Direct collapse of exceptionally heavy black holes in the merger-driven scenario

Lorenz Zwick\textsuperscript{1⋆}, Lucio Mayer\textsuperscript{1}, Lionel Haemmerlé\textsuperscript{2}, Ralf S. Klessen\textsuperscript{3,4}

\textsuperscript{1}Center for Theoretical Astrophysics and Cosmology, Institute for Computational Science, University of Zurich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
\textsuperscript{2}Département d’Astronomie, Université de Genève, chemin des Maillettes 51, CH-1290 Versoix, Switzerland
\textsuperscript{3}Universität Heidelberg, Zentrum für Astronomie, Institut für Theoretische Astrophysik, Albert-Ueberle-Str 2, D-69120 Heidelberg, Germany
\textsuperscript{4}Universität Heidelberg, Interdisziplinäres Zentrum für Wissenschaftliches Rechnen, Im Neuenheimer Feld 205, D-69120 Heidelberg, Germany

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We revisit the conditions present in supermassive discs (SMDs) formed by the merger of gas-rich, metal-enriched galaxies at red-shift \( z \sim 10 \). We find that SMDs naturally form hydrostatic cores which go through a rapidly accreting supermassive star phase, before directly collapsing into massive black holes via the general relativistic instability. The growth and collapse of the cores occurs within \( \sim 5 \times 10^5 \) yr from the formation of the SMD, producing bright electromagnetic, neutrino and gravitational wave transients with a typical duration of a few minutes and, respectively, a typical flux and a typical strain amplitude at Earth of \( \sim 10^{-8} \) erg s\(^{-1}\) cm\(^{-2}\) and \( \sim 4 \times 10^{-21} \). We provide a simple fitting formula for the resulting black hole masses, which range from a few \( 10^6 M_\odot \) to \( 10^8 M_\odot \) depending on the initial SMD configuration. Crucially, our analysis does not require any specific assumption on the thermal properties of the gas, nor on the angular momentum loss mechanisms within the SMD. Led by these findings, we argue that the merger-driven scenario provides a robust pathway for the rapid formation of supermassive black holes at \( z > 6 \). It provides an explanation for the origin of the brightest and oldest quasars without the need of a sustained growth phase from a much smaller seed. Its smoking gun signatures can be tested directly via multi-messenger observations.

Key words: Hydrodynamics; quasars: supermassive black holes; galaxies: high-redshift; X-rays: bursts; gravitational waves; methods: analytical;

1 INTRODUCTION

After the first radio observations and key theoretical insights of the 1960ies (see, e.g. Bolton et al. 1963; Schmidt 1963; Zel’dovich 1964), recent surveys have found several quasars with estimated masses exceeding \( 10^9 M_\odot \) at redshift \( z > 7 \), i.e. only 700 million years after the Big Bang (see, e.g. Mazzucchelli et al. 2017; Lyke et al. 2020; Wang et al. 2021). These observations are both impressive and unexpected. They call for satisfying theoretical models, able to both explain the existence of such behemoths, as well as the uncountable number of smaller supermassive black holes (SMBHs) that likely exist at high redshift. Or perhaps it would be more reasonable for such extreme outliers to form via completely independent mechanisms from the large majority of SMBHs? We will return to this question in the conclusion of this work.

Barring primordial BHs scenarios (for some recent constraints, see, e.g. Bellomo et al. 2018; Zhou et al. 2022a,b), proposed mechanisms for early SMBH formation can be broadly separated into two categories. In the so-called “light seeds” scenario, BHs can form at \( z \gtrsim 15 \) as the last stage in the evolution of population III stars (see, e.g. Madau & Rees 2001; Abel et al. 2002; Schneider et al. 2002; Hirano et al. 2014). Despite the large stellar masses, population III stellar evolution models consistently predict compact remains of around \( \lesssim 10^2 M_\odot \). Considering the fact that Eddington-limited accretion can typically only double BH masses on a Salpeter time of \( \lesssim 5 \times 10^7 \) yr (Eddington 1921; Salpeter 1964), it has become doubtful whether light seed models can satisfactorily explain the enormous masses of high-redshift quasars (see, e.g. Alvarez et al. 2008; Smith et al. 2018; Zhu et al. 2020; Sassano et al. 2021). A more promising origin for SMBHs is provided by “direct collapse” models (DCMs). What defines a DCM is the possibility for large amounts of gas to collapse into a “heavy seed”, typically with a mass of around \( \sim 10^4 M_\odot \) to \( \sim 10^5 M_\odot \), which provides a significant jump-start for the growth towards higher masses. However, DCMs often require a very specific balance between angular momentum loss and cooling mechanisms in order for

\* E-mail: zwicklo@ics.uzh.ch

© 2022 The Authors
gas structures to contract without fragmenting or reaching new equilibrium states. The question of BH growth is also not completely resolved, since several hundred million year stretches of Eddington-limited accretion are still necessary for a $10^9 M_\odot$ seed to reach observed SMBH masses by $z \sim 7$. While possible in principle, such sustained growth rates require rather compliant gas reservoirs and have (arguably) proven difficult to achieve in simulations where accretion is not assumed a priori. Several works have shown that seeds are only able to grow by a few orders of magnitude before further accretion is progressively quenched by stellar feedback, supernovae or even out-flowing winds (see, e.g. Fiacco & Rossi 2016; Lupi et al. 2016; Regan et al. 2019; Sassano et al. 2022). Due to the many uncertainties in the modelling of high-redshift galaxies, the degree to which some (or all) DCMs might be too idealised is still unclear, and no consensus has yet been reached with regards to which might be the most realistic. We refer the reader to the many insightful reviews on the advantages and disadvantages of different DCMs (see, e.g. Natarajan 2011; Volonteri 2012; Haiman 2013; Mayer & Bonoli 2019; Woods et al. 2019) and we will mention specific ones throughout this work.

In recent years, Mayer et al. have proposed a scenario in which many of the typical problems specific to DCMs might be largely by-passed. This pathway to the formation of SMBHs originates from simulations of major gas rich galaxy mergers at $z \sim 10$ (Mayer et al. 2010; Bonoli et al. 2014; Mayer et al. 2015) and is therefore referred to as the “merger-driven” DCM. In these simulations, tidal torques and hydrodynamical instabilities caused by the merger induce gas inflow rates in excess of $\sim 10^4 M_\odot$ yr$^{-1}$. Enormous quantities of gas are rushed into the centre of the newly merged proto-galaxy, where they form a rotationally supported disc-like gas structure. These “supermassive discs” (SMDs) typically contain upwards of $10^9 M_\odot$, localised in a region of only $\lesssim 1$ pc. In contrast to typical proto-galactic discs, they are opaque to radiation, stable to fragmentation and contain up to Solar metallicity due to their relatively late formation. The argument in Mayer et al. (2010) and Mayer et al. (2015) (hereafter MI and MII, respectively) is that SMDs represent an optimal environment for the rapid formation of an SMBH, either through an intermediate supermassive star (SMS) phase (see, e.g. Haemmerle et al. 2020, 2021), or directly via the collapse of a general relativistic (GR)-unstable region. The advantages of the merger-driven scenario are that the formation of SMDs does not require any specific restrictions on thermodynamical conditions of the gas, nor any specific angular momentum loss mechanism other than the ones provided by the hydrodynamics of galaxy mergers. However, the exact process by which an SMD might produce a BH is unspecified, since its formation is likely to occur well below the resolution limit of the aforementioned simulations. The aim of this work is thus set: revisit the peculiar conditions present in SMDs to determine the mass of BHs predicted by the merger-driven scenario.

