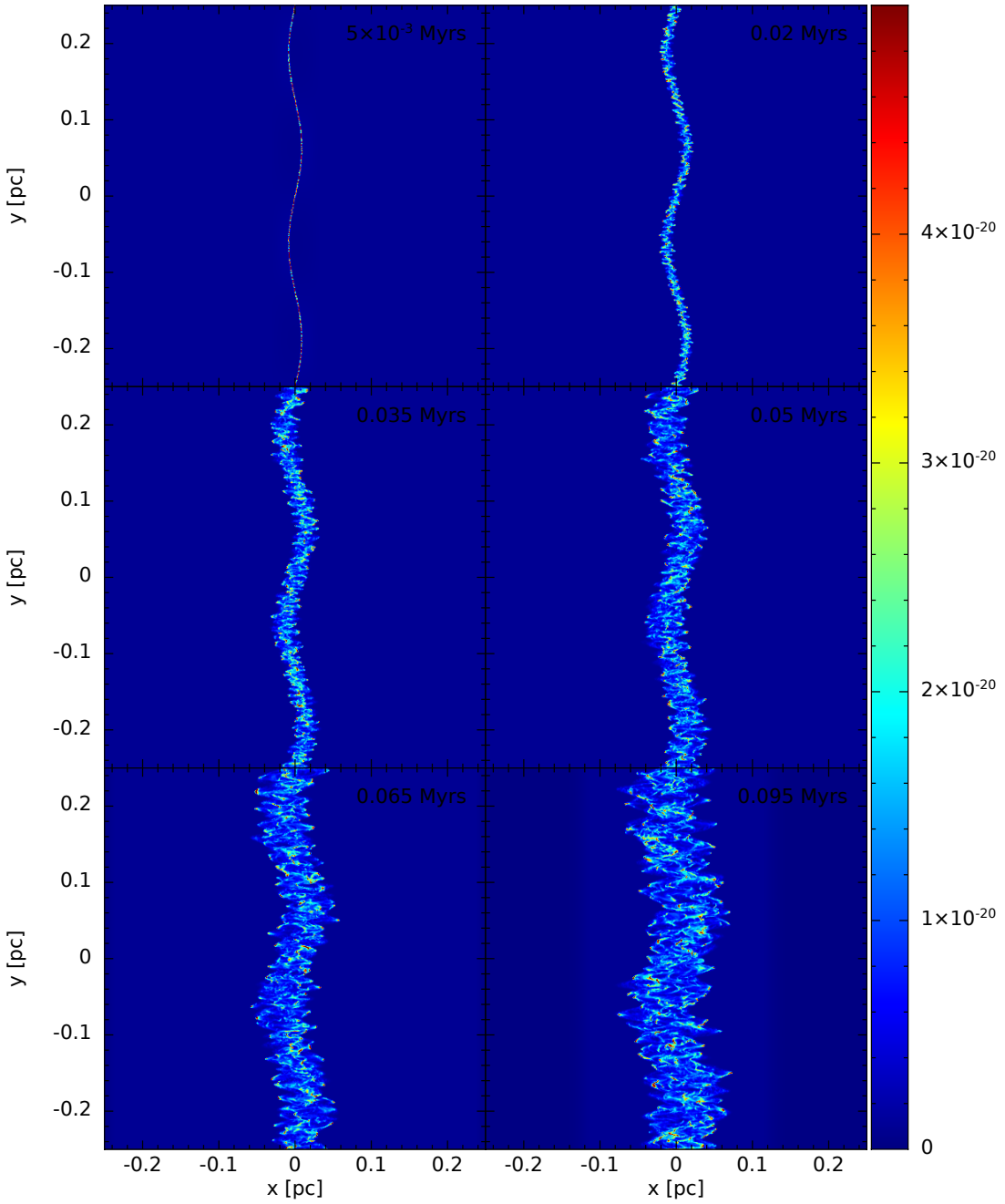


Figure B.4 – Cross-sections of density ($\mathrm{g\,cm^{-3}}$) in the xy plane for 1DHR simulation of monochromatic perturbations of wavenumber $k = 4\,\mathrm{pc}^{-1}$.

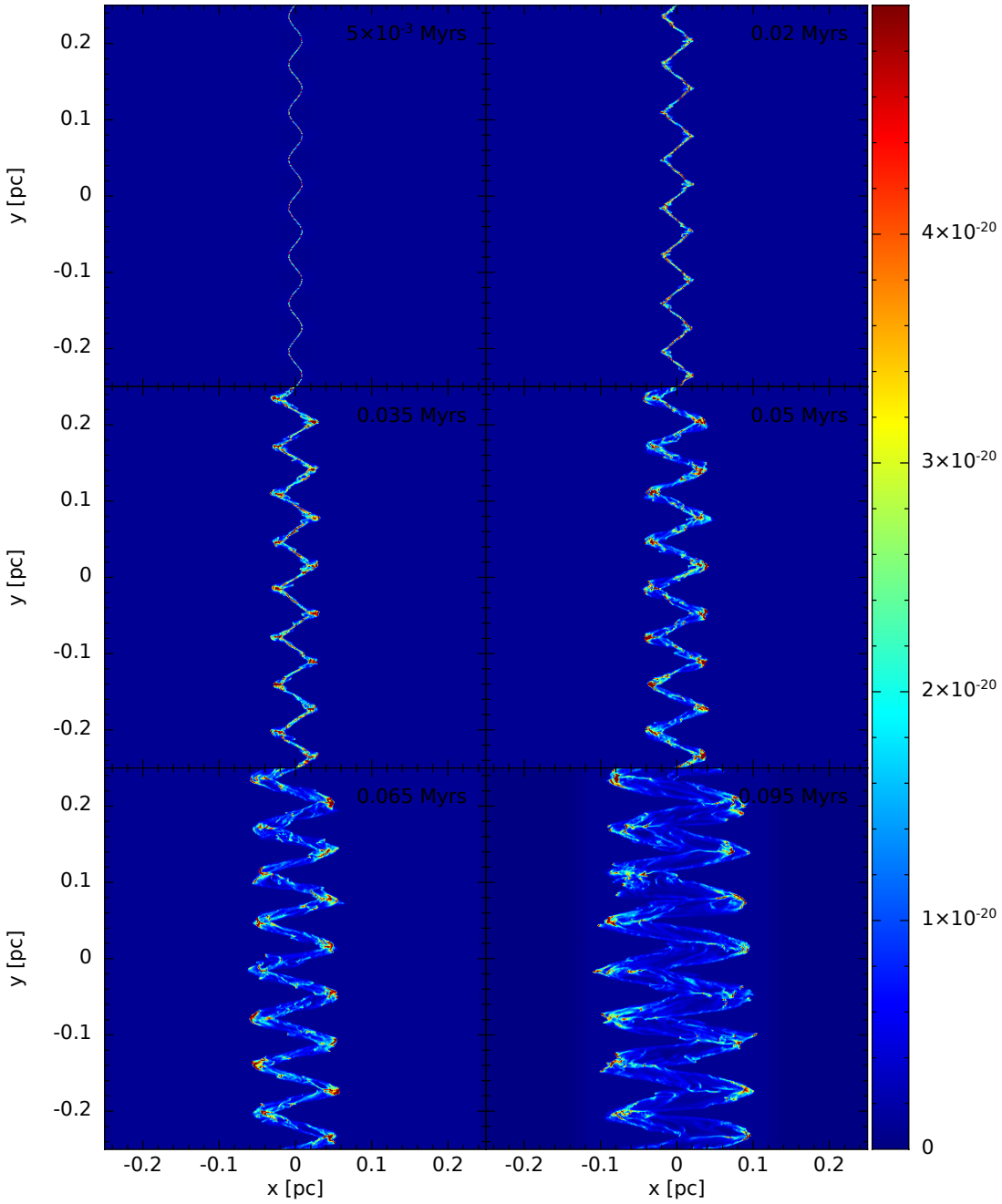


Figure B.5 – Cross-sections of density ($g\,cm^{-3}$) in the xy plane for 1DHR simulation of monochromatic perturbations of wavenumber $k = 16\,pc^{-1}$.

