Binary formation and mass function variations in fragmenting discs with short cooling times

R.D.Alexander[1,2]⋆, P.J.Armitage[2,3] and J.Cuadra[2]

[1] *Sterrewacht Leiden, Universiteit Leiden, Niels Bohrweg 2, 2300 RA, Leiden, the Netherlands*
[2] *JILA, University of Colorado, Boulder, CO 80309-0440, USA*
[3] *Department of Astrophysical and Planetary Sciences, University of Colorado, Boulder, CO 80309-0391, USA*

ABSTRACT
Accretion discs at sub-pc distances around supermassive black holes are likely to cool rapidly enough that self-gravity results in fragmentation. Here, we use high-resolution hydrodynamic simulations of a simplified disc model to study how the outcome of fragmentation depends upon numerical resolution and cooling time, and to investigate the incidence of binary formation within fragmenting discs. We investigate a range of cooling times, from the relatively long cooling time-scales that are marginally unstable to fragmentation down to highly unstable cooling on a time-scale that is shorter than the local dynamical time. The characteristic mass of fragments decreases with reduced cooling time, though the effect is modest and dependent upon details of how rapidly bound clumps radiate. We observe a high incidence of capture binaries, though we are unable to determine their final orbits or probability of survival. The results suggest that faster cooling in the parent disc results in an increased binary fraction, and that a high primordial binary fraction may result from disc fragmentation. We discuss our results in terms of the young massive stars close to the Galactic Centre, and suggest that observations of some stellar binaries close to the Galactic Centre remain consistent with formation in a fragmenting accretion disc.

Key words: accretion, accretion discs – instabilities – hydrodynamics – methods: numerical – stars: formation – Galaxy: centre

1 INTRODUCTION

In recent years the topic of gravitational instabilities in accretion discs has been the subject of a great deal of theoretical and numerical research. Such gravitational instabilities have applications both to accretion disc theory, in terms of the transport of angular momentum, and also to theories of star and planet formation, as under certain circumstances the instability can lead to fragmentation (e.g. Durisen et al. 2007; Lodato 2007). In particular, the fragmentation of accretion discs has been invoked to explain both the formation of gas giant planets (e.g. Boss 1997; Mayer et al. 2002), and also to explain the formation of young, massive stars observed close to super-massive black holes at the centres of galaxies (e.g., Polyachenko & Shukman 1977; Levin & Beloborodov 2003).

In order for a gravitationally unstable disc to fragment, two conditions must be met. The first is that the disc must be sufficiently cold, or sufficiently massive, to satisfy the Toomre (1964) criterion

$$Q = \frac{c_{\mathrm{s}}\kappa}{\pi G \Sigma} \lesssim 1 \qquad (1)$$

Here c_{s} is the local sound speed, κ the epicyclic frequency (which in a thin disc is approximately equal to the local orbital frequency),

and Σ the disc surface density. In addition, in order to fragment the disc must cool sufficiently rapidly to overcome compressional heating during collapse. Gammie (2001) used a local analysis, and found the fragmentation boundary to be

$$t_{\mathrm{cool}} < \beta_{\mathrm{crit}}\Omega^{-1}, \qquad (2)$$

where t_{cool} is the local cooling time-scale, Ω is the orbital frequency, and $\beta_{\mathrm{crit}} \simeq 3$. Subsequent studies using global simulations have verified the validity of this criterion (Rice et al. 2003, 2005; Mejía et al. 2005; Boley et al. 2007), although the exact value of β_{crit} required for fragmentation can be larger by a factor of 2–3 depending on the details of the thermal physics adopted (cooling law, equation of state, etc.). To date, however, most work has either made use of simple thermal physics (e.g. Rice et al. 2003, 2005; Nayakshin et al. 2007), or used more realistic thermal physics in the specific case of self-gravitating protoplanetary discs (e.g. Cai et al. 2007; Stamatellos & Whitworth 2008). More generally, most studies of fragmentation have tended to look at cases of critical cooling (with $\Omega t_{\mathrm{cool}} \simeq \beta_{\mathrm{crit}}$), despite the fact that some discs, notably those around black holes, are thought to cool much more rapidly (e.g. Bertin & Lodato 2001; Goodman 2003).

In this paper we use numerical hydrodynamics to study the influence of rapid cooling on the fragmentation properties of grav-

itationally unstable accretion discs. In Section 2 we present the details of our numerical simulations, in which we adopt a cooling law consistent with radiative cooling in an optically thick disc. We use this model to study the non-linear evolution of the instability over a wide range in cooling times. We also consider in detail the effects of numerical resolution on our calculations, varying the number of fluid elements by a factor of 64 between our lowest- and highest-resolution calculations. We present our results in Section 3, first considering issues of numerical convergence and then addressing the effects of varying the cooling time. In Section 4 we discuss the limitations of our analysis, and the implications of our results for the formation of stars near to the Galactic Centre, before summarizing our conclusions in Section 5.

2   NUMERICAL METHOD

The simulations presented in this paper use the publicly-available smoothed-particle hydrodynamics (SPH) code GADGET2 (Springel 2005). We have modified the code to include a simple cooling prescription (discussed below), which is representative of the cooling expected in a real accretion disc. We adopt the standard Monaghan & Gingold (1983) prescription for the SPH viscosity (with $\alpha_{\rm SPH} = 1.0$), using the "Balsara-switch" (Balsara 1995) to limit the artificial shear viscosity (as specified in Equations 11–12 of Springel 2005). We allow for a variable gravitational softening length, and fix the SPH smoothing and gravitational softening lengths to be equal throughout. We use the standard Barnes-Hut formalism to compute the gravitational force tree (as described in Springel 2005), and use $N_{\rm ngb} = 64 \pm 2$ as the number of SPH neighbours. Typically, a handful of SPH particles end up in rather close orbits around the central gravitating mass and, if left unchecked, result in unreasonably short time-steps. Consequently we use a single sink particle for the central gravitating mass. The sink particle simply accretes all gas particles that pass within its radius (as described in Cuadra et al. 2006), and we set the sink radius to be 1/4 of the inner disk radius. This is simply a numerical convenience adopted in order to save computing time, and has no physical effect on the simulations. The simulations were run on the TUNGSTEN Xeon Linux cluster at NCSA[1], using 64 parallel CPUs for the highest-resolution runs.

Numerical simulations of gas disc dynamics suffer from a basic computational limitation, namely that the pressure scale-height of the disc acts as a minimum length-scale that must always be well-resolved. Requirements on computing time are therefore very strongly dependent on the disc aspect ratio (the required time typically scales as $(H/R)^{-3}$ or steeper), and accurate global simulations of thin discs are very computationally expensive. In a self-gravitating disc with $Q \simeq 1$, the disc aspect ratio is approximately equal to the ratio of the disc mass to the central object mass, and to date most numerical simulations have focussed on relatively massive, thick discs (with $M_{\rm disc} \sim 0.1 M_*$, e.g., Rice et al. 2003, 2005; Boley et al. 2007). This is typical of young protoplanetary discs, but discs around black holes may in fact be much thinner ($H/R < 10^{-2}$). We therefore seek a method of simulating the behaviour expected in a very thin disc, but without incurring the severe computational penalty that usually results.