The paper is structured as follows: In Section 2, we briefly discuss the peculiar initial conditions, time-scales and physical processes that govern the evolution of SMDs. In Section 3, we use thermodynamical arguments to describe the accretion of the SMD onto a hydrostatic core via a generic angular momentum loss mechanism. We derive a mass-radius-temperature relation for the core and show that it can grow to significant size before approaching the GR instability regime. We provide a simple fitting formula for the mass of the resulting BHs, which have a typical scale of a few $10^7 M_\odot$. In Section 4, we discuss the implications as well as some caveats of our model. We focus on the direct observational consequences of the merger-driven DCM, highlighting possible electromagnetic and gravitational counterparts. Finally, we summarise our findings and present some concluding remarks in Section 5.

2 THE PHYSICS OF SUPERMASSIVE DISCS
2.1 The initial configuration

The formation of SMDs has been discussed extensively in several works (MI; Mayer et al. 2010; MII; Mayer et al. 2015, Bonoli et al. 2014, Mayer & Bonoli 2019). For the purposes of this work, we assume that the final stages of the aforementioned multi-scale simulations are representative of the fate of gas in major galaxy mergers at $z \sim 10$. In particular, we take the SMDs analysed in MII as a baseline case, but allow for important parameters such as mass, size and temperature to vary slightly from the simulated results, in case the latter are contaminated by numerical artefacts and spurious sub-grid physics.

During the final stages of the aforementioned simulations, large quantities of gas settle into a rotationally supported SMD containing upwards of $10^9 M_\odot$. The mass is localised in a region of size $\sim 1$ pc, yielding a typical average density of $\rho \sim 10^{-11}$ Kg/m$^3$, or $\sim 10^{10}$ particles per cubic centimetre. The gravitational compactness $K_\odot$ is defined as the size of an object divided by the Schwarzschild radius (Schwarzschild 1916) of its total mass. Typical values for SMDs read:
The initial thermal state of an SMD is determined by the dynamics of its formation. The competition between atomic cooling, shock heating and turbulent dissipation results in a disc which is approximately isothermal at a temperature of \( T_d \sim 7 \times 10^4 \) K, in which fragmentation is suppressed regardless of the presence of metals. Despite the different physics at play, this final state is comparable to the virial temperatures commonly expected in smaller, metal-free proto-galactic discs which are often invoked in DCMs (see, e.g. Rees & Ostriker 1977; White & Frenk 1991). At typical temperatures of \( \lesssim 10^4 \) K, the total thermal energy of the gas is not sufficient for spherical hydrostatic equilibrium and the inflows settle into a largely rotationally supported disc. In MI and MII, the resulting SMDs display a constant scale-height of approximately \( \sim 0.1R_d \), likely due to the resolution limit. In this paper, we assume the standard scaling for the disc’s scale-height, \( H \sim c_s/\omega \), where \( \omega \) is the angular velocity and \( c_s \), the isothermal sound speed, of sound for a mixture of gas and radiation.

Any self gravitating fluid is likely to exhibit some degree of mass concentration, i.e. a density profile of the form \( \rho \sim r^{-n} \). Pressure-less thin discs with arbitrary density profiles have been famously studied in Mestel (1963), and the allowed rotational configurations range from constant angular velocity \( (\rho \sim r^n \text{ and } v_t \sim r) \) to constant tangential velocity \( (\rho \sim r^{-2} \text{ and } v_t \sim r^2) \). A steep density profile also implies that the pressure gradient is dominated by radiation pressure and gas pressure, leading to a more complex behaviour of the sound speed and the scale-height. Temperature fluctuations can potentially also affect the balance of pressure. Whenever a system is optically thick however, radiative heat transfer can be approximated via the Rosseland diffusion equation (Rosseland 1936; Mihalas & Mihalas 1984). One consequence of diffusion is to smooth out any sufficiently strong temperature variability, suggesting that the isothermality seen in MI and MII is also a good approximation at scales that are below the resolution limits.

\[ K = \frac{c^2 R_d}{2GM_d} \sim 10^3 \left( \frac{10^9 M_\odot}{M_d} \right) \left( \frac{R_d}{1 \text{ pc}} \right), \]

where \( M_d \) is the mass of the SMD and \( R_d \) its radius. The dynamical time-scale \( t_{\text{dynamical}} \sim \rho^{-1/2} \) of such a system is close to 500 yr, only two orders of magnitude larger than the light-crossing time \( t_c \sim R_d c^{-1} \sim 3 \) yr, where \( c \) is the speed of light.

The unusual compactness of SMDs is the primary motivation for considering the GR instability criterion, which will be discussed in more detail in Section 2.3. A crucial difference between SMDs and more conventional proto-galactic discs is the fact that they form at the transition between the optically thin and optically thick regimes. Therefore, any further contraction will be strongly affected by adiabatic heating. Consider the mean free path \( \lambda_{\text{ph}} \) of a photon that undergoes Thomson scattering (Thomson 1851):

\[ \lambda_{\text{ph}} = \frac{1}{n_e \sigma_T} \sim A \times 10^{-5} \left( \frac{10^9 M_\odot}{M_d} \right) \left( \frac{R_d}{1 \text{ pc}} \right)^3 \text{ [pc]}, \]

where \( A \) corresponds to the average aspect ratio of the disc (which typically ranges from \( \sim 10^{-2} \) to \( 10^{-1} \)), \( n_e \) is the number density of free electrons, \( \sigma_T \) is the Thomson cross-section and we assumed full ionisation, which is appropriate for low density hydrogen at these temperatures (Saha 1921). Eq. 2 seems to suggest that SMDs should be considered opaque. However, a system with a short mean free path can still lose heat over a sufficiently long period of time. What determines if a system truly behaves adiabatically is the diffusion timescale \( t_{\text{diff}} \):

\[ t_{\text{diff}} = \frac{A^2 R_d^2 \lambda_{\text{ph}} c}{\lambda_{\text{ph}}^2} \sim 10^5 A \left( \frac{M_d}{10^9 M_\odot} \right) \left( \frac{1 \text{ pc}}{R_d} \right) \text{ [yr]}, \]

where we approximate photon paths as random walks. Due to the enormous accretion rates triggered by the galaxy merger \( (\sim 10^4 M_\odot \text{ yr}^{-1}) \), more and more mass is accumulated within a given central region, enabling the formation of an SMD in approximately \( 10^5 \) yr, a time comparable to Eq. 3. This suggests that the assembly of an SMD occurs mostly in an “effectively” optically thin regime, where radiation has time to escape the system before much more mass is added. The system finally becomes truly opaque only when a sufficient amount of mass \( (\sim 10^9 M_\odot) \) is concentrated within a region of a fixed scale \( (\sim 1 \text{ pc}) \). From this point on, the accretion of even more mass would not change the thermodynamic behaviour of the disc qualitatively, since it is already adiabatic as a whole\(^2\). Interestingly, this simple consideration yields a mass-radius relation for SMDs, which sets a natural density scale that nicely aligns results reported in MI (whereas it slightly differs from the results in MII, in which SMDs are more compact by a factor few).

---

1 SMDs are as gravitationally compact as the typical white dwarf, but with densities twenty order of magnitudes lower (Adams 1925; Koester & Chugamug 1990).

2 The addition of orders of magnitudes more mass might indeed change the thermodynamics of the SMD. This is however not a realistic possibility since the size of dark matter haloes and their gas reservoirs are limited in size at \( z \sim 10 \).

---

\(^2\) Merger Driven Direct Collapse
e.g. Agarwal et al. 2012; Latif et al. 2013; Regan et al. 2014; Habouzit et al. 2016; Agarwal et al. 2016; Johnson & Dijkstra 2017). Dynamical mechanisms involve the rapid transport of angular momentum, inducing a contraction of the disc on time-scales comparable with gravitational free-fall. While these processes can potentially circumvent the issue of fragmentation, they are conditional to specific triggers (such as a nearby supernova explosion, Colgate et al. 2003) or to the onset of global gravitational instabilities (e.g. Koushiappas et al. 2004; Begelman et al. 2006; Lodato & Natarajan 2006, 2007; Begelman 2008). A mechanism based on cold, turbulent gas inflows has been proposed recently and could circumvent some of the aforementioned challenges, producing DC BHs with masses of the order $\sim 10^3$ $M_\odot$ (Latif et al. 2022).