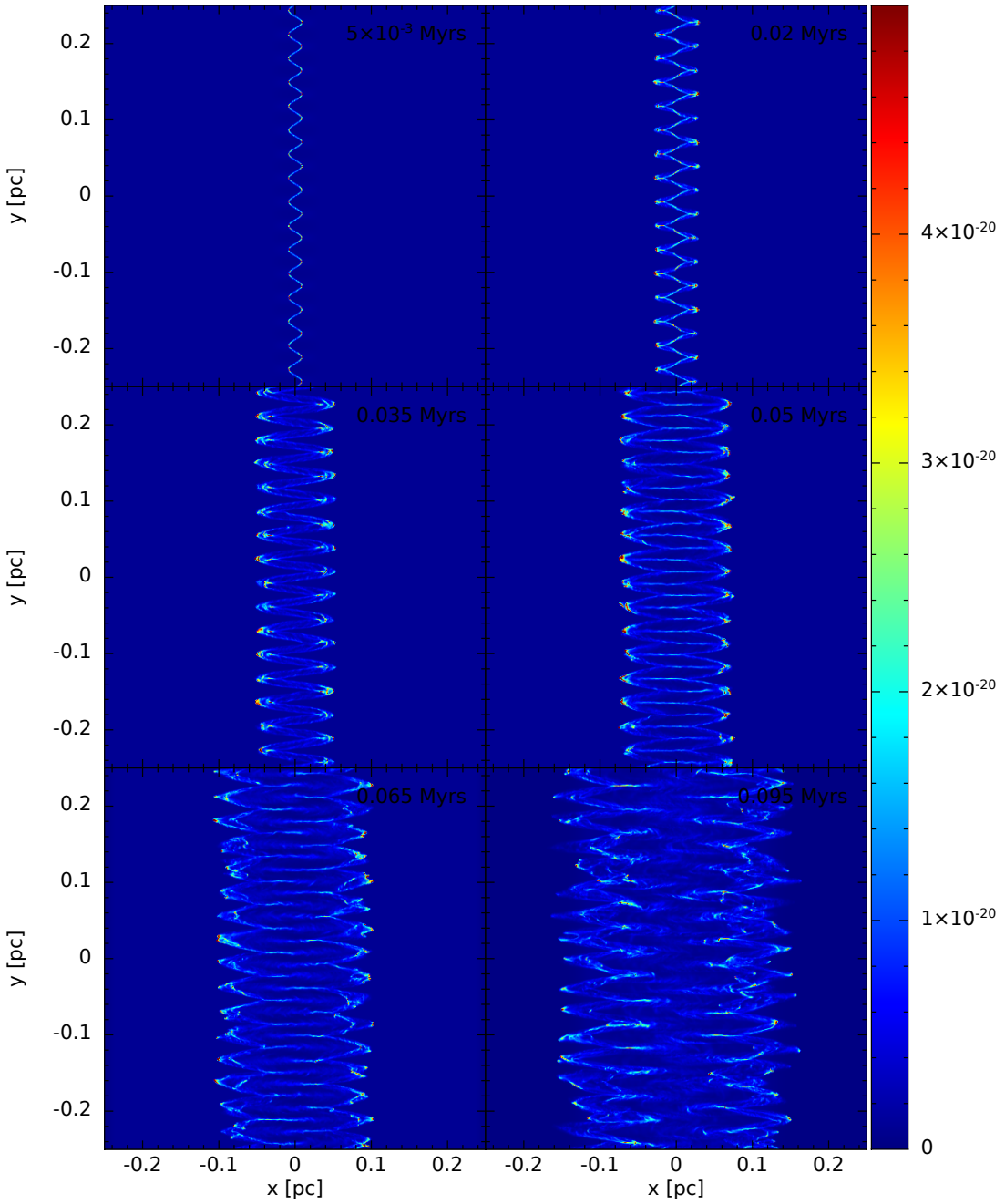


Figure B.6 – Cross-sections of density ($\mathrm{g\,cm^{-3}}$) in the xy plane for 1DHR simulation of monochromatic perturbations of wavenumber $k = 32\,\mathrm{pc}^{-1}$.

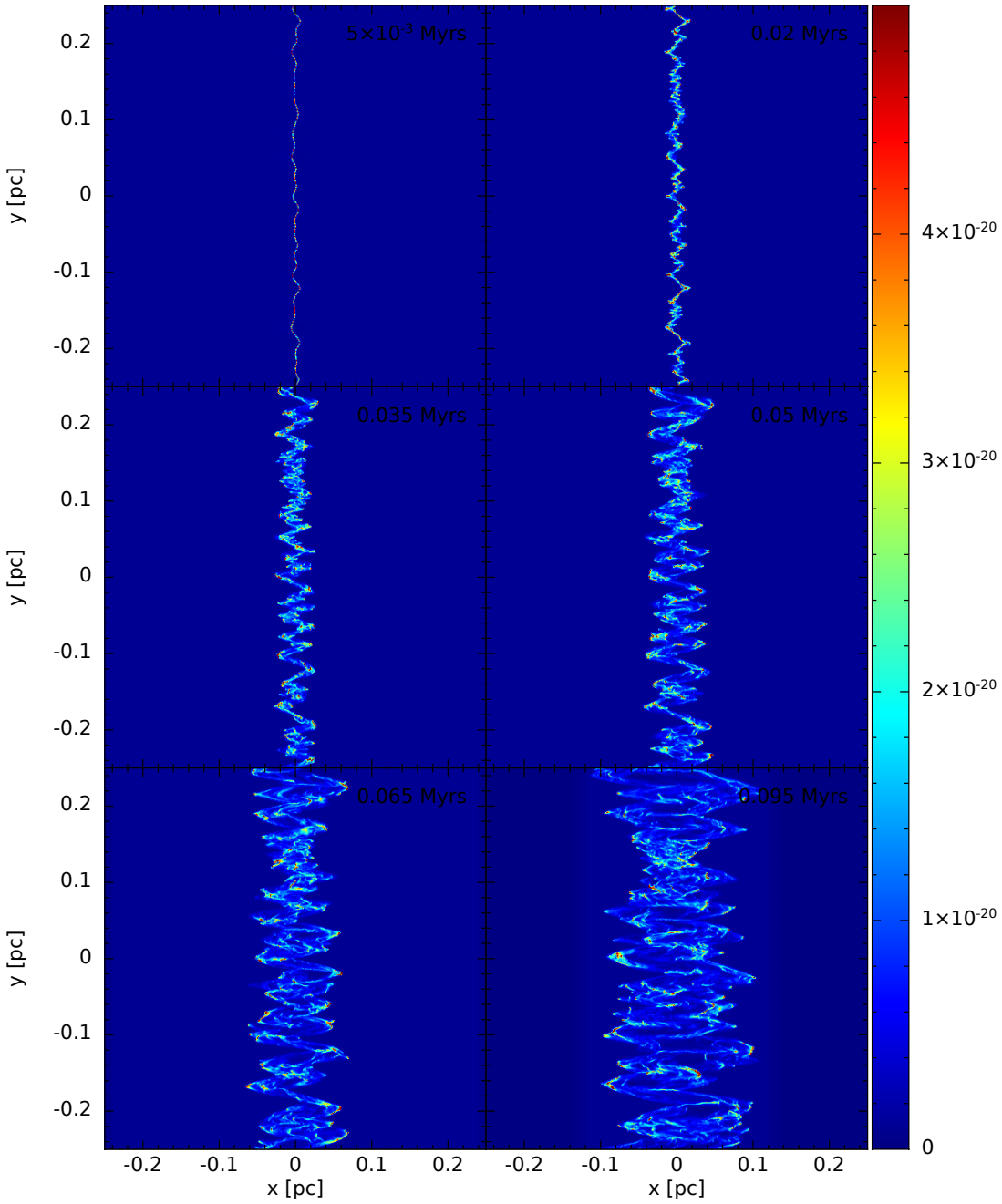


Figure B.7 – Cross-sections of density ($g\,cm^{-3}$) in the xy plane for 1DHR simulation of white noise perturbations.