We make use of an artificial disc model, which behaves like a thin self-gravitating disc while in fact remaining thick enough to be

well-resolved even at moderate resolution. Spiral density waves in self-gravitating discs tend to be dominated by a relatively narrow range in spiral mode number, with progressively thinner discs resulting in ever higher-order spiral waves (see the review by Lodato 2007). Moreover, as the disc thickness is reduced, the "local approximation" (that global modes of the instability are negligible) becomes more accurate. Consequently, we adopt a relatively large disc mass, but impose an artificial radial cooling profile that results in the disc fragmenting in a single, near-circular filament. The analysis of our simulations considers only this filament, with the rest of the disc regarded as a boundary condition: the dynamics of fragmentation in a single, near-circular filament are a good approximation to a real self-gravitating disc in the thin disc limit. With this set-up we are able to resolve the disc scale-height properly with only moderate numbers of SPH particles (see Section 2.3 below), and can thus run simulations with much larger particle numbers ("over-resolving" the disc thickness) to study the fragmentation process in detail when the cooling time is short.

2.1   Initial Conditions

We adopt initial conditions similar to those that have been adopted in previous studies of disc fragmentation (e.g. Rice et al. 2003, 2005; Alexander et al. 2008). We adopt a surface density profile

$$\Sigma(r) \propto r^{-1} \tag{3}$$

and a disk temperature that scales as

$$T(r) \propto r^{-1/2} \,, \tag{4}$$

where r is the (cylindrical) radius. The disc sound speed $c_s \propto T^{1/2}$ is normalized so that the Toomre parameter (Equation 1) is equal to 2.0 at the outer boundary. The disc is initially vertically isothermal, so the vertical structure of the disc is given by

$$\rho(r, z) = \frac{\Sigma(r)}{\sqrt{2\pi}H} \exp\left(-\frac{z^2}{2H^2}\right) \,, \tag{5}$$

where $H = c_{\rm s}/\Omega$ is the disc scale-height (here $\Omega = \sqrt{GM_*/r^3}$ is the Keplerian orbital frequency.). With this set-up the initial disc is marginally gravitationally stable, and the instability is allowed to develop in a physical manner as the disc cools. We define a disc mass of $M_{\rm d} = 0.1 M_*$ (where M_* is the mass of the central gravitating mass), and model the disc using $N_{\rm SPH}$ SPH particles of mass $m = M_{\rm d}/N_{\rm SPH}$. We adopt a system of units where $M_* = 1$, the inner edge of the disc is at $r = 1$, and the time unit is the orbital period at $r = 1$. (Consequently, $G = 4\pi^2$ in code units.) The initial distribution of the SPH particles is obtained by randomly sampling the disc density profile described above, with each particle given a circular velocity in the $x - y$ plane. We adopt an adiabatic equation of state throughout, with adiabatic index $\gamma = 7/5$ (as expected for a disc consisting primarily of molecular hydrogen).

2.2   Cooling

The principle aim of this study is to investigate the effect of variable cooling rates on the fragmentation properties. At each time-step, we allow the internal energy of the ith particle, u_i, to cool as

$$\left(\frac{du}{dt}\right)_{{\rm cool},i} = -\frac{u_i}{t_{\rm cool}} \,. \tag{6}$$

The cooling time-scale $t_{\rm cool}$ is defined in terms of the dimensionless parameter β and the orbital frequency

---

$$t_{\text{cool}} = \frac{\beta}{\Omega}. \tag{7}$$

As mentioned in Section 1, previous studies have found that the fragmentation boundary lies in the range $\beta_{\text{crit}} \simeq 3\text{–}7.5$, with a weak dependence on the adopted equation of state (e.g. Gammie 2001; Rice et al. 2003, 2005). Here, in order to study the effects of variable cooling, we define β to be a function of radius, density and time:

$$\beta = \mathcal{R}(r)D(\rho)T(t). \tag{8}$$

As discussed above, the radial dependence of β is chosen to mimic the behaviour of a tightly-wound spiral density pattern (as expected in a thin disc; e.g. Lodato 2007), but without the excessive demands on computing resources that usually arise when simulating very thin discs. We choose $\mathcal{R}$ so that the disc fragments in a ring at $r = 2$, and can therefore save computing power by considering a limited range in radius. The only requirement on the radial range is that the boundaries of the disc do not influence the fragmentation, so we fix the inner boundary at $r = 1$ and the outer boundary at $r = 3$. The radial dependence of the cooling time is chosen to be a Gaussian profile centred on $r_0 = 2$

$$\mathcal{R}(r) = 8.0 \times \left[1 - \frac{1}{2} \exp\left(-\frac{(r - r_0)^2}{2\Delta^2} \right) \right], \tag{9}$$

with $\Delta = 0.25$. Thus $\mathcal{R}(r)$ varies from 4–8, with the minimum value at $r = r_0$. The normalisation parameters are chosen so that the disc is initially marginally stable to fragmentation everywhere except close to $r = r_0$, and the value of Δ is chosen so that the radial length-scale over which the cooling time varies is a few times larger than the pressure scale-height of the disc.

In addition, we allow the cooling time to vary with the local density. In a real self-gravitating disc we expect radiative cooling to be optically thick (see the discussion in Section 4), so the cooling rate is density-dependent. If cooling is optically thick then the radiative cooling time-scale is proportional to the photon diffusion time-scale. In the case of constant opacity κ, the time, t_{D}, for a photon to diffuse a length L is given by

$$t_{\text{D}} = \frac{L^2}{c} \rho \kappa, \tag{10}$$

where c is the speed of light. In a collapsing (thinning) disc the appropriate length-scale is the disc thickness h, and the density $\rho \simeq \Sigma/h$. Thus, for 2-D collapse we have

$$t_{\text{D}} \propto h \propto \rho^{-1}, \tag{11}$$

and we see that the cooling time-scale decreases with increasing density ρ. By contrast, in a collapsing spherical clump of mass m and radius h, we have $\rho \simeq m/h^3$, so

$$t_{\text{D}} \propto h^{-1} \propto \rho^{1/3}, \tag{12}$$

and therefore the cooling time-scale increases (i.e. cooling becomes less efficient) as the clump becomes more dense. This increase in the cooling time-scale can lead to significant suppression of cooling as the forming fragments reach high density, and consequently we adopt a density-dependence with the following form:

$$D(\rho) = \max\left[1, \left(\frac{\rho}{\rho_0} \right)^{1/3} \right]. \tag{13}$$

Thus for densities below ρ_0 the cooling time is independent of the density; at higher densities the cooling time increases as $\rho^{1/3}$. Empirically, we choose $\rho_0 = 1.0$ in code units (central object mass /

N_{SPH}	τ	t_{frag}	$\beta'(t_{\text{frag}})$
200,000	∞	8.5	4.0
1,600,000	∞	8.2	4.0
12,800,000	∞	8.1	4.0
200,000	5	6.6	2.1
1,600,000	5	6.4	2.2
200,000	3	5.7	1.3
1,600,000	3	5.7	1.3
200,000	2	5.0	0.5
1,600,000	2	4.8	0.6
12,800,000	2	4.8	0.6

Table 1. List of simulations run, showing resolution and physical parameters for each simulation. t_{frag} denotes the time (in code units) when the first fragment forms, and $\beta'(t_{\text{frag}})$ indicates the approximate value of the local cooling parameter $\beta'(t_{\text{frag}}) = \mathcal{R}(r_0)T(t_{\text{frag}})$ (i.e. neglecting the density dependence) when fragmentation occurs.

inner disc radius[3]), as this is typical of the local density in our simulations at the point where bound, roughly spherical objects begin to form. In practice the weak dependence of the cooling time on density means that our results do not depend significantly on the exact value of ρ_0.