Crucially, SMDs provide a realistic cosmological object in which not only dynamical, but also secular mechanisms generally act on time-scales comparable to the diffusion timescale. The ratio between the diffusion time-scale and dynamical time-scales is at least $\sim A \times 10^3$, and increases dramatically when considering denser sub-regions of the disc. To assess secular mechanisms, we consider the example of viscous dissipation by means of magnetic gravito-turbulence, a key player in the formation of SMDs (as discussed in MI and MII) that is found to be ubiquitously important in astrophysical disc (see, e.g. Hawley et al. 1995; Pessah et al. 2007; Abramowicz & Fragile 2013; Prieto et al. 2013). The ratio of the diffusion and viscous dissipation time-scales $t_{\text{visc}}$ can be estimated as follows:

$$\frac{t_{\text{diff}}}{t_{\text{visc}}} \sim A^2 \frac{2 \alpha R_\text{d} c_s}{\lambda_{\text{visc}}} \frac{c_s}{c}, \quad (4)$$

where we assumed an $\alpha$-viscosity prescription (Shakura & Sunyaev 1973) and used the standard scaling $t_{\text{visc}} \sim \alpha^{-1} R_\text{d}^2 c_s^{-1} H^{-1}$. At a temperature of a few thousand Kelvin, the speed of sound in gases is generally much smaller than the speed of light. Consider however, the ratio of thermal energy in a typical SMD that is stored in radiation ($E_{\text{rad}}$) rather than gas ($E_{\text{gas}}$):

$$\frac{E_{\text{rad}}}{E_{\text{gas}}} \sim A \times 10^3 \left( \frac{10^9 M_\odot}{M_\text{d}} \right) \left( \frac{R_\text{d}}{\Gamma_{\text{adi}}} \right)^3 \left( \frac{T_\text{d}}{7000 \text{ K}} \right)^3, \quad (5)$$

scaled here for values typical to SMDs. Whenever a system is radiation-pressure dominated, the speed of sound approaches the value appropriate for a photon fluid, i.e $c/\sqrt{3}$ (see, e.g. Mihalas & Mihalas 1984). This implies that the ratio $c_s/c$ in Eq. 4 is actually a fraction of order unity rather than a small number. If we assume values for the $\alpha$ parameter of $\sim 0.2$, which are expected in fully ionised discs (see, e.g. Lodato & Rice 2004; King et al. 2007), we can evaluate Eq. 4 for a typical SMD:

$$\frac{t_{\text{diff}}}{t_{\text{visc}}} \sim A^2 10^3 \left( \frac{M_\text{d}}{10^9 M_\odot} \right) \left( \frac{1 \text{ pc}}{R_\text{d}} \right)^2, \quad (6)$$

where we replaced the number $\alpha c_s c^{-1}$ as a fraction of order $10^{-1}$, consistently with the arguments above. Eq. 6 shows that, for reasonable values of the aspect ratio, SMDs are expected to dissipate their rotational energy in a time that is comparable or smaller than a diffusion time-scale, i.e. smaller than at most $\sim A \times 10^5$ yr. Even more so when considering sub-regions of the disc, where the viscous time-scale decreases substantially and the diffusion time-scale is affected by the higher average density. To re-iterate, the dynamical time-scale in SMDs is typically of the order 500 yr, while both the diffusion and the viscous time-scales are on the order of $5 \times 10^5$ yr. The crucial implication of Eq. 6 is that the merger-driven scenario does not require a dynamical angular momentum loss mechanism, nor a source of ionising radiation in order to circumvent the problem of fragmentation. Rather, the evolution of the SMD naturally occurs in a regime where cooling is negligible, i.e. the adiabatic regime, even in the slowest possible case of secular angular momentum loss.

### 2.3 GR instability

The unlikely combination of a high gravitational compactness and a low density is suggestive that SMDs should be close to the GR-unstable regime (without considering the effect of their geometry). The GR instability relies on the relativistic corrections to the equation of radial momentum, that translate into an increase of the critical value for the first adiabatic exponent $\Gamma_1$, above the Newtonian value of 4/3. An analytic criterion for GR instability has been derived by Chandrasekhar (1964) for the general case of a hydrostatic, spherical mass. This increase scales with gravitational compactness of the system, For a polytrope (i.e. for homogeneous entropy distribution) the instability is reached when:

$$\Gamma_1 = \frac{4}{3} < \frac{\kappa 2GM}{c^2 R}, \quad (7)$$

where $M$ is the system’s mass, $R$ its radius and $\kappa$ is a factor of order unity that depends on the polytropic index. The GR instability is global, i.e it is triggered at the scale of the whole system. The expectation is therefore that the collapse cannot be halted until a compact object is formed. Several numerical GR simulations have confirmed this, showing that a large fraction of the cloud’s mass (60% to 90%) is retained in the final BH (see, e.g. Sari & Hawke 2009; Reisswig et al. 2013). In the context of DCMs, the GR instability is invoked as a mechanism to destabilise supermassive stars, which can potentially lead to the formation of highly massive BHs (see, e.g. Haemmerlé 2020, 2021; Woods et al. 2021). The adiabatic index of a mixture of gas and radiation $\Gamma$ depends on its temperature and density. It is determined by the ratio between gas pressure and the total pressure, $\beta$, which is approximated by the following expression in the case of radiation dominated gases:

$$\Gamma_1 = \frac{32 - 27\beta}{24 - 21\beta} \approx \frac{4}{3} + \frac{\beta}{6}. \quad (8)$$
We model the latter by defining an enclosed mass profile \( M_c \) for the SMD, which we approximate as a simple power law:

\[
M_c = M_d \left( \frac{r}{R_d} \right)^{3-n}.
\]  

(15)

Here the radius \( r < R_d \) is an Eulerian radial coordinate defined within the interior of the SMD. The parameter \( n \) determines the degree of mass concentration, roughly corresponding to a density profile with a scaling of \( \rho \sim r^{-n} \), or a tangential velocity profile of \( v_t \sim r^{1-n/2} \). Since all values of \( n \) between 0 and 2 are consistent with self-gravitating, rotationally supported discs, we keep \( n \) as a free parameter. Higher resolution simulations of the formation of SMDs can reveal if specific values of \( n \) are preferred and our results can be evaluated accordingly.

Given Eqs. 9 to 15, we can find a characteristic radius at which the enclosed thermal energy \( E_T \) is equal to the enclosed gravitational energy \( E_G \). This is an important boundary which separates the region of an SMD that is rotationally supported from the one that is supported by gas or radiation pressure. By definition, the structure of the latter should not be strongly affected by rotation and can be described as an approximately spherical core in hydrostatic equilibrium. We estimate the core radius \( R_c \) by setting \( E_G \sim E_T \), and solving for the radius:

\[
E_G \sim \int \frac{G M_c \rho d \tau}{T} \sim \int P dV.
\]  

(16)

Here \( G \) is Newton’s constant and \( \rho_d \) is a density profile consistent with Eq. 15 and the scale height of the SMD. Simple explicit solutions for the core radius can be found by considering either gas or radiation pressure:

\[
\left( \frac{R_{\text{gas}}}{R_d} \right)^{2-n} = 2 \cdot 5 - 2n \cdot \frac{k_B T}{3-n} \cdot \frac{\rho_d}{\rho} \cdot \frac{R_d}{R_d} \cdot \left( \frac{T}{7000 \text{ K}} \right)
\]  

(17)

\[
\left( \frac{R_{\text{rad}}}{R_d} \right)^{2-n} = 16 \pi 5 - 2n \cdot \frac{\sigma_B T^4}{3-n} \cdot \frac{\rho_d}{\rho} \cdot \frac{R_d}{R_d} \cdot \left( \frac{T}{7000 \text{ K}} \right)
\]  

(18)

while the general solution can be found by solving Eq. 16 numerically. Eqs. 17 and 18 should be understood as describing the mass and size of a virialised region, which will likely settle into a hydro-static core resembling a radiation dominated supermassive star (SMS) on a local dynamical timescale. Within the scope of this work, we will not be able to precisely model the evolution and the internal structure of these virialised regions, and will make no distinction between them and the SMS they might produce. We discuss this caveat in detail in section 4.2, and expect our methods to correctly capture the relevant global parameters such as total mass, energy content and approximate size.