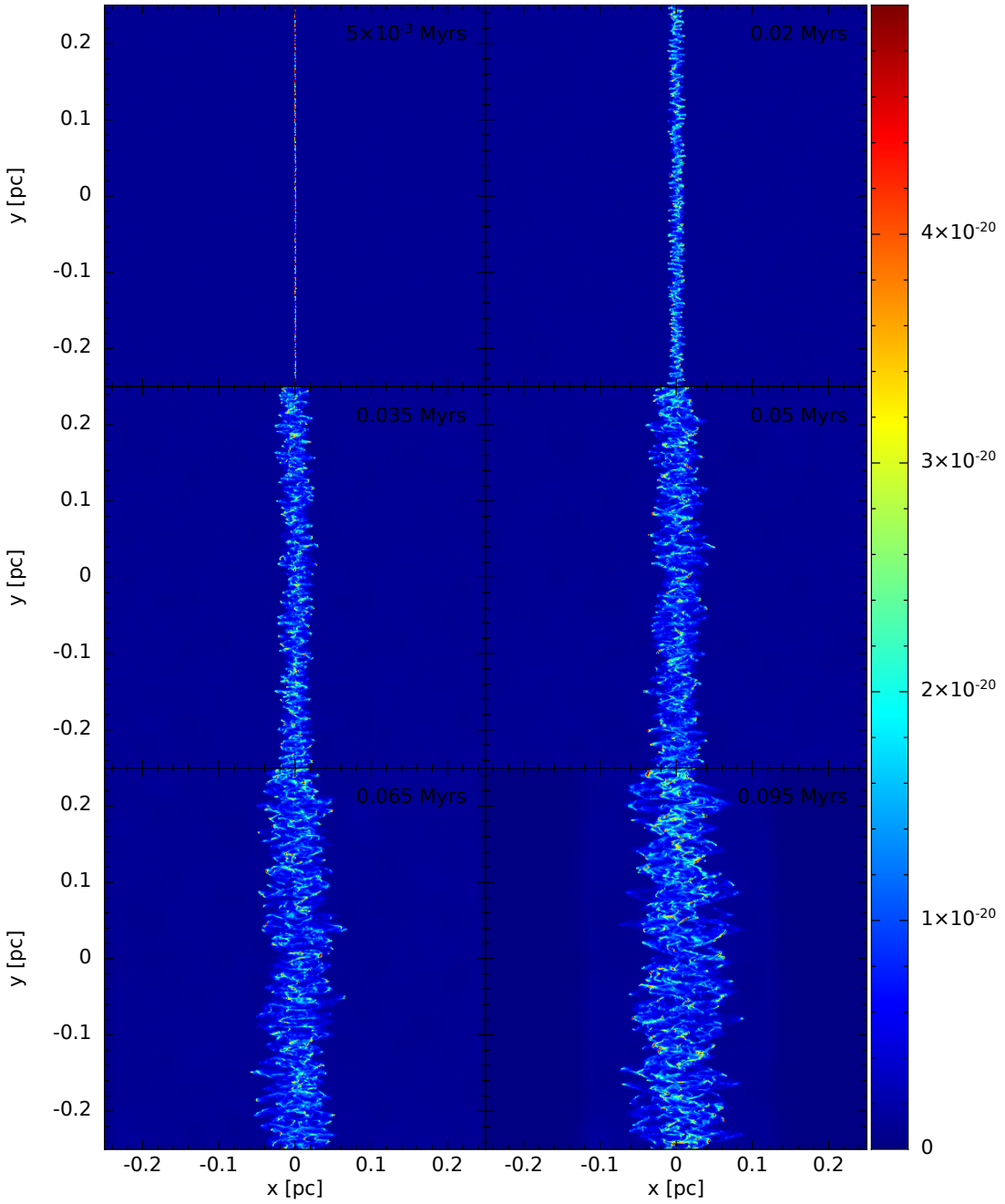


Figure B.8 – Cross-sections of density ($\mathrm{g\,cm^{-3}}$) in the xy plane for 1DHR simulation with subsonic turbulence.

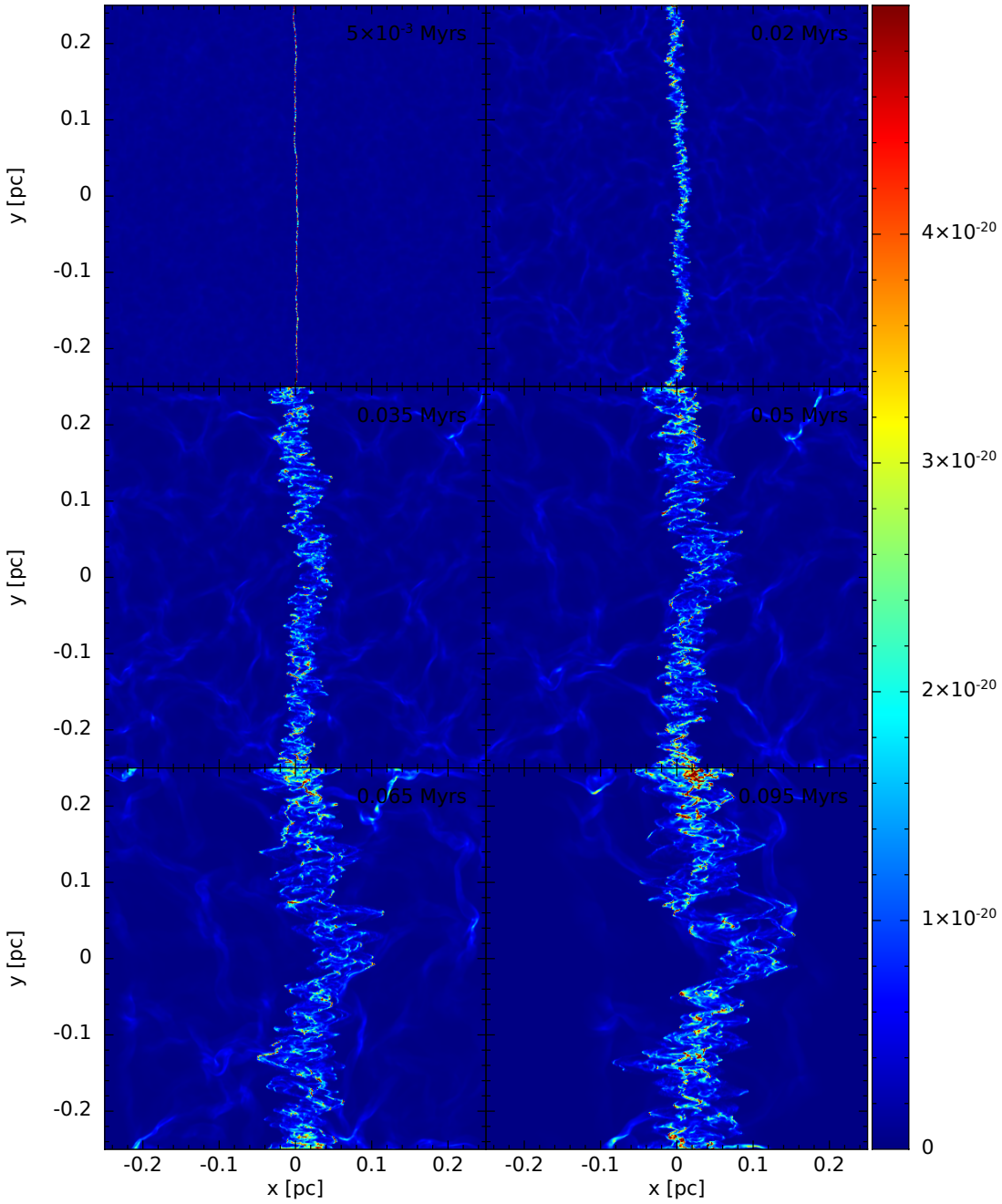


Figure B.9 – Cross-sections of density ($g\,cm^{-3}$) in the xy plane for 1DHR simulation with supersonic turbulence.

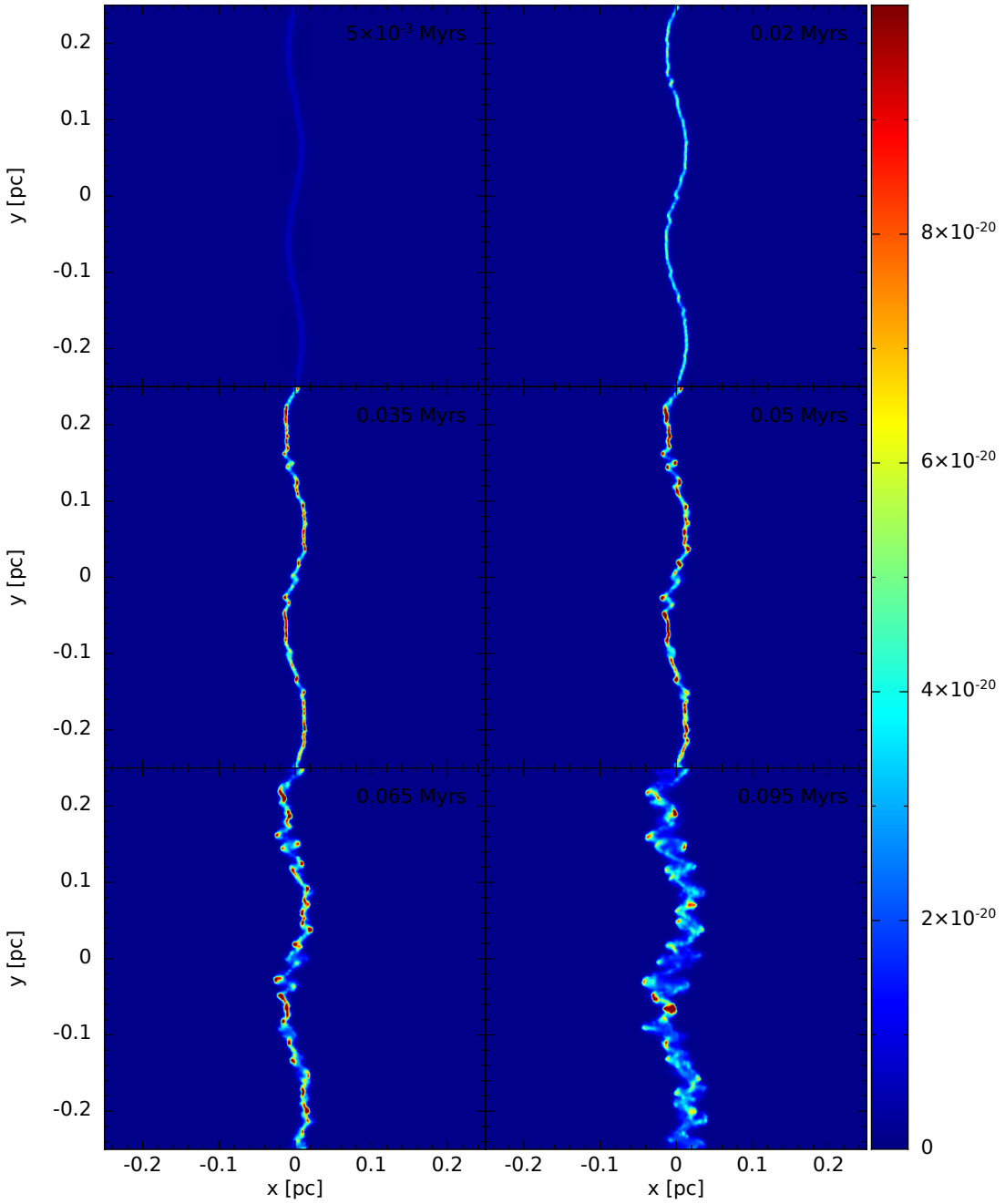


Figure B.10 – Cross-sections of density ($\mathrm{g\,cm^{-3}}$) in the xy plane for 2D simulation of monochromatic perturbations of wavenumber $k_y = k_z = 4\,\mathrm{pc}^{-1}$.