Lastly, in order to study the effect of very rapid cooling we allow the cooling rate to vary with time in the simulations. The motivation for this is that one cannot simply begin a numerical simulation with a cooling time that is much shorter than the dynamical time-scale, as the resulting collapse and fragmentation merely amplifies noise in the initial conditions and has little or no physical meaning. Instead, one must allow the instability to develop in a physical manner before the cooling time is allowed to become short. We follow the approach of Clarke et al. (2007), and allow the cooling time to decrease as

$$T(t) = 1 - \left(\frac{1}{\tau} \frac{t}{t_{\text{orb}}(r_0)} \right) \tag{14}$$

Here t_{orb} is the local orbital period, and τ is a dimensionless parameter that we use to parametrize the cooling rate. We note that the form of $T(t)$ allows the cooling time to become very small (and indeed negative) at large t; in practice, however, our simulation run for short enough times that this is not a problem. We adopt values of the cooling parameter $\tau = 2, 3, 5, \infty$ (in the latter case the cooling time is constant with time). With this set-up the disc is marginally unstable to fragmentation at $r = 2$ for $\tau = \infty$, while smaller values of τ result in cooling times that are up to order of magnitude shorter. At the end of our simulations the smallest values of β are in the range 0.2–0.3 (see also Table 1).

2.3 Simulations

The issue of numerical resolution is critical when considering calculations of fragmentation, as insufficient resolution can result in spurious suppression or amplification of fragmentation. We have conducted a number of calculations using a range of numbers of SPH particles for different values of τ (see Table 1). Resolution requirements for SPH simulations of fragmentation were presented and tested by Bate & Burkert (1997), and extended for the specific case of self-gravitating discs by Nelson (2006). In this case the Jeans mass must be resolved into $\gtrsim 10N_{\text{ngb}}$ particles, and the disc scale-height H must be resolved into at least 4 SPH smooth-

ing lengths in the saturated state of the instability[2]. For most of the models considered here the second of these is the more stringent condition, but correctly resolving the Jeans mass becomes more demanding in the simulations with $\tau = 2$ (see Section 3.1). We choose our lowest resolution so that we marginally satisfy these conditions for the case $\tau = \infty$, and thus choose $N_{\rm SPH} = 200,000$. We then move to higher resolution, first increasing the particle number by a factor of 8 (to $N_{\rm SPH} = 1.6 \times 10^6$), and then by a further factor of 8 (to $N_{\rm SPH} = 1.28 \times 10^7$). In a perfectly spherical configuration increasing $N_{\rm SPH}$ by a factor of 8 results in a factor of 2 ($= 8^{1/3}$) increase in the spatial resolution. However, in the disc geometry considered here we find that a factor of 8 increase in $N_{\rm SPH}$ in fact improves the spatial resolution by a factor of approximately 2.5. For convenience, we subsequently refer to the simulations with $N_{\rm SPH} = 200,000$ as "low resolution" (LR), $N_{\rm SPH} = 1.6 \times 10^6$ as "medium resolution" (MR), and $N_{\rm SPH} = 1.28 \times 10^7$ as "high resolution" (HR). For reasons of computation time we ran the HR simulations only for the extreme cases $\tau = \infty$ and $\tau = 2$; the LR and MR runs simulations were conducted across the full range in τ.

An additional benefit of moving to resolutions higher than the minimum requirement is the potential to resolve binary and multiple systems. Roughly speaking, we expect that individual objects will only be able to fragment into multiple systems if they cool on a time-scale shorter than the local dynamical time-scale. In a self-gravitating disc with $Q \simeq 1$ the dynamical time-scale of a single fragment is, to first order, equal to the local orbital time-scale. Consequently, we do not expect the formation of "fission" binaries in discs with $\beta > 1$ (although binaries may still form by dynamical capture events), as thermal pressure acts to smooth out density fluctuations on a time-scale shorter than the cooling time-scale. Thus the minimum resolution requirement is sufficient in simulations with $\beta > 1$. However, with faster cooling it may be possible to form binary and multiple systems via the dynamical fission of single clumps. Consequently our high-resolution simulations act not only as convergence tests, but also provide a means of studying the small-scale fragmentation that may occur in discs with short cooling times.

In order to compare the fragmentation in different simulations we must first identify individual fragments ("clumps") in our simulations, and we do this simply by adopting a density threshold of $\rho = 10.0$ (i.e. ten times larger than the density above which the cooling time begins to slow). We identify clumps as objects which have peak densities above this threshold, and estimate clump masses by moving outward in spherical shells one (minimum) SPH smoothing length in radius, until the mean density drops below the threshold value: all particles within this radius are then deemed to be part of the clump.

One consequence of using a fixed density threshold in this manner is that clumps are identified at somewhat earlier times in simulations with higher resolution, as the central regions of individual clumps are no longer "smeared out" over as long a length-scale. This is an artefact of our analysis, rather than a defect in the simulations, and we solve this simply by comparing simulations at different resolutions relative to the time at which the first clump is identified. In practice this difference is very short: the delay in t between the identification of the first fragment at the highest and
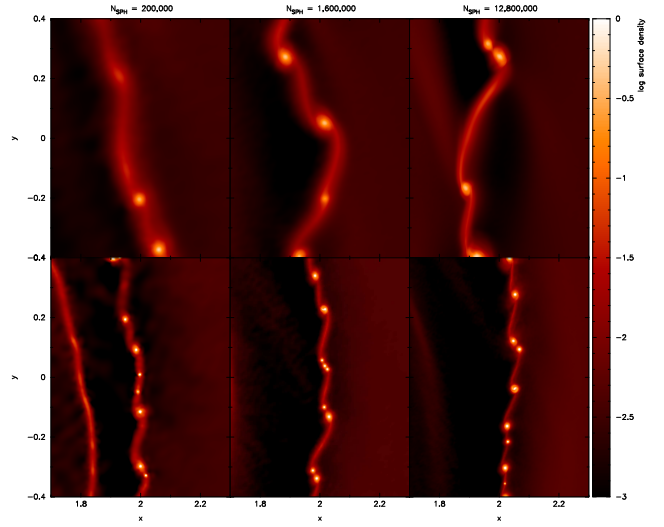
Figure 1. Plots of the disc surface density in the simulations with the slowest ($\tau = \infty$; top row) and fastest ($\tau = 2$; bottom row) cooling times, using different numerical resolutions. From left-to-right, the three columns show the LR, MR and HR simulations respectively, and each plot is made 1/8 of a local orbital period (0.35 time units) after the first fragment forms. For slow cooling, just below the fragmentation threshold, we see that the fragmentation process is well-resolved in all our simulations, although more substructure is apparent in the higher-resolution simulations. With fast cooling the disc is colder when it fragments, and higher resolution is needed in order to resolve the disc scale-height: in this case we see that the LR simulation is somewhat under-resolved (e.g. the density peaks at $r \simeq 1.8$ are numerical artefacts), but that $N_{\rm SPH} = 1.6 \times 10^6$ appears sufficient to resolve the fragmentation process properly.