The results of Eq. 19 are visualised in Fig. 2 for several values of the initial temperature of the SMD. The size of the hydrostatic core varies significantly depending on the parameter \( n \), i.e. the degree of matter concentration: SMDs with

3 The description of a mixture of gas and radiation technically requires three separate adiabatic indexes. Here we only discuss the relevant one for the GR instability criterion.

4 Electrons and protons contribute equally to gas pressure in an ideal, relaxed plasma (see, e.g. Stoltzmann 2001)
shallower density profiles \((n \lesssim 1)\) form large radiation pressure dominated cores, while more concentrated discs form smaller, gas pressure dominated cores. The presence of such a core suggests that the inevitable contraction of SMDs caused by angular momentum loss should be thought as an accretion process onto a spherical, pressure supported central region.

### 3.2 Accretion of the disc onto the core

In order to properly describe the accretion of mass and energy onto the central core, we consider a shell of fluid elements that is loosing or dissipating rotational energy by the mechanisms detailed in Section 2.2. In order for a shell to contract radially, an amount of specific rotational energy proportional to the enclosed gravitational energy must be removed from every fluid element. Because of the arguments detailed in Section 2.2, we expect sub-regions of the SMD to contract adiabatically, i.e. for cooling losses to be neglectable. Furthermore, in the case of secular angular momentum loss mechanisms, rotational energy is slowly dissipated within the gas, adding a heating term to its energy balance. We can model all contributions to the internal energy with the following equation:

\[
E(x) = E_{\text{gas}} \left( \frac{r_0}{x} \right)^2 + E_{\text{rad}} \left( \frac{r_0}{x} \right) + \int x \frac{GM_s}{r^2} \, dr, \tag{19}
\]

where \(r_0\) is the initial radius and \(x\) is the current radius of the fluid element. Here the first two terms describe the adiabatic heating of gas and radiation specific energy from the initial values \(E_{\text{gas}}\) and \(E_{\text{rad}}\), while the integral term accounts for the dissipated rotational energy. The latter must be proportional to the gravitational energy at the radius \(r\) in order for rotational support to be significantly affected.

In the case of dynamical mechanisms, which can transport angular momentum without necessarily dissipating its associated energy, we still expect the integral term in Eq. 19 to appear alongside the adiabatic contributions. The reason for this is slightly subtler and is due to an effect observed in contracting, self-gravitating fluids known as turbulent support (see, e.g. Larson 1981; Murray & Chang 2015; Li 2018; Xu & Lazarian 2020). If a seed of turbulence is present within a fluid element (which is the case in SMDs), any global contraction causes it to heat up, i.e. to increase in turbulent mean-root-square velocity (Robertson & Goldreich 2012; Hennebelle 2021). In the case of radial self-gravitating flows, turbulence is thus driven to a steady state in which a significant fraction of the inflows’ kinetic energy is redirected into turbulent eddies, supporting the system against its own gravity. Since the inflow velocity is ultimately sourced by the gravitational potential, the amount energy that is dissipated by turbulent viscosity will necessarily be proportional to the gravitational energy. Essentially, turbulent heating acts to “locally virialise” the internal energy in a contracting fluid, leading to an integral term in Eq. 19 even in the case of a dynamical angular momentum loss mechanism (with a possible prefactor of order unity, which we neglect). The interesting implication of this local virialisation is that the energy content of fluid elements is unaffected by the details of the angular momentum loss mechanism, in particular the time-scale on which it acts (provided it is shorter than a diffusion time-scale, as discussed in Section 2.2). This is a stark contrast between the merger-driven scenario and other DCMs and is partially responsible for the robustness of the predicted BH masses in Section 3.3.

Eq. 19 shows that, as a gas shell contracts radially, it will accrue a certain amount of internal energy along its path. This contraction is halted only once the conditions for a new equilibrium configuration are reached, i.e. when the total internal energy within a certain radius is comparable to the gravitational self energy. For the considered density profiles, this condition can be approximated mathematically in terms of the specific energy, which reads:

\[
E(x) \sim \frac{GM_s}{x}. \tag{20}
\]

We can solve Eq. 20, in combination with Eq. 19, to find the new equilibrium radius of every shell of fluid elements within the SMD. If the system is dominated by gas pressure, the new equilibrium radius \(r_{eq}\) reads:

\[
r_{eq}^{\text{gas}} = r_0 \left( \frac{2k_B R_d T_0}{G M m_p} \right)^{1/2} \left( \frac{R_d}{r_0} \right)^{(2-n)/2}. \tag{21}
\]

If it is radiation pressure dominated, it reads:

\[
r_{eq}^{\text{rad}} = r_0 \left( \frac{8 \pi R_d^4 \sigma_B T_0^4}{3 c G M^2 (3-n)} \right) \left( \frac{R_d}{r_0} \right)^{(2-2n)}. \tag{22}
\]

The general solution can be found by solving Eq 20 numerically. Eq. 21 and Eq. 22 are describing how SMDs transition from rotationally supported discs to smaller, spherical gas clouds in hydrostatic equilibrium. This is exactly what is expected, since SMDs lose angular momentum while retaining most of their thermal energy (see Eq. 3). The picture is thus one of a growing core embedded within a gaseous disc that contracts from the inside outwards. We define a contraction factor, \(\kappa\), which encodes the global change in geometry from a disc of radius \(R_d\) to a spherical cloud of radius \(K R_d\) and is found by evaluating either Eq. 21 or Eq. 22 at the disc radius \(R_d\). An explicit expression can be calculated for a gas pressure or radiation pressure dominated system:

\[
\kappa_{\text{gas}} = \left( \frac{2k_B R_d T_0}{G M m_p} \right)^{1/2} \approx 5 \times 10^{-3} \left( \frac{R_d}{\text{pc}} \right)^{1/2} \left( \frac{10^9 M_\odot}{M} \right)^{1/2} \left( \frac{T}{7000 \, \text{K}} \right)^{1/2}, \tag{23}
\]

\[
\kappa_{\text{rad}} = \left( \frac{8 \pi R_d^4 \sigma_B T_0^4}{3 c G M^2 (3-n)} \right) \approx 1.3 \times 10^{-2} \left( \frac{R_d}{\text{pc}} \right)^4 \left( \frac{10^9 M_\odot}{M} \right)^2 \left( \frac{T_0}{7000 \, \text{K}} \right)^4. \tag{26}
\]

In both gas and radiation pressure dominated cases, the factor has a typical value of \(10^{-2}\), suggesting that SMDs are
likely to contract significantly once rotational support is completely lost. A more accurate result can be found by solving Eq. \ref{eq:adiabatic_index} numerically.