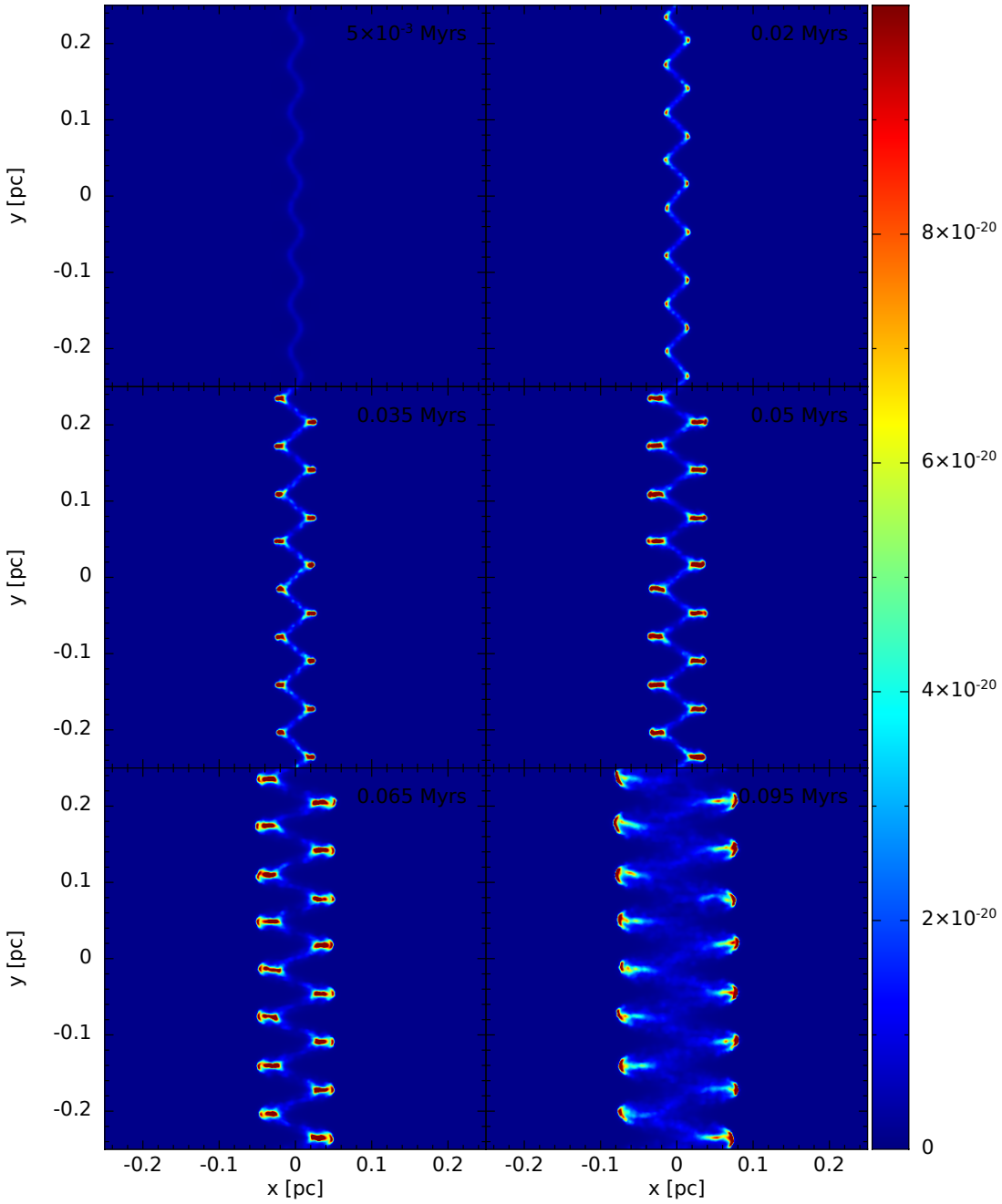


Figure B.11 – Cross-sections of density ($\mathrm{g\,cm^{-3}}$) in the xy plane for 2D simulation of monochromatic perturbations of wavenumber $k_y = k_z = 16\,\mathrm{pc}^{-1}$.

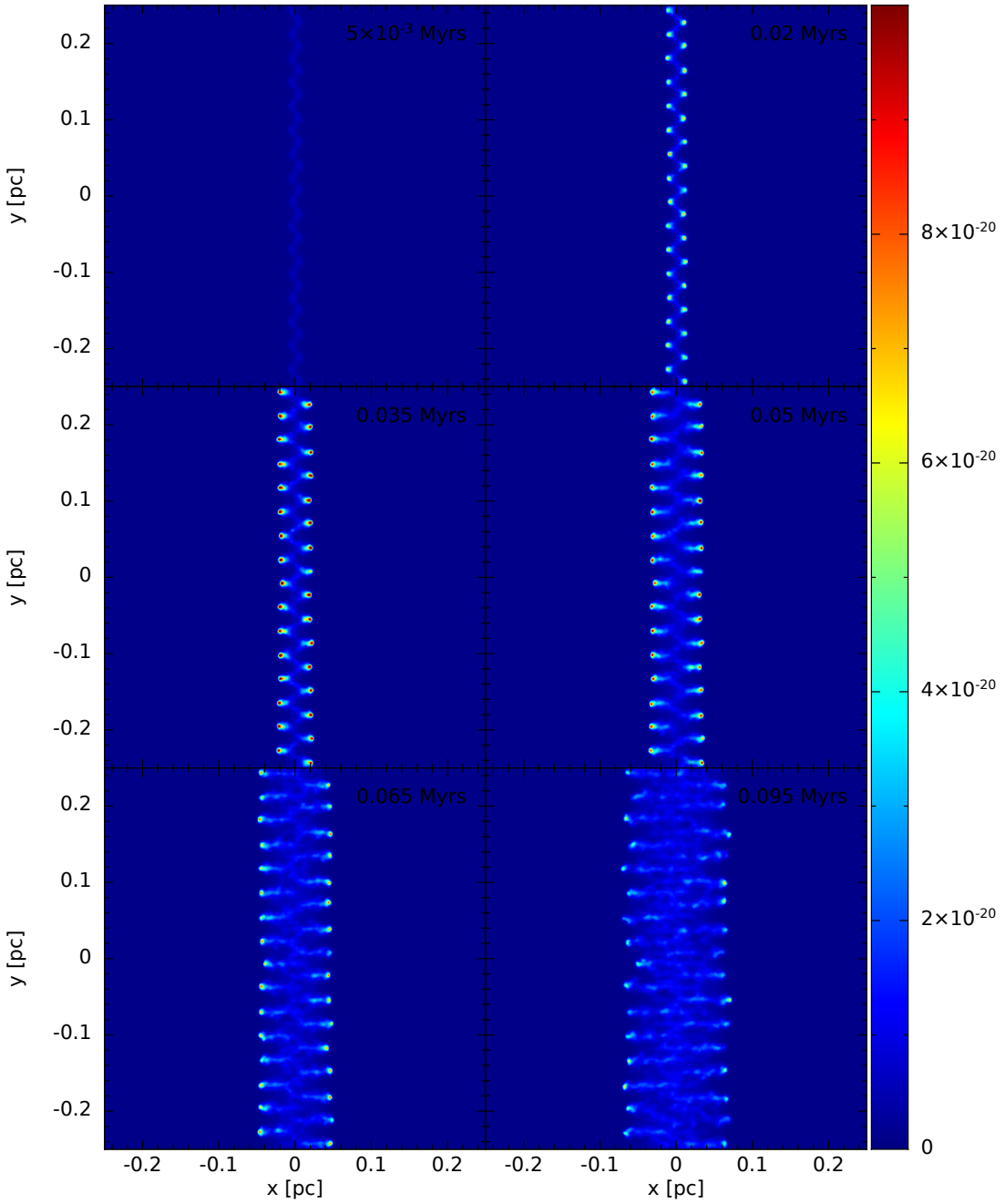


Figure B.12 – Cross-sections of density ($\mathrm{g\,cm^{-3}}$) in the xy plane for 2D simulation of monochromatic perturbations of wavenumber $k_y = k_z = 32\,\mathrm{pc}^{-1}$.