lowest resolutions is a few percent of the total simulation time (see Table 1). Overall this is a rather crude method for measuring the clump mass function, and a more sophisticated approach could, for example, allow for non-spherical clumps, or test every particle in order to determine how many individual particles are bound to each density peak (e.g. Alexander et al. 2008). In practice, however, such "improvements" create at least as many problems as they solve, because the treatment of, for example, binary and multiple systems, or "circumstellar" discs, becomes ambiguous. The method used here is robust across different resolutions, and is sufficiently accurate for our needs: *a posteriori* checks show that negligible fraction of clumps identified in this manner are not in fact bound objects

3   RESULTS

3.1   Numerical convergence

We first seek to compare our results at difference numerical resolutions, in order to test convergence in our simulations. The non-linear development of gravitational instabilities essentially amplifies seed fluctuations in the density field, and it is not possible to construct identical noise fields at different particle numbers. Consequently we instead seek statistical convergence, measuring when the first bound objects appear, how many such objects we find, and how much mass is present in these objects. We identify clumps using the method described in Section 2.3, and use the mass functions derived in this manner to compare different simulations.

The results of our numerical convergence tests are shows in

---

[2] Note that the definition of the SPH smoothing length in Springel (2005) differs from that in most of the SPH literature, being larger by a factor of 2. Throughout this paper we discuss only the standard definition of the smoothing length (e.g. Monaghan 1992).
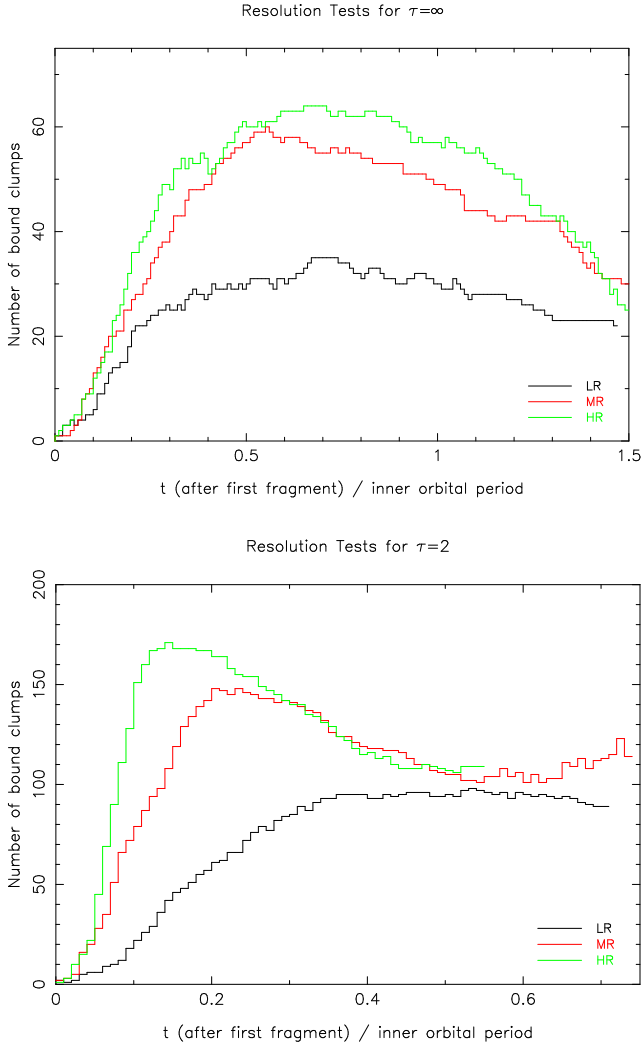
Figure 2. Results of numerical convergence tests. The upper panel shows the evolution of the number of clumps for the simulations with $\tau = \infty$. The origin of the time axis is the time when the first clump is identified, and the LR, MR and HR simulations are denoted by black, red and green lines respectively. Although more binaries are resolved at higher resolution the total mass in clumps does not vary significantly between the runs, indicating a good degree of numerical convergence. The lower panel shows the same plot for the simulations with $\tau = 2$. In this case the LR simulation is rather under-resolved, but good convergence is found between the MR and HR simulations.

Figs.1 and 2[3]. In the case of slow cooling ($\tau = \infty$) we see that the fragmentation process is well-resolved in all of our simulations. In terms of the criteria laid out by Bate & Burkert (1997) and Nelson (2006), we resolve the disc scale-height at the mid-plane into approximately 5.3 SPH smoothing lengths in the LR run, 12.5 smoothing lengths in the MR run, and 29.6 smoothing lengths in the HR run. In addition the characteristic fragment mass is well-resolved throughout so, as noted in Section 2.3, the LR simulation marginally satisfies the minimum resolution requirement. Runs with higher resolution reveal more substructure, resolving binaries and "circumstellar" discs, but fragmentation is independent of resolution. Subsequent interactions between clumps, however, are not

independent of resolution: this is clearly seen in Fig.2, where the number of clumps beyond $\Delta t \simeq 0.3$ differs significantly between the LR and HR runs. This is due to mergers of binary clumps, and indicates that the interactions between such binaries are not properly resolved. This occurs because although the global disc is well-resolved the many "circumstellar" discs are not, so the time-scale on which binary pairs merge is governed by the rate of numerical dissipation of energy in these "sub-discs". Consequently we do not attach great significance to our results beyond the time at which binaries start to merge. Before clump mergers start to become significant the total mass in clumps varies by less than 20% between the LR and HR runs, indicating that good numerical convergence has been achieved.

In the case of fast cooling ($\tau = 2$) the disc is colder when it fragments, so higher resolution is needed in order to resolve the Jeans mass. Here, the pressure scale-height of the disc is again marginally resolved in the LR run (into 4.5 SPH smoothing lengths at the disc mid-plane), but the characteristic fragment mass is $\simeq 4N_{\rm ngb}m$. Thus in this case, as expected, we see that the LR simulation is slightly under-resolved (by a factor of approximately 2 in particle number), but both resolution requirements are satisfied comfortably in the MR and HR runs. Binary mergers are somewhat more significant in the MR simulation than in the HR simulation (again due to under-resolving "circumstellar" discs), but otherwise there is good agreement between the MR and HR runs. In both the $\tau = \infty$ and $\tau = 2$ cases good statistical convergence is found between the MR and HR runs, so we conclude that $N_{\rm SPH} = 1.6 \times 10^6$ is sufficient to resolve the fragmentation process properly throughout all our simulations.

3.2 Cooling

In Section 3.1 we found that the MR runs were sufficient to resolve the fragmentation process properly across the full range of cooling times. We now compare the MR runs across the full range of cooling times, in order to study the effect of the cooling rate on the fragmentation process. The results of the MR runs are shown in Figs.3–5. Fig.3 shows the evolution of the disc surface density in the four different simulations, and it is clear that the cooling rate has a dramatic effect on the fragmentation process. Slow cooling leads to the formation of larger, more massive fragments with larger separations, while faster cooling results in a larger number of less massive fragments, with smaller separations. Faster cooling also leads to the formation of more binary and multiple systems, although numerical dissipation results in many of these systems merging rapidly in our simulations. Additionally, the fragmentation process appears to be more dynamic in the case of slow cooling, with individual objects showing greater perturbations from circular orbits. This occurs because in the case of slow cooling the collapsing filament has time to become unstable to bending modes before it fragments, while with faster cooling the filament fragments before it becomes significantly bent.