Once a given shell has contracted to its new equilibrium radius, it has by construction added exactly the amount of mass and thermal energy needed for a core of mass $M_0$ to be in hydrostatic equilibrium, where $M_0$ is the enclosed mass within the shell’s initial radius $r_0$. This means that, with the aid of the solutions to Eq. \ref{eq:adiabatic_index}, we can find a simple relationship for the mass ($M_e$) and size ($R_e$) of the hydrostatic core once a given shell is added to it:

$$M_e(R_e) = M_d \left( \frac{r_0(R_e)}{R_d} \right)^{3-n},$$  \hspace{1cm} (27)

where $r_0$ is the initial radius of the specific shell that achieves equilibrium at the radius $R_e$. We can find explicit solutions for the gas and the radiation dominated regimes by simply inverting Eqs. \ref{eq:gas_mass} and \ref{eq:rad_mass}:

$$M_e^{\text{gas}}(R_e) = M_d \left( \frac{R_e}{K_{\text{gas}}R_d} \right)^{\frac{6}{n-2}},$$  \hspace{1cm} (28)

$$M_e^{\text{rad}}(R_e) = M_d \left( \frac{R_e}{K_{\text{rad}}R_d} \right)^{(n-3)/(2n-1)}. \hspace{1cm} (29)$$

As expected, the core’s mass approaches the mass of the whole SMD once its radius becomes $KR_d$. Numerical results for the mass-radius relation that include both gas and radiation pressure are shown in the top panel of Figure 3, for several choices of the parameter $n$. In general, the initial degree of mass concentration strongly determines the growth of the cores. At high masses however, the cores tend to approach the same radius regardless of the value of $n$. This behaviour is due to the cores always transitioning from being gas pressure to radiation pressure dominated. Therefore, their radius at a mass of $10^9 M_\odot$ must approach the value $K_{\text{rad}}R_d$, which is derived in the same limit. The total amount of mass and internal energy is always dominated by the accretion of new layers (for $n \lesssim 2$). Therefore, the temperature within the core must be mostly determined by the temperature of the last accreted shell, which is by construction equal to the virial temperature of the enclosed mass. We can therefore assume that the cores always re-establish local isothermal equilibrium at the virial temperature appropriate for their new radius $r_{eq}$, every time a new shell is added.

As discussed previously, the mass-radius relationship for the core relies exclusively on the assumptions of adiabaticity and virialisation of rotational energy. It is therefore a generic feature of the evolution of SMDs that does not depend on the details of angular momentum loss. Adiabaticity can be broken in the presence of significant amounts of nuclear fusion, which would provide an additional heating term. For the purposes of this paper, we adopt a temperature limit of $\sim 2 \times 10^7$ K for the ignition of hydrogen within the core (see, e.g. Stahler \& Palla 2004), and consider any core whose virial temperature exceeds this limit to be affected by nuclear fusion. As is visible in the lower panel of Figure 3, only SMDs with high degrees of mass concentration ($n \gtrsim 1.7$) produce cores in which nuclear fusion is likely. In this case, the cores would tend to heat up and increase in radius faster than what our relations predict, possibly evolving along supermassive star tracks (see, e.g. Hosokawa et al. 2012; Haemmerlé 2021). We refer the reader to section 4.2 for a more thorough discussion.

### 3.3 Formation of a BH via the GR instability

As more and more shells lose angular momentum, the hydrostatic core increases in size, mass and internal energy. The relative contribution between gas pressure and radiation pressure changes in favour of the latter. As discussed in Section 2.3, the criterion for GR instability is met whenever the adiabatic index is sufficiently close to the critical value of 4/3. For the purposes of this work, we deem the core to be GR-unstable whenever its average adiabatic index crosses over a phenomenological value $\gamma_p$ given by the following equation:

$$\gamma_p = \frac{4}{3} + \frac{1}{2} \frac{GM_c}{c^2 R_c},$$  \hspace{1cm} (30)
Changing the prefactor mass in discs with shallower density profiles (centred towards the centre of the disc. The initial instability is triggered at smaller scales if more mass is concentrated, while SMDs with shallow density profiles produce large masses, since it only requires the seeds to accrete at Eddington levels for a few Salpeter time-scales rather than dozens. The additional mass can be provided by the rest of the SMD, which must eventually settle into an accretion disc around the newly formed seed, likely powering quasar emission.

In general, more concentrated SMDs produce smaller BH seeds, while SMDs with shallow density profiles produce large BH seeds. The trend is simply due to the fact that the GR instability is triggered at smaller scales if more mass is concentrated towards the centre of the disc. The initial temperature of the SMD has a noticeable effect on the final BH mass in discs with shallower density profiles ($n \lesssim 1.4$). This is a consequence of the relative amount of mass in the core that is primarily supported by gas pressure versus radiation pressure. A higher initial temperature produces less dense, puffier cores that become unstable at larger scales, thus enclosing more mass. The total initial mass and radius of the disc have a negligible effect on the final masses, since the general density scale of SMDs is set by the transition between the optically thin and adiabatic evolution regimes (see Section 2.1) and its value cannot deviate significantly from $\sim 10^{-11}$ Kg/m$^3$. Less dense discs with a size of $\sim 1$ pc might still produce some form of compact objects, just not via the specific mechanism proposed in this paper. Much denser discs are unlikely to exist at $z \sim 10$ due to the limitations of dark matter halo masses. If one still allowed the density to vary, denser SMDs would typically produce slightly lighter BHs, while the reverse is also true. We stress that the collapse of these exceptionally heavy seeds is rapid. From the formation of the SMD at $z \sim 10$, at most $\sim 5 \times 10^5$ yr are required for the core to approach the GR-unstable regime. The core then collapses on a local dynamical time, which ranges from days to months for the relevant densities. To summarise all the possibilities, we provide a simple formula which accurately fits (to a factor of order unity) the numerical results within a range of reasonable parameters for the initial SMD, i.e $2 < n < 0$ and $5000$ K $< T_0 < 11000$ K. The final BH masses ($M_{BH}$) are given by the following expression:

$$\frac{M_{BH}}{M_\odot} \sim 3 \times 10^7 f(T_0, n)$$

$$f(T_0, n) = \left(1 - \frac{n}{2}\right)^{2.8} + \frac{2}{3} \left(\frac{T_0}{7000 \text{ K}}\right)^4 \left(1 - \left(\frac{4n}{5}\right)^2\right)^{1.8},$$

where the latter term must be set to 0 as soon as it becomes undefined ($n > 1.25$). We expect this fitting formula to be accurate (within the limits of our analysis) for SMDs in which the core is unlikely to ignite, i.e. with a parameter $n \lesssim 1.7$. It can be applied as a sub-grid model in cosmological simulations of major mergers as well as population models that have the aim of reproducing the observed statistics of high-redshift quasars and the BH mass function (see, e.g. Bonoli et al. 2014; Spinoso et al. 2023).

4 DISCUSSION

4.1 Smoking gun observational signatures

The hydrostatic cores described in Section 3.2 grow in size over the course of the accretion process, reaching radii of $\sim 10^{-7}$ pc to $10^{-2}$ pc before collapsing via the GR instability. Because of their large size and relatively high temperature ($\sim 10^6$ K), we expect them to be bright sources of thermal radiation. Neglecting for a moment the fact that the cores are embedded in opaque discs, we can describe their bolometric luminosity, $L_c$, by means of the Stefan-Boltzmann law:

$$L_c = 4\pi R_c^2 \sigma T_c^4.$$  

The growth of the cores occurs on a time-scale comparable to the angular momentum loss time-scale, i.e. $\lesssim 10^5$ yr. Over
this period, the temperature variation is limited (as shown in the bottom panel of figure 3) and we expect the luminosity to roughly increase as $\sim R_c^2$. Once the GR instability is reached, the core collapses on a much shorter dynamical time-scale. This rapid contraction causes a violent increase in temperature, both due to adiabatic heating and virialisation by turbulent dissipation. The temperature in a contracting adiabatic system scales as $T \sim R^{-1}$. We therefore expect the luminosity of the cores to roughly scale as $R_c^{-2}$ during the collapse, mirroring the trend during the growth phase. However, a couple of effects can taper down the luminosity in the latter stages of collapse. Firstly, the rapid change in size will detach the core from the rest of the disc. It will then be able to freely radiate thermal energy into the cavity it leaves behind. This introduces a cooling term in the energy balance, which is described by Eq. 33. Secondly, whenever temperatures in the core reach values of $\sim 10^9$ K, pair annihilation and the URCA process can induce a large outward flux of neutrinos (see, e.g. Itoh et al. 1989; Qian & Woosley 1996; Dutta et al. 2004). Neutrinos can carry a significant amount of energy, adding an additional cooling term that is highly sensitive to temperature. For the purposes of this work, we adopt a simple approximation to the neutrino cooling rate $Q_\nu$, also used in Begelman et al. (2006) in the context of DCMs:

$$Q_\nu \approx 3 \times 10^{15} \left( \frac{T_c}{10^9 K} \right)^9 \left[ \frac{\text{erg}}{\text{s cm}^2} \right].$$  \hspace{1cm} (34)