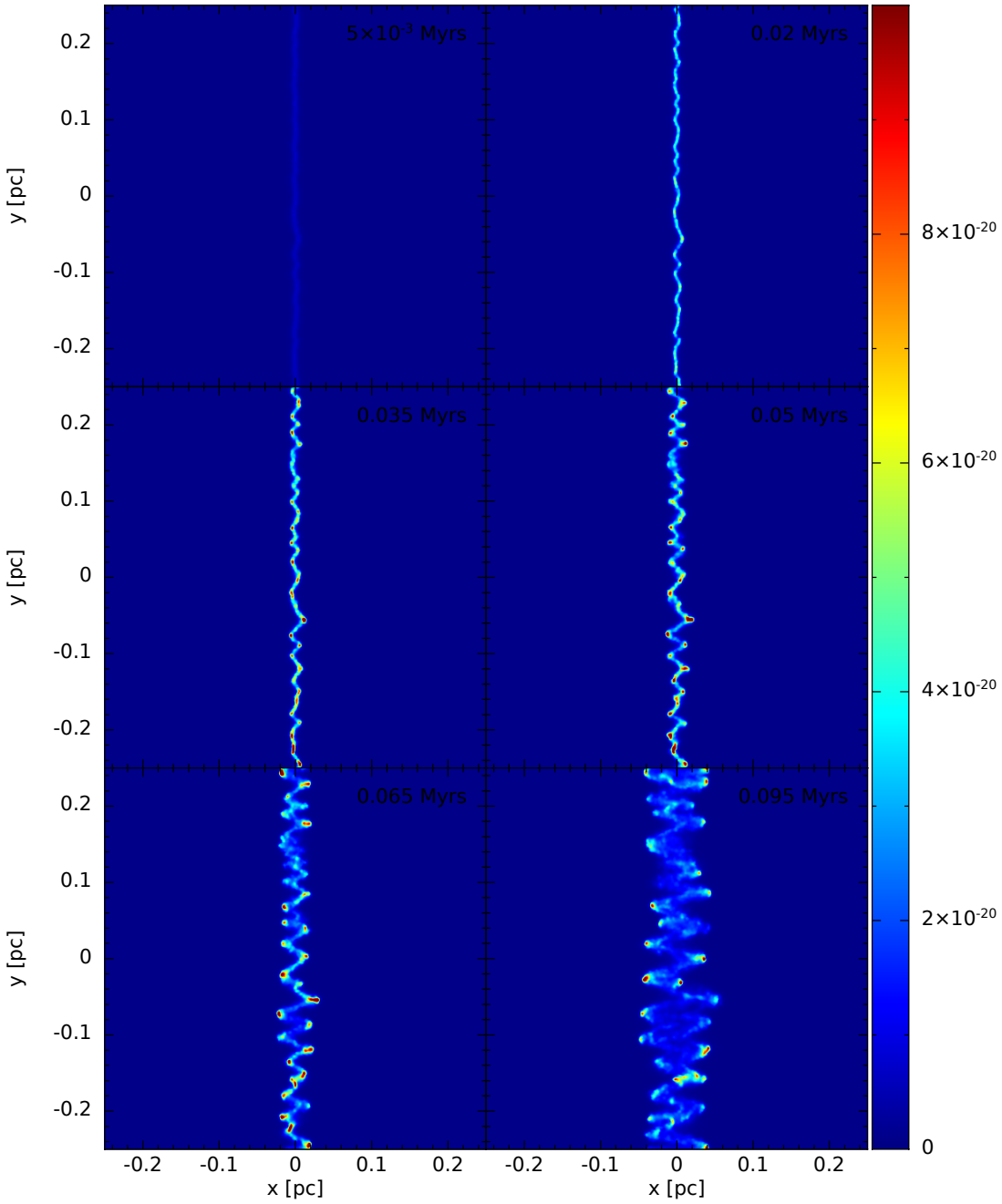


Figure B.13 – Cross-sections of density ($\mathrm{g\,cm^{-3}}$) in the xy plane for 2D simulation of white noise perturbations.

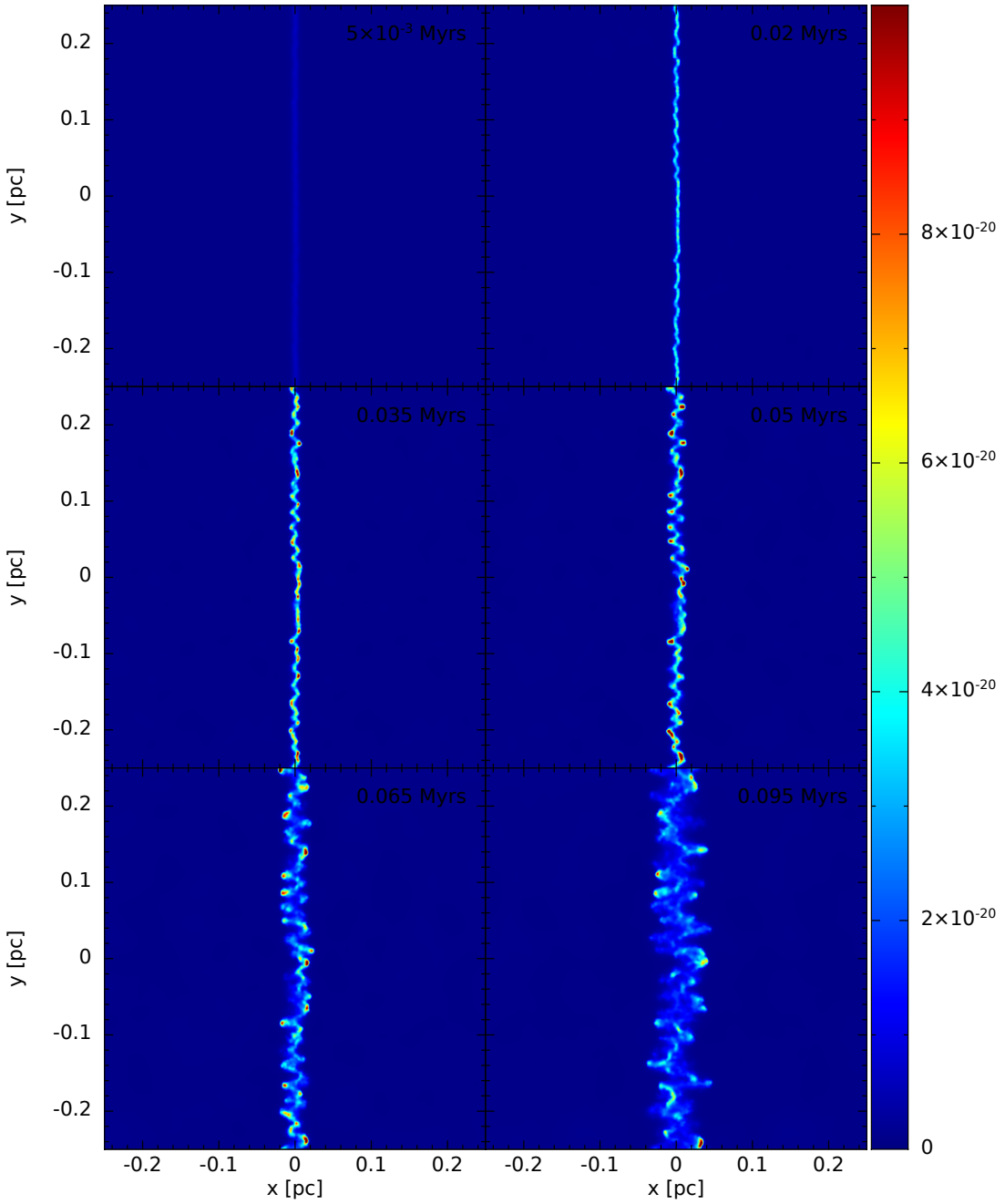


Figure B.14 – Cross-sections of density ($\mathrm{g\,cm^{-3}}$) in the xy plane for 2D simulation with subsonic turbulence.