This qualitative assessment is seen more quantitatively in Fig.4, which shows the number and total mass of the bound fragments in each simulation. Although mass is accreted into fragments more rapidly when the cooling is fast, the total mass in fragments does not differ significantly with different cooling times. However, the simulations with faster cooling produce many more fragments, with correspondingly lower masses, and the incidence of multiple systems is greatly increased.

Fig.5 shows how the clump mass function (MF) varies as a function of the local cooling time. We compare the MFs at a point
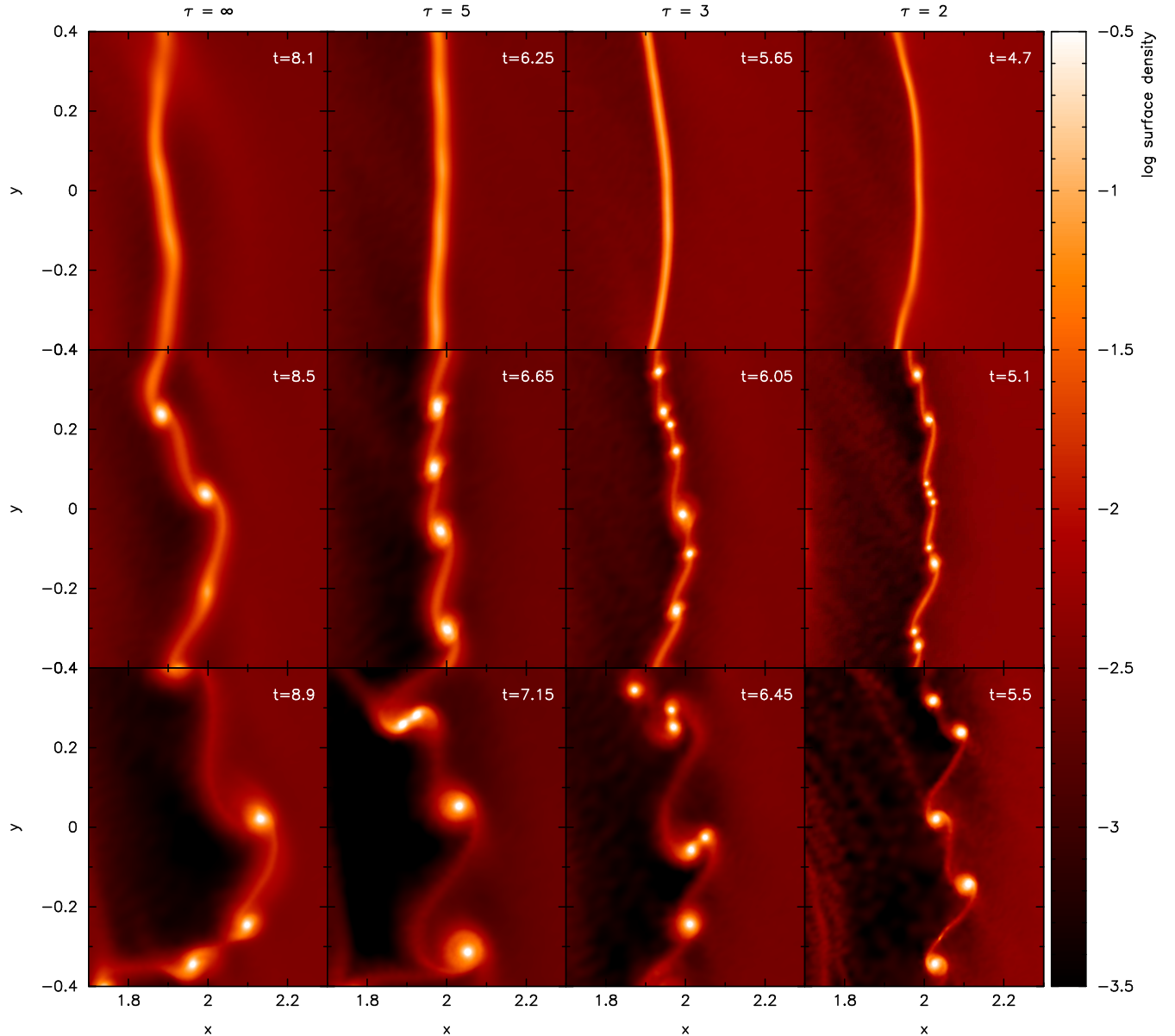
[3] Visualisations of our SPH simulations were created using SPLASH: see Price (2007) for details.

Figure 3. Plots of the disc surface density in the MR simulations, showing how fragmentation develops as a function of cooling time. From left-to-right, the four columns show the evolution of the simulations with $\tau = \infty$, 5, 3 and 2 respectively. The plots are made in a reference frame that is co-rotating with the disc at $r = 2$, so that only relative motions are evident in the time evolution. (In this frame the disc shear flow appears to move upwards for $x < 2$ and downwards for $x > 2$.) It is clear from the snapshots that faster cooling results in the formation of more, smaller fragments, as well as a higher incidence of multiple systems. Note also the large number of binary mergers seen between the second and third snapshots in the simulations with fast cooling.

where an approximately equal fraction of the disc mass is bound into clumps in each simulation, and before the (under-resolved) binary mergers begin to dominate the simulations with the fastest cooling. We therefore compare the MFs at 0.7 local cooling times after the formation of the first fragment (see Fig.4), although we note that the MF comparison is qualitatively similar throughout the simulations. All of the MFs are reasonably well-fit by log-normal functions, in agreement with previous studies of disc fragmentation (e.g. Nayakshin et al. 2007; Alexander et al. 2008). We find that faster cooling results in a smaller characteristic fragment mass. This is in agreement with the previous study of Nayakshin et al. (2007), which also found that the characteristic fragment mass decreased as the cooling rate increased. In our simulations the local (density-independent) cooling time when fragments form varies by

a factor of $\simeq 7$ between the slowest- ($\tau = \infty$) and fastest-cooling ($\tau = 2$) models, and the characteristic mass of the respective MFs differs by a factor of $\simeq 4$. We also find that the MFs become somewhat broader with shorter cooling times, but given the relatively small numbers of clumps formed in the $\tau = \infty$ and $\tau = 5$ runs this result is not statistically significant. The range in characteristic masses we find is much smaller than that seen by Nayakshin et al. (2007), who found that the characteristic mass differed by a factor of $\simeq 100$ between their simulations with $\beta = 3$ and $\beta = 0.3$. We attribute this to the differences between the cooling laws adopted here and those used by Nayakshin et al. (2007). In particular, we note that the density dependence of our cooling law (which suppresses cooling at high densities) was not included in the simulations of Nayakshin et al. (2007), and the likely consequence of

this omission is to enhance fragmentation as the density in clumps increases. We therefore suggest that the variation in characteristic mass with cooling time in real discs is less pronounced than that found by Nayakshin et al. (2007), and conclude that large variations in the cooling time-scale result in only modest variations in the clump MF.