Finally, gravitational redshift, $z_\text{g}$, can significantly dim the luminosity of a system that is approaching its own gravitational radius. The value for gravitational redshift is given by the following formula:

$$z_\text{g} = \left( 1 - \frac{\text{R}_\text{s}}{\text{R}_\text{c}} \right)^{-1/2} - 1,$$ \hspace{1cm} (35)

and it decreases the bolometric luminosity by a factor $(1 + z_\text{g})^{-3}$ (see, e.g. Peebles 1993). In Figure 5 we show the bolometric luminosity tracks of several SMD cores as a function of their radius, for different values of the parameter $n$. To produce the tracks, we consider the effects of accretion, adiabatic heating, virialisation, black body radiation, neutrino cooling and gravitational redshift on the thermodynamics of the cores. We express the results in terms of the expected flux at a luminosity distance appropriate for $z = 10$ source. This would represent the flux as seen by an observer on the Earth, assuming that radiation is able to freely escape the host proto-galaxy. We consider this process to be a first step in assessing the direct detectability of a merger-driven DCM event as a source of radiation, though significantly more modelling is required for a realistic prediction.

As can be seen in the figure, the tracks initially follow the growth of the core. During this phase, the flux typically peaks at values of $\sim 10^{-12}$ erg s$^{-1}$ cm$^{-2}$ and is mostly concentrated in the ultraviolet and soft X-ray bands. The duration of this phase depends on the angular momentum loss mechanism, which can range from $\sim 10^6$ yr to $\sim 10^9$ yr. Once the GR instability is triggered, the cores start to contract and the flux rapidly increases by several orders of magnitudes due to the rise in temperature. The signatures of this phase are extremely bright, hard x-ray (and potentially gamma-ray) transients that peak at bolometric flux of $\sim 10^{-8}$ ergs s$^{-1}$ cm$^{-2}$. The duration of this phase depends on the size of the core as it becomes unstable. In general, the peak luminosity is limited to an extremely short period of a few seconds to minutes. During this phase, the core detaches from the rest of the disc due to its rapid collapse, leaving behind a cavity at the centre of the disc. Depending on the configuration and the disc geometry, it possible that radiation may escape from the proto-galaxy, as suggested by the cartoon in Figure 1. The very last stages of the tracks differ qualitatively for SMDs with shallow ($n \lesssim 1$) or steep ($n \gtrsim 1$) initial density profiles. For the former, the flux is smoothly reduced by gravitational redshift in the moments before a BH is formed. For the latter, the temperatures are high enough for neutrino cooling to suddenly become efficient. While the electromagnetic luminosity is reduced, neutrino cooling induces a flux of particles with a scale of $\sim 10^{-7}$ ergs s$^{-1}$ cm$^{-2}$. Neutrinos are much more likely to freely exit the proto-galaxy, and the collapse of an SMD core could in principle be detectable as a point source transient by current and future instruments (see, e.g. Abbasi et al. 2011; IceCube Collaboration et al. 2013; Hallmann et al. 2021).

The electromagnetic and particle transients are not the only smoking guns of the merger-driven DC scenario. If a small amount of angular momentum is retained within the core at the moment it becomes unstable (which we discuss in Section 4.2), we would expect its collapse to not be exactly spherically symmetric, i.e. have a non vanishing quadrupole mass moment. According to GR, this would imply the generation of a gravitational wave burst (Einstein 1916a,b; Misner...
et al. 1973; Chia et al. 1977; Fryer & New 2011). The gravitational collapse of oblate spheroids has been studied in several works (see, e.g. Lin et al. 1965; Fujimoto 1968; Larson 1972; Goodman & Binney 1983). For the purposes of this analysis, we assume that cores contract homologously, i.e. the ratio between their semi-major axes, $a$, and semi-minor axes, $b$, remains constant. The quadrupole moment, $\mathcal{M}$, of an oblate spheroid is given by the following formula:

$$\mathcal{M} = \frac{2}{5} M c R_c^2 \epsilon^2$$

(36)

$$\epsilon^2 = \frac{1 - (a/b)^2}{(a/b)^{4/3}}$$

(37)

where we replaced the geometric-mean radius of the spheroid with the core radius $R_c$ and defined an ellipticity $\epsilon$. In the quadrupole approximation, the typical amplitude, $h$, of a gravitational wave depends on the second derivative of the mass quadrupole:

$$h \sim \frac{G}{c^3 D_l} \mathcal{M},$$

(38)

where $D_l$ is the luminosity distance. Dimensionally, the amplitude produced by a collapsing spheroid is therefore given by the following simple expression:

$$h \sim \epsilon^2 \frac{G M c}{c^3 D_l} \left(\frac{v}{c}\right)^2$$

(39)

$$\approx 4 \times 10^{-21} \left(\frac{\epsilon}{10^{-2}}\right)^2 \left(\frac{M_c}{10^7 M_\odot}\right) \left(\frac{10 \text{ Gpc}}{D_l}\right) \left(\frac{v}{c}\right)^2,$$

(40)

where we introduced a characteristic velocity $v \sim \dot{R}_c$, which becomes comparable to the speed of light during the late stages of collapse. Scaled with values typical to cores in SMDs, we find that the amplitude of the burst is comparable to the sensitivity of both existing ground based and future space based detectors for an ellipticity of $\epsilon \gtrsim 10^{-2}$. The typical frequency of a burst, $f_b$, at its loudest point (when $v \sim c$) is comparable to the reciprocal of the light-crossing time for the final BH. For values typical to SMD cores it reads:

$$f_b \sim \frac{\epsilon^3}{2GM_c} \approx 10^{-2} \left(\frac{10^7 M_\odot}{M_c}\right) \text{[Hz]},$$

(41)

which places the bursts squarely in the preferred frequency band for the upcoming Laser Interferometer Space Antenna (Amaro-Seoane et al. 2017; Baker et al. 2019; Amaro-Seoane et al. 2022; Auclair et al. 2022) and TianQin detectors (Luo et al. 2016; Mei et al. 2021).

We can roughly estimate the event rate at which the transients discussed above are expected to occur. An estimate of the rate of appropriate major mergers that can produce SMDs is provided in Mayer & Bonoli (2019), ranging between $10^{-7}$ to $10^{-5}$ mergers per Mpc$^3$. Here we consider an interesting redshift window between $z = 7$ and $z = 12$, where the lower value is given by the furthest observed high mass quasars while the larger one is an arbitrary boundary above which major mergers are not expected. We can simply multiply the merger rate with the co-moving volume between $z = 7$ and $z = 12$, which amounts to $\sim 1500$ Gpc$^3$ for a flat ΛCDM universe (Wright 2006). We find that the total amount of appropriate mergers amounts to $\sim 10^5$ to $\sim 10^7$. Dividing the number by 500 million years, i.e. the elapsed time between $z = 7$ and $z = 12$, we find that we should expect a rate of $\sim 10^{-2}$ to $\sim 1$ gravitational, neutrino and electromagnetic transients per year, assuming that every suitable major merger inevitably leads to a DC event. The likelihood for detecting a counterpart is boosted significantly with instruments sensitive enough to observe the growth and the initial collapse phases of the cores due to their longer duration.