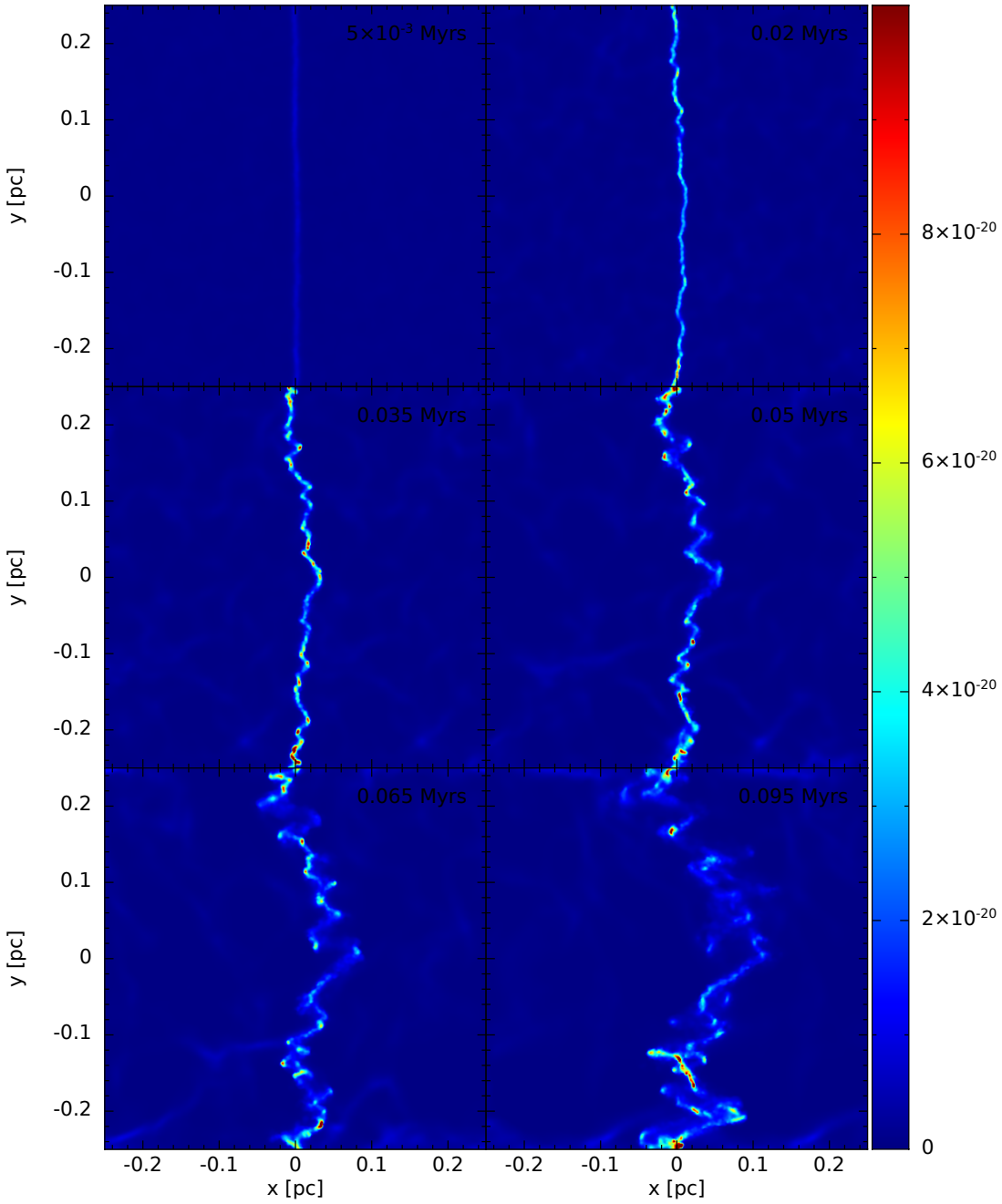


Figure B.15 – Cross-sections of density ($g\,cm^{-3}$) in the xy plane for 2D simulation with supersonic turbulence.

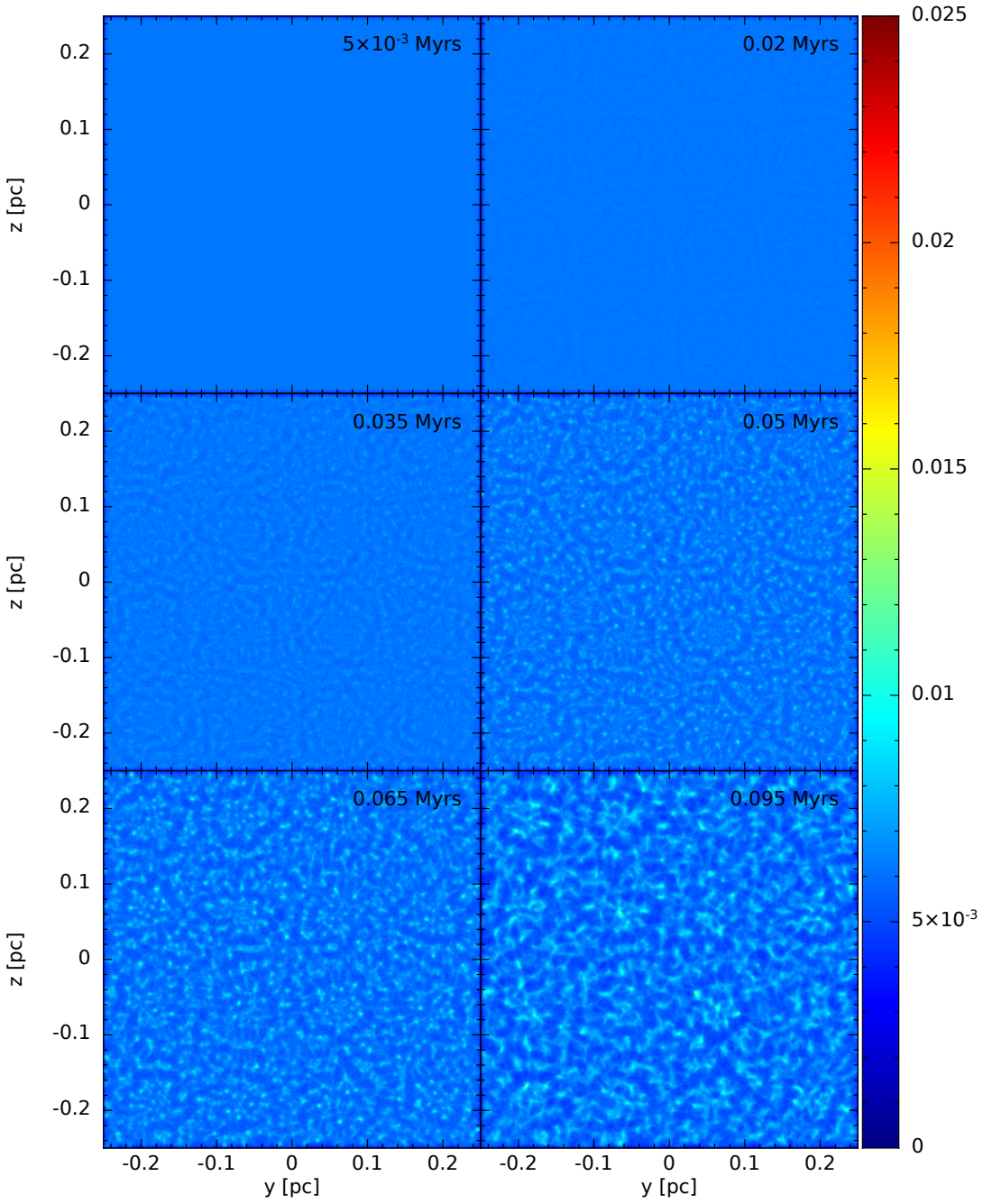


Figure B.16 – Column density ($\mathrm{g\,cm^{-2}}$) in the yz plane for 2D simulation of monochromatic perturbations of wavenumber $k_y = k_z = 4\,\mathrm{pc}^{-1}$.