As mentioned above, a further difference between the fast and slow cooling simulations is in the fraction of multiple and binary fragments that are formed. The binary fraction is difficult to quantify due to the rapid merging of the binaries that form, but the differences are clearly seen in Fig.3. With critical cooling ($\tau = \infty$) most objects form as singles, and although some binaries form due to dynamical capture events the overall binary fraction is low. However, when the cooling is faster the disc is thinner (colder) when it fragments, and consequently the clumps are more closely-spaced when they first form. This leads to many more capture events, and we see many more binaries and multiples in the simulations with faster cooling. Indeed, in the simulations with $\tau = 3$ and $\tau = 2$ almost all the clumps appear in binary or multiple systems. We see no fragmentation of single clumps after the initial fragmentation occurs, and all of the binary and multiple systems observed form via dynamical capture events (in which the disc gas acts as a "third body"). We attribute the absence of any "fission" binaries to the density dependence of the cooling time, as the increase in β with increasing clump density means that the value of β in the high-density clumps rarely, if ever, falls below unity. In this case (as explained in Section 2.3) we do not expect to see sub-clump fragmentation, and it seems that the formation of binaries by fission is unlikely in real discs with optically thick cooling. The high incidence of capture binaries can easily be understood, as faster cooling leads to a lower temperature, and therefore a smaller disc scale-height H, in the fragmenting disc. The length-scale that is most unstable to gravitational instabilities is $\simeq H$, so the typical fragment mass is $\simeq \pi \Sigma H^2$ and the typical fragment separation is $\simeq H$. Faster cooling results in a smaller value of H, which in turn results in smaller fragment masses and smaller separations between fragments. In our simulations these binaries tend to merge rapidly due to the high rate of numerical dissipation in their "circumstellar" discs (as seen between $t = 5.1$ and $t = 5.5$ in the $\tau = 2$ simulation in Fig.3), but in reality such mergers would likely be much less efficient. Consequently, we find that fast cooling in fragmenting self-gravitating discs leads to a slight decrease in the characteristic fragment mass, and a dramatic increase in the incidence of binary and multiple objects.

4 DISCUSSION

Our simulations are robust and well-understood from a numerical perspective, but concerns remain when we consider the physical interpretation of our results. The most obviously unphysical aspect of our simulations is our "thin disc approximation", using an artificial cooling law to allow fragmentation characteristic of a thin disc to be modelled at moderate numerical resolution. It would be preferable to construct global models of thin discs, but this is not computationally feasible with current techniques, especially in the case of short cooling times. Our disc model is effectively a local approximation, so the primary concern regards the boundary conditions. In our models the real physical boundaries (at $r = 1$ and $r = 3$) are sufficiently far from the region of interest that they have no influence on the results, and the radial variation of the cooling time is negligible in the fragmentation region. However, in a real
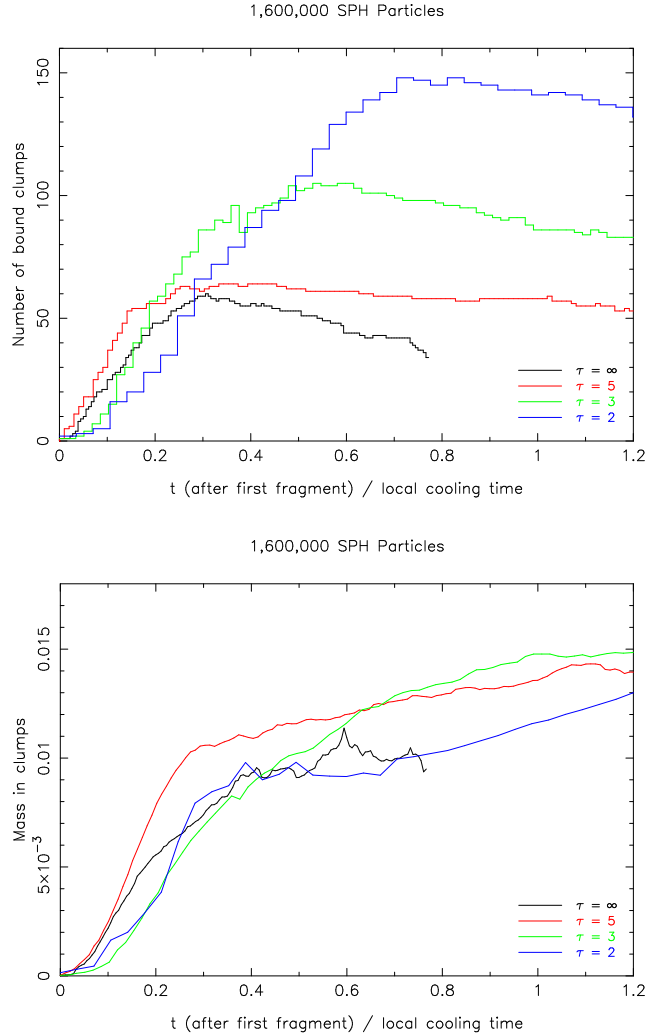

Figure 4. Fragmentation as a function of cooling time. The upper panel shows the number of clumps as a function of time in each of the four MR simulations, while the lower panel shows the total mass in the clumps. The origin of the time axis is set at the point where the first fragment forms, and time is plotted in units of the local cooling time (i.e. $\beta'(t_{\text{frag}})/\Omega(r_0)$, with the values of $\beta'(t_{\text{frag}})$ taken from Table 1). The total mass in clumps is independent of the cooling time, but we see that faster cooling results in the formation of more objects (with lower masses).

thin disc there are many such filaments, and the typical radial separation between filaments can be as little as a few times H. Dynamical interactions between adjacent filaments seem likely, especially once bound, point-mass-like objects have formed, and the likely outcome of this increased rate of clump-clump interactions is the formation of even more capture binaries than seen in our simulations. In some of our simulations (those with slower cooling) some of the bound clumps undergo significant radial scattering even in the absence of adjacent filaments (moving as far as $\Delta r \simeq 0.3$ from their initial location), and consequently their cooling times change in an unphysical way. However, this only occurs at late times in the simulations (to which we do not attach great significance), and moreover in such dense clumps the dominant term in the cooling time is the density dependence. Consequently we are satisfied that our artificial cooling model is robust, and although our simulations should be regarded as numerical experiments rather than
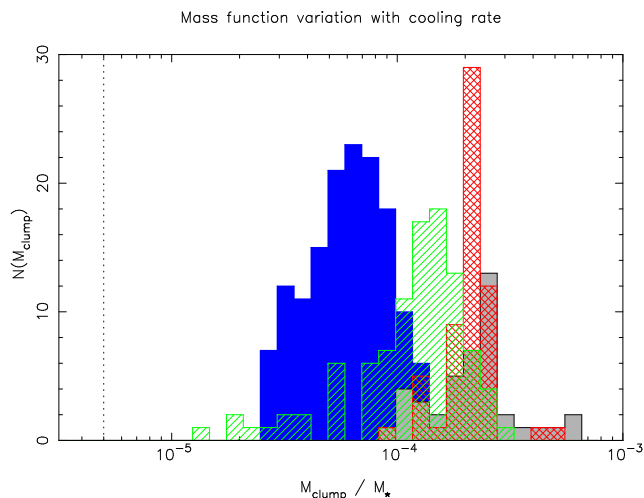
Mass function variation with cooling rate

Figure 5. Mass functions of clumps in simulations with different cooling times. The mass functions were computed 0.7 local cooling times after the formation of the first fragment (i.e. at 0.7 on the time axis in Fig.4). The black/grey, red, green and blue MFs denote the simulations with $\tau = \infty$, 5, 3 and 2 respectively. The dotted line denotes the approximate (mass) resolution limit of $10N_{ngb}m$; all of our clumps are above this limit by at least a factor of 5–10. We see clearly that faster cooling results in a smaller characteristic mass, and a somewhat broader mass function.

physical simulations, we are confident attaching limited physical significance to our results.