We look forward to the feed-back of the x-ray, neutrino and gravitational wave astronomy communities to assess whether these signatures are detectable in the foreseeable future, or might even be present in existing wide field surveys.

### 4.2 Caveats

The largest caveat of our analysis is the simplistic treatment of the hydrostatic cores, which are likely to settle into structures resembling SMSs. The arguments used in section 3.2 only allow for an order of magnitude estimate of their properties, e.g. the radius, mass and energy content. A more precise analysis would require full solutions to the equation of hydrostatic equilibrium, which necessarily implies the use of numerical models. In particular, the latter show that the structure of SMSs depend critically on their mass accretion rate. For rates greater than $0.01 M_\odot$ yr$^{-1}$, accretion prevents thermal relaxation on fully convective, polytropic structures. The SMS evolves instead as a ‘red supergiant proto-star’ (Hosokawa et al. 2012, 2013; Haemmerlé et al. 2018a), whose internal structure consists of a convective core produced by H-burning, a radiative zone containing most of the stellar mass, and a convective envelope that covers a dominant fraction of the photospheric radius. These structures can be approximated by ‘hylotropes’ (Begelman 2010), in particular for rates $\gtrsim 10^2 M_\odot$ yr$^{-1}$, at which the contraction is adiabatic (Haemmerlé et al. 2019). The hydrodynamical simulations of Mayer et al. (2010, 2015) show accretion rates up to $\sim 10^5 M_\odot$ yr$^{-1}$ down to a fraction of a pc during the assembly of the SMD. Gravitational free-fall can maintain such rates down to the stellar surface only if its mass exceeds $\sim 10^5 M_\odot$. Milder rates up to $\sim 1000 M_\odot$ yr$^{-1}$ can be driven if the attractor’s mass exceeds $\sim 1000 M_\odot$ (Haemmerlé et al. 2021), which is comparable to our estimates for the mass of the initial hydrostatic cores in SMDs. A dynamical angular momentum loss mechanism (as discussed in section 2) could supply accretion rates close to the maximum value allowed by gravitational free-fall. At such rates, masses up to $10^8 - 10^9 M_\odot$ could be reached before the final collapse, which corresponds to similar estimates as those of section 3.3.

Another important factor we neglect to model explicitly is rotation. Because of the prominent role of radiation pressure, SMSs must have surface velocities lesser than 10% of the Keplerian velocity (Haemmerlé et al. 2018b; Haemmerlé & Meynet 2019). As a consequence, the centrifugal acceleration represents at maximum a percent of the gravitational acceleration, implying that rotation has a negligible impact.

---

7 As shown in Bonoli et al. (2014), these rates are consistent with the statistics of high red-shift quasars.
on the hydrostatic structure and that deformation remains small. On the other hand, rotation impacts significantly the dynamical stability of the structure (Fowler 1966; Bisnovatyi-Kogan et al. 1967; Baumgarte & Shapiro 1999; Shibata et al. 2016; Haemmerlé 2021). The mass at which rapidly accreting SMSs reach the GR instability depends on both the accretion rate of mass and angular momentum. Without any angular momentum, reaching masses $\gtrsim 10^6 \, M_\odot$ before becoming unstable typically requires accretion rates larger than $1000 \, M_\odot$ yr$^{-1}$ (Haemmerlé 2020). However, even a small amount of angular momentum postpones this limit by orders of magnitude, and maximally rotating SMSs can remain stable potentially up to $\sim 10^9 \, M_\odot$ (Fowler 1966; Haemmerlé 2021). Since the growth of the cores described in section 3.2 results from the accretion of a rotationally supported disc, it is likely that some amount of angular momentum will be retained even in the final hydrostatic structure. Some combination of high accretion rates and retention of angular momentum is exceedingly likely in SMD cores, strongly supporting that the high final BH masses reported in Figure 4 are realistic estimates despite the simplifying assumptions.

The final issue that requires further discussion is fragmentation. As shown in MI and MII, the innermost region of SMDs (whithin $\sim 1$ pc) is stabilised by the extreme shear forces, preventing the formation of stars. However, it is possible that significant fragmentation may occur at the outer edges, where the Toomre parameter approaches its critical value. Thus, the consequence of fragmentation would be to enable the formation of a stellar disc with a typical size of a few pc. Understanding the interplay between fragmentation, migration and relaxation processes within the SMD and its stellar counterpart would require significant modelling, and would be best tackled by a purpose built numerical simulation. If the scenario lead to a significant flux of stars towards the central region, it could perhaps enable an extreme version of the stellar bombardment model proposed in Tagawa et al. (2020), in which runaway mergers of stars take over the mass and energy balance of the growing SMD core. Both the dynamics of the stellar disc and the evolution and structure of SMD cores under bombardment are extremely interesting and deserve further investigation.

5 SUMMARY & CONCLUSION

In this work we revisited the merger-driven scenario in order to assess its promise as a DCM. In Section 2, we show that the thermodynamical evolution of SMDs is fundamentally different than more typical proto-galactic discs. Firstly, SMDs are optically thick and therefore evolve adiabatically over periods comparable with the diffusion time-scale of $\lesssim 5 \times 10^5$ yr (see Eq. 3). Fragmentation is therefore naturally suppressed without the requirement of an external source of ionising radiation. Secondly, SMDs are expected to lose angular momentum on time-scales comparable or shorter than a diffusion time-scale, both in the case of dynamical as well as secular mechanisms (see Eq. 4 for the case of turbulent viscous dissipation). In combination with the adiabaticity, this implies that our analysis is independent of the choice of an angular momentum loss mechanism. We summarise the DCM as follows: As an SMD loses rotational support, more and more concentric shells accrete onto to a central spherical hydrostatic core (see Section 3.2). Every shell adds a specific amount of mass and internal energy, which determines the core’s mass-radius and temperature relation (shown in figure 3). Eventually, the cores approach the GR instability regime and collapse into a compact object, producing several observational counterparts that can be considered smoking gun signatures of the model. The main findings of our analysis are listed here:

- The mass of BHs produced in the merger-driven scenario has a typical scale of a few $10^7 \, M_\odot$, ranging from $10^6 \, M_\odot$ to $10^9 \, M_\odot$ depending on the initial disc configuration (see figure 4). SMDs are therefore optimal locations to produce fully fledged SMBHs by direct collapse. The collapse occurs very rapidly, at most $\sim 5 \times 10^5$ yr after the formation of an SMD at $z \sim 10$. A simple fitting formula for the final BH mass is provided by Eq. 31 and can be implemented as a sub-grid model in future simulations.

- The collapse of an SMD core produces a characteristic luminosity curve, shown in figure 33. It exhibits an extremely bright ($\sim 10^{-8}$ erg s$^{-1}$ cm$^{-2}$) but short ($\sim$ few s) x-ray transient. In certain cases, the electromagnetic signature might be accompanied by a burst of neutrinos ($\sim 10^{-7}$ erg s$^{-1}$ cm$^{-2}$ in flux) and of gravitational wave radiation ($\sim 10^{-21}$ in amplitude, see Eq. 38), potentially detectable by current and future instruments. If identified, these counterparts would represent a smoking gun signature of the merger driven DCM. Their expected rate is $\sim 10^{-2}$ to $\sim 1$ per year.

The merger-driven scenario has the advantage of not requiring any specific assumptions on the thermodynamics or on the angular momentum loss mechanisms of proto-galactic discs, provided that they are massive enough to behave adiabatically over long time-scales. This is a unique feature, which is at odds with other DCMs. In this paper, we have shown that once an SMD forms, it is inevitable that it will produce a large BH on a very rapid time-scale, without requiring any further additional assumption other than conventional hydrodynamics of ideal gas flow and the large-scale dynamics imposed by the merger. The trade-off is that only relatively rare major mergers of galaxies well exceeding the typical mass of star forming galaxies at $z > 10$ can form SMDs, implying that this flavour of DC must be rare by construction (see Bonoli et al. 2014; Mayer & Bonoli 2019, for an estimate of the occurence rate of such systems, and its agreement with the abundance of high-z quasars.).