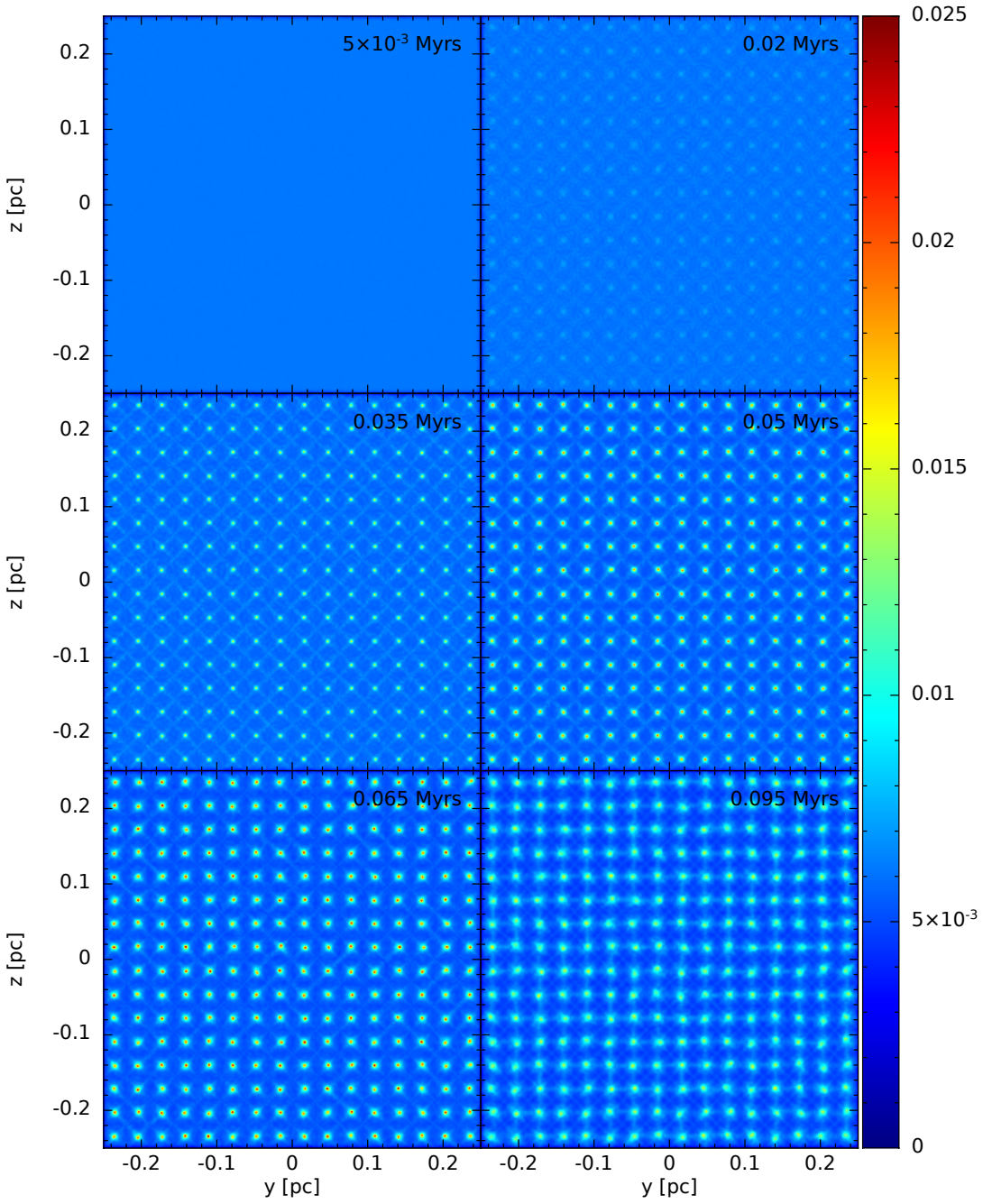


Figure B.17 – Column density ($\mathrm{g\,cm^{-2}}$) in the yz plane for 2D simulation of monochromatic perturbations of wavenumber $k_y = k_z = 16\,\mathrm{pc}^{-1}$.

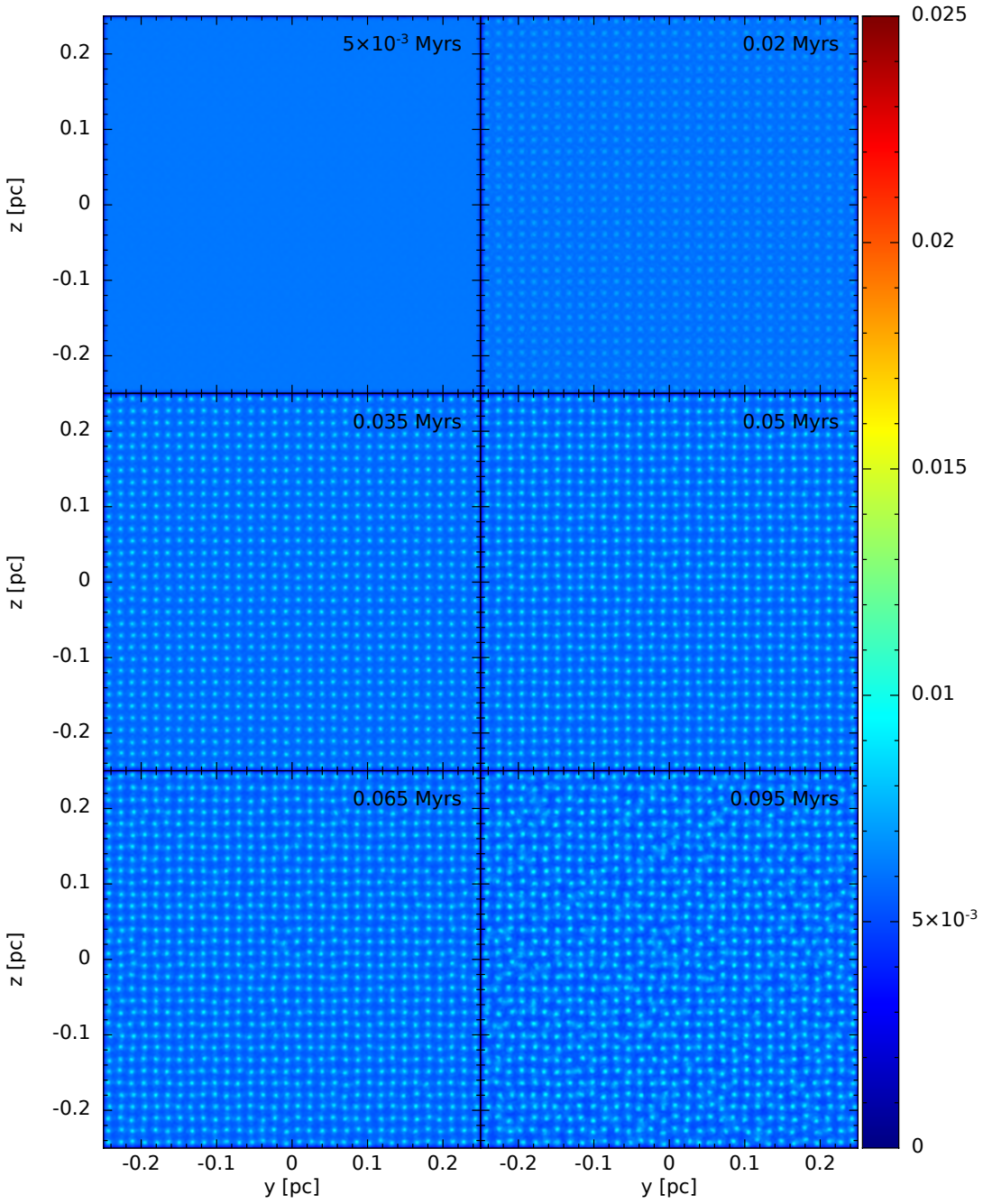


Figure B.18 – Column density ($\mathrm{g\,cm^{-2}}$) in the yz plane for 2D simulation of monochromatic perturbations of wavenumber $k_y = k_z = 32\,\mathrm{pc^{-1}}$.

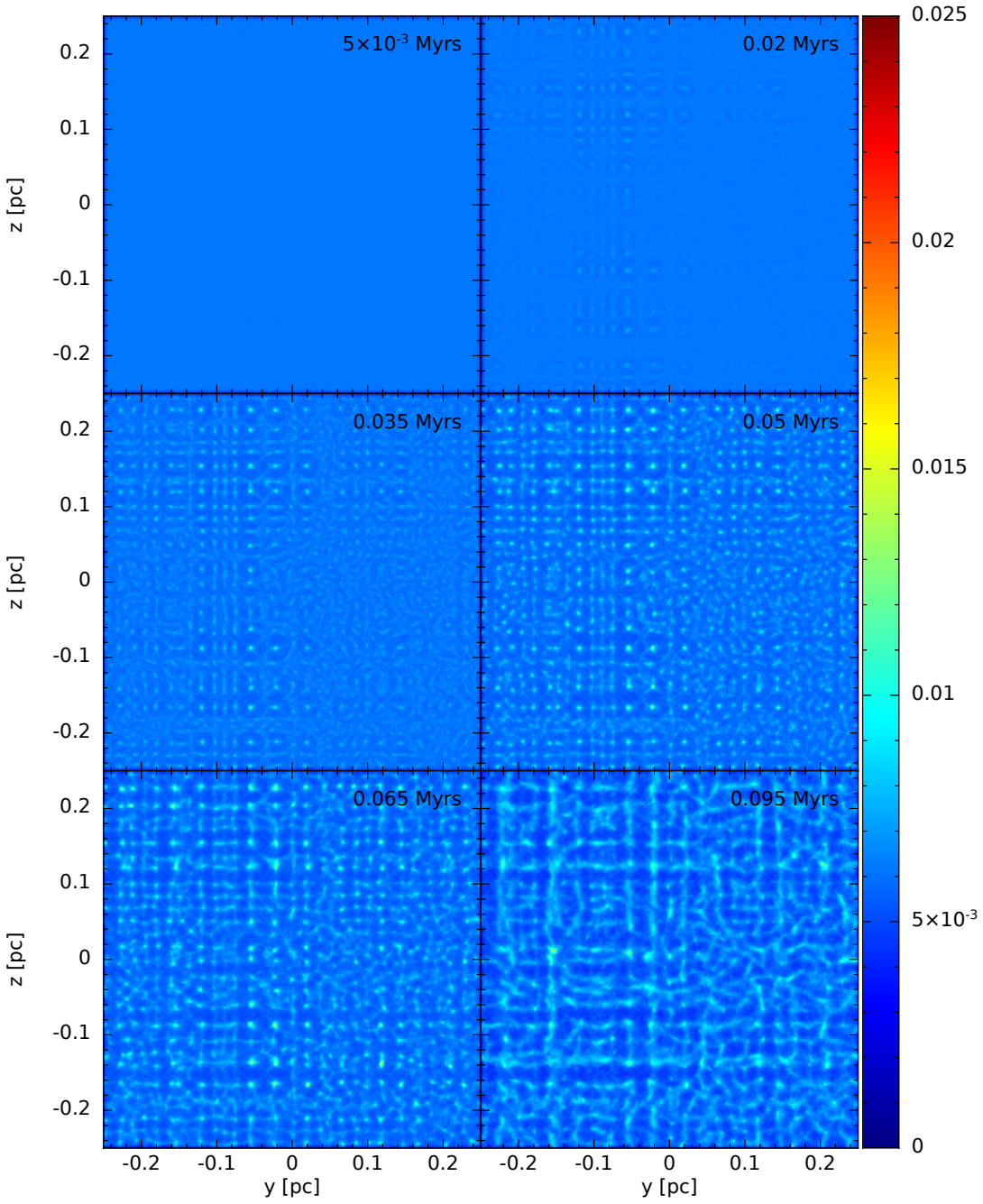


Figure B.19 – Column density ($g\,cm^{-2}$) in the yz plane for 2D simulation of white noise perturbations.