An additional concern regards the density dependence of our cooling law. Fixing the cooling time to scale as $\rho^{1/3}$ is correct only if the opacity is constant and the disc is optically thick, but it is not clear that these assumptions hold in reality. In real self-gravitating discs the temperatures are expected to be ~ 100K, so dust grains dominate the disc opacity (e.g. Bell & Lin 1994; Semenov et al. 2003). In the case of a fragmenting disc with $Q = 1$, the disc is optically thick if the orbital period is $\lesssim 1000$yr (Levin 2007). Real discs are generally only thought to be self-gravitating in their outer regions, at radii $\gtrsim 50$AU in the case of protostellar discs (e.g. Stamatellos & Whitworth 2008) or at radii $\gtrsim 0.1$pc in the case of a disc around a $\sim 10^6 M_\odot$ black hole (e.g. King & Pringle 2007), and in both cases this suggests that fragmenting discs are marginally optically thick. Individual clumps (with higher densities) thus have optical depths much greater than unity, so the optically thick approximation appears valid across a wide range of disc scales. It is clear, however, that the opacity is not strictly constant in real discs. Dust opacity is essentially independent of density, but varies strongly as a function of temperature: the opacity scales as T^2 for $T \lesssim 160$K, and is approximately constant for 160K$\lesssim T \lesssim 1000$K (Bell & Lin 1994). It is reasonable to assume that the temperature will increase somewhat during collapse, leading to an increase in the dust opacity, so it seems likely that our adopted function $D(\rho)$ somewhat over-estimates the efficiency of radiative cooling. Thus cooling is likely slower than predicted by our simulations in real fragmenting clumps. Models of core-collapse during star formation predict only moderate increases in temperature, however, so the uncertainty due to the temperature dependence of dust opacity is likely small. We note in passing that we also neglect further complexities of the collapse process, such as changes in gas composition with increasing temperature (due to dust sublimation, H_2 dissociation, etc.): such issues are well beyond the scope of this work.

An additional complication is that is still not clear whether or not it is possible to form a disc with $\beta < 1$. A coherent disc

structure cannot be assembled in the presence of such rapid cooling, as the growing disc would simply fragment on a dynamical time-scale. Various authors have argued in favour of rapid cooling in self-gravitating discs around massive black holes (e.g. Bertin & Lodato 2001; Goodman 2003), but these arguments generally rely on the assumption of roughly steady-state conditions in the disc, rather than considering disc formation. Rafikov (2007) argued that convective cooling results in an cooling time of $\beta \simeq 0.1$, but again does not consider disc formation in detail. By contrast, Nayakshin (2006, see also Levin 2007) argues against the presence of rapid cooling in a fragmenting, star-forming disc around a massive black hole, instead arguing that the disc mass is accreted gradually. The cooling time therefore decreases slowly as the disc mass is accumulated, so fragmentation occurs at $\beta \simeq \beta_{crit}$. If, however, the disc was formed in a rapid "accretion event", it seems that fragmentation could occur with $\beta < 1$. In addition the thermodynamics of the disc formation process are not fully understood, and one can conceive of a scenario where the disc is assembled with a longer cooling time and then begins to cool faster as it evolves (e.g. Clarke et al. 2007). One possible mechanism appeals to the "opacity gap" at 10^3–10^4K: if a disc was assembled from warm, $\gtrsim 1500$K gas then the condensation of dust at ~ 1000–1500K would result in a dramatic increase in the cooling rate as the temperature dropped, potentially resulting in a cooling time-scale shorter than the dynamical time-scale. Our simple model is not able to address this issue in detail, but this problem provides fertile ground for future research.

Despite the fact that our simulations show good numerical convergence with increasing particle number, numerical resolution is still an issue that must be considered with care. In particular, as noted in Section 3.1, the dynamical interactions between different clumps are under-resolved even in our highest-resolution simulations. This is a common feature of hydrodynamic simulations which seek to model large dynamic ranges in length and density, and in our simulations this results in unphysically rapid mergers of binary and multiple systems (as seen, for example, in Fig.3). We treat this problem by simply stopping our calculations before such under-resolved mergers begin to dominate the simulations, and attach only limited significance to our results beyond the point at which the first such mergers begin to happen. [Note, however, that real mergers between fragments are expected to occur in such systems, and indeed have been invoked to explain the large discrepancy between the typical fragment mass and the observed stellar masses near the Galactic Centre (Levin 2007).] A more sophisticated approach could, for example, make use of "sink particles" (e.g. Bate et al. 1995), modified thermal physics (e.g. Jappsen et al. 2005), or sub-resolution physics (e.g. Nayakshin et al. 2007) in order to follow the calculations further forward in time, but in our models (and in particular in the simulations with fast cooling) it is not clear that such approaches would yield more accurate results. Here, a better approach may be to "re-simulate" a small region of our computational domain at much higher resolution in order to study the inter-clump dynamics in more detail, but this is beyond the scope of this paper.

4.1 Application to the Galactic Centre

We now consider the implications of our results for the formation of stars near the Galactic Centre (GC). Many young, massive stars are known to exist close to the black hole at the centre of the Galaxy (Ghez et al. 1998; Genzel et al. 2003; Ghez et al. 2005; Paumard et al. 2006), and many of the stars at $\simeq 0.1$pc are known to form

a coherent ring or disc (e.g. Genzel et al. 2003)[4]. A popular theory for the formation of these stars is fragmentation of an accretion disc due to gravitational instabilities (e.g. Levin & Beloborodov 2003; Goodman & Tan 2004; Nayakshin & Cuadra 2005), and recent hydrodynamic simulations have found that the properties of the stars formed in this manner are broadly consistent with the observed GC stellar disc (e.g. Nayakshin et al. 2007; Alexander et al. 2008). However, to date most of these simulations have used greatly simplified thermal physics, adopting scale-free cooling and/or simplified equations of state, and given the critical role of cooling in the development of gravitational instabilities this treatment is less than ideal.

Additionally, for numerical reasons most previous simulations of the formation of the GC stellar disc have made use of relatively slow cooling time-scales (with $\beta \simeq 3$), although Nayakshin et al. (2007) did run a single model with $\beta = 0.3$ (as discussed in Section 3.2). However, analytic estimates of the cooling times in black hole discs generally find that the cooling time should be short compared to the dynamical time (e.g. Goodman 2003; Rafikov 2007) although, as discussed above, how such systems form remains uncertain. Estimates of the cooling parameter in a fragmenting disc at the GC range from $\beta \sim 0.1$ (Rafikov 2007) to $\beta = \beta_{\rm crit} \simeq 3$ (Nayakshin 2006), which is approximately the range spanned by our simulations. As discussed above, we find that a dramatic decrease in the cooling time leads to only a modest decrease in the characteristic stellar mass, and given the dependence of the characteristic mass on other disc properties (notably the disc aspect ratio H/R), this is unlikely to have a significant effect on the MF of the GC stellar disc. However, our simulations indicate that fast cooling results in a very high primordial binary/multiple fraction (of order unity), and although binaries do form when cooling is slower, they are less common. We therefore suggest that, if the GC stellar disc formed by fragmentation of a gaseous accretion disc, a high binary fraction may be indicative of rapid cooling in the parent accretion disc.