Returning to the question posed in the introduction, the merger-driven scenario implies the possibility that high red-shift BHs with varying masses might form by different mechanisms. In this scenario, an uncountable number of smaller BH seeds can form early, by any of the several models present throughout the literature, including the other more conventional DC channels. These might then slowly grow by accretion or by repeated mergers to become one of the many BHs, ranging from $10^5 \, M_\odot$ to $10^8 \, M_\odot$, that we identify at the centre of galaxies in our local universe. The BH mass function at high redshift could therefore start off rather “under-weight”, and remain so up until the sudden appearance of merger-driven DC BHs at $z \sim 10$. These would form fast and late, by a completely separate mechanism from the majority. They would instantly surpass their older, tamer counterparts, becoming visible as TON618, Holmberg...
15A and many others. Incidentally, these heaviest of SMBHs tend to reside in elliptical early type galaxies, the likely outcome of the massive mergers required to form SMDs in the first place (see, e.g. Colla et al. 1970; Cattaneo & Bernardi 2003; McConnell & Ma 2013; López-Cruz et al. 2014; Yoon et al. 2019).

In conclusion, it is hard to argue which answer would be more satisfying: a single DCM, in which the most massive outliers are explained as improbable statistical flukes? Or rather the co-existence of several models, which can explain BHs in different mass ranges more naturally? Perhaps the answer must come from astronomical observation. With a new generation of detectors coming, we are certainly hopeful that a careful analysis of the redshift–dependent BH mass function, as well as the constant monitoring of the electromagnetic, neutrino and gravitational wave skies will be sufficient to untangle the question of early BH formation, be it merger-driven or otherwise.

ACKNOWLEDGEMENTS

The authors acknowledge support from the Swiss National Science Foundation under the Grant 200020_192092. We acknowledge Pedro R. Capelo, Andrea Derdzinski, and Mudit Garg for insightful discussions. LH has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 833925, project STAREX). RSK acknowledges financial support from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme.4 The Milky Way System (funding ID 50002206). We thank the Kavli Institute for Theoretical Physics at UC Santa Barbara for hospitality during the development of this work. in particular the program BINARY22 supported in part by the National Science Foundation under Grant No. NSF PHY-1748958.

DATA AVAILABILITY

Our python scripts will be shared upon reasonable request.

REFERENCES

Abbasi R., et al., 2011, A&A, 535, A109
Abel T., Bryan G. L., Norman M. L., 2002, Science, 295, 93
Abramowicz M. A., Fragile P. C., 2013, Living Reviews in Relativity, 16, 1
Adams W. S., 1925, Proceedings of the National Academy of Science, 11, 382
Agarwal B., Khochfar S., Johnson J. L., Neistein E., Dalla Vecchia C., Libio M., 2012, MNRAS, 425, 2854
Agarwal B., Smith B., Glover S., Natarajan P., Khochfar S., 2016, MNRAS, 459, 4209
Alvarez M. A., Wise J. H., Abel T., 2008, in O’Shea B. W., Heger A., eds, American Institute of Physics Conference Series Vol. 990, First Stars III. pp 432–434, doi:10.1063/1.2905657
Amaro-Seoane P., et al., 2017, arXiv e-prints, p. arXiv:1702.00786
Amaro-Seoane P., et al., 2022, arXiv e-prints, p. arXiv:2203.06016
Auclair P., et al., 2022, arXiv e-prints, p. arXiv:2204.05434
Baker J., et al., 2019, arXiv e-prints, p. arXiv:1907.06482
Bauer M. T. W., Shapiro S. L., 1999, ApJ, 526, 941
Begelman M. C., 2009, in O’Shea B. W., Heger A., eds, American Institute of Physics Conference Series Vol. 990, First Stars III. pp 489–493 (arXiv:0709.0548), doi:10.1063/1.2905669
Begelman M. C., 2010, MNRAS, 402, 673
Begelman M. C., Volonteri M., Rees M. J., 2006, MNRAS, 370, 289
Bellomo N., Bernal J. L., Racanelli A., Verde L., 2018, J. Cosmology Astropart. Phys., 2018, 004
Bisnovatyi-Kogan G. S., Zel’dovich Y. B., Novikov I. D., 1967, Soviet Astr., 11, 419
Bolton J. G., Gardner F. F., Mackey M. B., 1963, Nature, 199, 682
Bonoli S., Mayer L., Callegari S., 2014, MNRAS, 437, 1576
Cattaneo A., Bernardi M., 2003, MNRAS, 344, 45
Chandrasekhar S., 1964, ApJ, 140, 417
Chia T. T., Chau W. Y., Henriksen R. N., 1977, ApJ, 214, 576
Colgate S. A., Cen R., Li H., Currier N., Warren M. S., 2003, ApJ, 598, L7
Colla G., et al., 1970, A&A, 1, 281
Dutta S. I., Ratković S., Prakash M., 2004, Phys. Rev. D, 69, 023005
Eddington A. S., 1921, Zeitschrift fur Physik, 7, 351
Einstein A., 1916a, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin, pp 688–696
Einstein A., 1916b, Annalen der Physik, 354, 769
Eisenstein D. J., Loeb A., 1995, ApJ, 443, 11
Fiacconi D., Rossi E. M., 2016, MNRAS, 455, 2
Fowler W. A., 1966, ApJ, 144, 180
Fryer C. L., New K. C. B., 2011, Living Reviews in Relativity, 14, 1
Fujimoto M., 1968, ApJ, 152, 523
Goodman J., Binney J., 1983, MNRAS, 203, 265
Habouzit M., Volonteri M., Latif M., Dubois Y., Peirani S., 2016, MNRAS, 463, 529
Haemmerlé L., 2020, A&A, 644, A154
Haemmerlé L., 2021, A&A, 650, A204
Haemmerlé L., Meynet G., 2019, A&A, 623, L7
Haemmerlé L., Woods T. E., Klessen R. S., Heger A., Whalen D. J., 2018a, MNRAS, 474, 2757
Haemmerlé L., Woods T. E., Klessen R. S., Heger A., Whalen D. J., 2018b, ApJ, 853, L3
Haemmerlé L., Meynet G., Mayer L., Klessen R. S., Woods T. E., Heger A., 2019, A&A, 632, L2
Haemmerlé L., Mayer L., Klessen R. S., Hosokawa T., Madau P., Bromm V., 2020, Space Sci. Rev., 216, 48
Haemmerlé L., Klessen R. S., Mayer L., Zwick L., 2021, A&A, 652, L7
Haiman Z., 2013, in Wiklind T., Mobasher B., Bromm V., eds, Astrophysics and Space Science Library Vol. 396, The First Galaxies. p. 293 (arXiv:1203.6075), doi:10.1007/978-3-642-32362-1_6
Hallmann S., Clark B., Glaser C., Smith D., 2021, arXiv e-prints, p. arXiv:2107.08910
Hawley J. F., Gammie C. F., Balbus S. A., 1995, ApJ, 440, 742
Hennebelle P., 2021, A&A, 655, A3
Hirano S., Hosokawa T., Yoshida N., Umeda H., Omukai K., Chiaki G., Yorke H. W., 2014, ApJ, 781, 60
Hosokawa T., Omukai K., Yorke H. W., 2012, ApJ, 756, 93
Hosokawa T., Yorke H. W., Inayoshi K., Omukai K., Yoshida N., 2013, ApJ, 778, 178

MNRAS 000, 1–13 (2022)
This paper has been typeset from a TeX/LaTeX file prepared by the author.