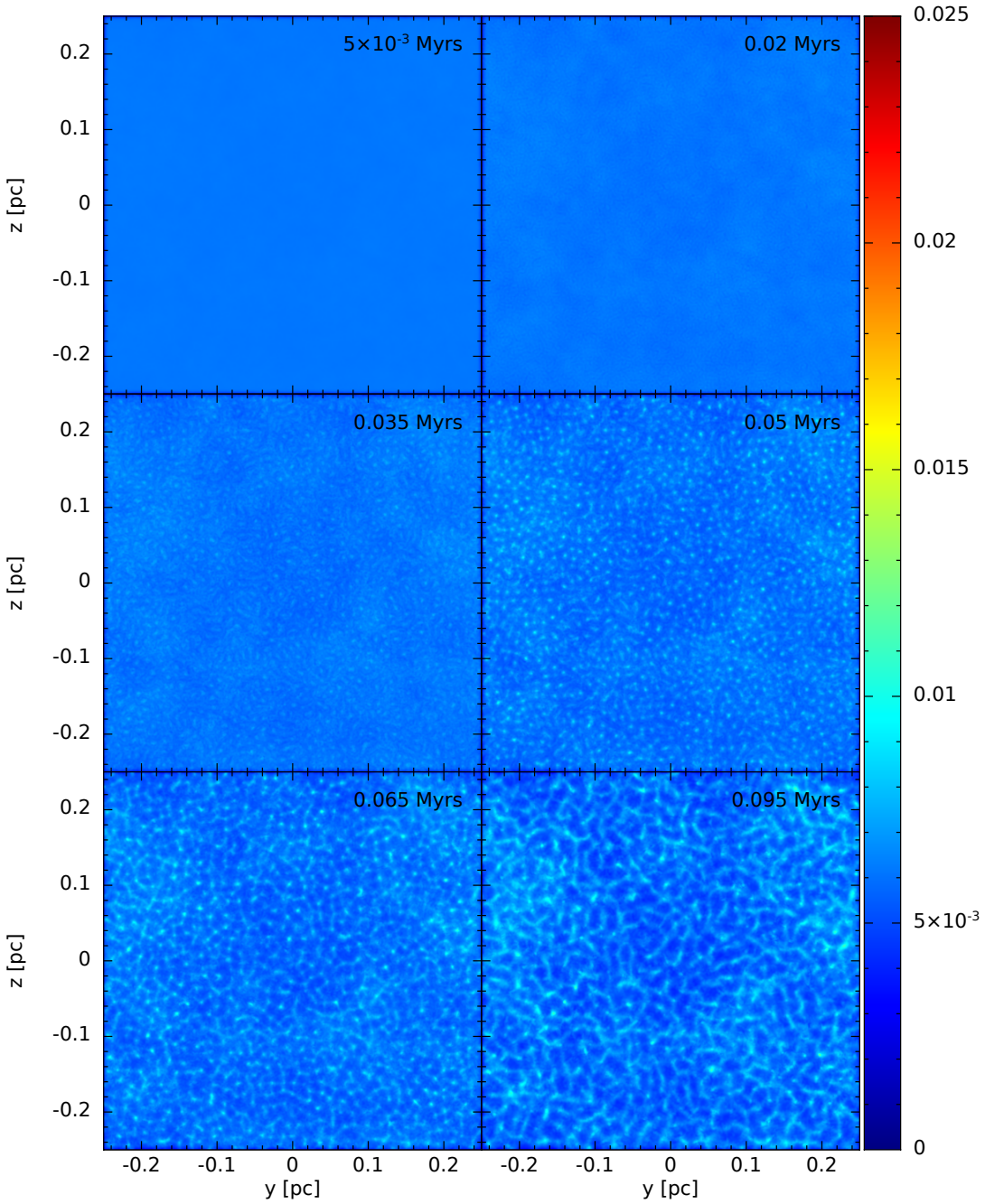


Figure B.20 – Column density ($\mathrm{g\,cm^{-2}}$) in the yz plane for 2D simulation with subsonic turbulence.

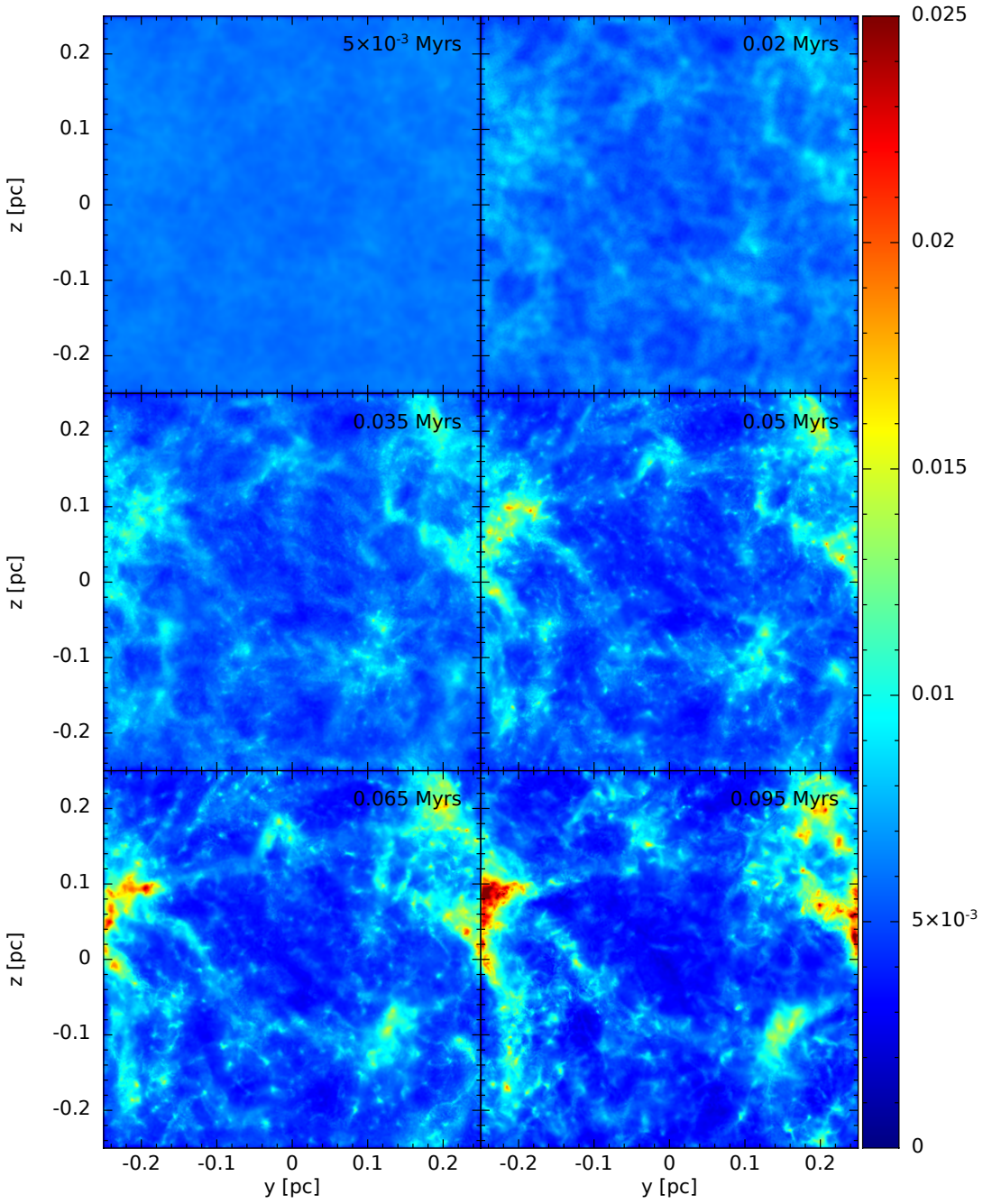


Figure B.21 – Column density ($\mathrm{g\,cm^{-2}}$) in the yz plane for 2D simulation with supersonic turbulence.