It is not clear, however, if this prediction provides a means of distinguishing between different proposed scenarios for the formation of the GC stellar disc. The massive star population in the field is known to have a high multiplicity fraction, and the binary fraction for Wolf-Rayet stars in the solar neighbourhood is $\sim 40\%$ (van der Hucht 2001). Little is currently known about binary statistics in the GC stellar cluster, with neither observations nor dynamical arguments (e.g. Cuadra et al. 2008) providing real constraints. It is entirely possible (and in fact likely) that many unresolved close binaries near the GC may be masquerading as single objects in current data, but some binary systems have been observed. In particular IRS 16SW, an eclipsing binary with a period of $\simeq 19.5$ days, appears to be a contact binary consisting of two approximately equal-mass Wolf-Rayet stars with masses $\gtrsim 50 M_\odot$ (Martins et al. 2006; Peeples et al. 2007; Rafelski et al. 2007). Systems of this type appear to be a plausible outcome of fast cooling in a fragmenting, self-gravitating accretion disc, and we suggest that such systems should be common if the stellar disc at the GC formed via fragmentation of a gaseous accretion disc.

---

[4] Note that there is some controversy as to whether one or two discs of young stars exist at the GC (Paumard et al. 2006; Lu et al. 2006). For the purposes of this discussion this is unimportant: the existence of one disc of young stars is sufficient to motivate our arguments, and they are not significantly altered by the existence of a second disc.

5 SUMMARY

We have used numerical hydrodynamic simulations to investigate the role of the cooling rate in the fragmentation of gravitationally unstable accretion discs. Our simulations make use of a simple cooling law, designed to mimic the behaviour of radiative cooling in an optically thick, geometrically thin disc. We performed detailed numerical convergence tests, and are satisfied that our simulations are well-resolved. We find that decreasing the cooling time results in the formation of a larger number of fragments, but with a correspondingly smaller characteristic mass. We also observed a high incidence of binaries, all of which form by dynamical capture. Rapid cooling, on a time-scale shorter than the local dynamical time, results in a dramatic increase in the incidence of binary and multiple systems, although we are not able to determine the final properties of such systems. We have considered our results in the context of the young stellar disc(s) at the Galactic Centre, and find that - at the very least - observations of some stellar binaries close to the Galactic Centre do not rule out formation in a fragmenting accretion disc.

REFERENCES

Alexander, R.D., Armitage, P.J., Cuadra, J., Begelman, M.C., 2008, ApJ, 674, 927
Balsara, D.S., 1995, J. Comput. Phys., 121, 357
Bate, M.R., Burkert, A., 1997, MNRAS, 288, 1060
Bate, M.R., Bonnell, I.A., Price, N.M., 1995, MNRAS, 277, 362
Bell, K.R., Lin, D.N.C., 1994, ApJ, 427, 987
Bertin, G., Lodato, G., 2001, A&A, 370, 342
Boley, A.C., Hartquist, T.W., Durisen, R.H., Michael, S., 2007, ApJ, 656, L89
Boss, A.P., 1997, Science, 276, 1836
Cai, K., Durisen, R.H., Boley, A.C., Pickett, M.K., Meija, A.C., 2008, ApJ, 673, 1138
Clarke, C.J., Harper-Clark, E., Lodato, G., 2007, MNRAS, 381, 1543
Cuadra, J., Nayakshin, S., Springel, V., di Matteo, T., 2006, MNRAS, 366, 358
Cuadra, J., Armitage, P.J., Alexander, R.D., 2008, MNRAS (Letters), in press (arXiv:0804:3596)
Durisen, R.H., Boss, A.P., Mayer, L., Nelson, A.F., Quinn, T., Rice, W.K.M., 2007, in Reipurth, B., Jewitt, D., Keil, K., eds, *Protostars & Planets V*, Univ. Arizona Press, Tuscon, 607
Gammie, C.F., 2001, ApJ, 553, 174
Genzel, R., et al. 2003, ApJ, 594, 812

Ghez, A.M., Klein, B.L., Morris, M., Becklin, E.E. 1998, ApJ, 509, 678

Ghez, A.M., Salim, S., Hornstein, D., Tanner, A., Lu, J.R., Morris, M., Becklin, E.E., Duchêne, G. 2005, ApJ, 620, 744

Goodman, J., 2003, MNRAS, 339, 937

Goodman, J., Tan, J.C. 2004, ApJ, 608, 108

Jappsen, A.-K., Klessen, R.S., Larson, R.B., Li, Y., Mac Low, M.-M., 2005, A&A, 435, 611

Levin, Y., Beloborodov, A.M. 2003, ApJ, 590, L33

Levin, Y., 2007, MNRAS, 374, 515

Lodato, G., 2007, La Rivista del Nuovo Cimento, 30, 293

Lu, J., Ghez, A.M., Hornstein, S.D., Morris, M., Matthews, K., Thompson, D.J., Becklin, E.E., 2006, J. Phys. Conf. Series, 54, 279

King, A.R., Pringle, J.E., 2007, MNRAS, 377, L25

Martins F., et al., 2006, ApJL, 649, L103

Mayer, L., Quinn, T., Wadsley, J., Stadel, J., 2002, Science, 298, 1756

Mejía, A.C., Durisen, R.H., Pickett, M.K., Cai K., 2005, ApJ, 619, 1098

Monaghan, J.J., 1992, ARA&A, 30, 543

Monaghan, J.J., Gingold, R.A., 1983, J. Comput. Phys., 52, 374

Nayakshin, S., 2006, MNRAS, 372, 143

Nayakshin, S., Cuadra, J., 2005, A&A, 437, 437

Nayakshin, S., Cuadra, J., Springel, V., 2007, MNRAS, 379, 21

Nelson, A.F., 2006, MNRAS, 373, 1039

Paumard, T., et al. 2006, ApJ, 643, 1011

Peeples, M.S., Bonanos, A.Z., DePoy, D.L., Stanek, K. Z., Pepper, J., Pogge, .W., Pinsonneault, M. H., Sellgren, K., 2007, ApJL, 654, L61

Polyachenko, V.L., & Shukman, I.G. 1977, Soviet Ast. Letters, 3, 134

Price, D.J., 2007, PASA, 24, 159

Rafikov, R.R., 2007, ApJ, 662, 642

Rafelski, M., Ghez, A.M., Hornstein, S.D., Lu, J.R., Morris, M., 2007, ApJ, 659, 1241

Rice, W.K.M., Armitage, P.J., Bate, M.R., Bonnell, I.A., 2003, MNRAS, 339, 1025

Rice, W.K.M., Lodato, G., Armitage, P.J., 2005, MNRAS, 364, L56

Semenov, D., Henning, Th., Helling, Ch., Ilgner, M., Sedlmayr, E., 2003, A&A, 410, 611

Springel, V., 2005, MNRAS, 364, 1105

Stamatellos, D., Whitworth, A. P., 2008, A&A, 480, 879

Toomre, A., 1964, ApJ, 139, 1217

van der Hucht, K.A., 2001, NewAR, 45, 135