Star formation in accretion discs and SMBH growth

Alexander J. Dittmann and M. Coleman Miller

Department of Astronomy and Joint Space-Science Institute, University of Maryland, College Park, MD 20742-2421, USA

Accepted 2020 February 11. Received 2020 January 23; in original form 2019 November 20

ABSTRACT

Accretion discs around active galactic nuclei (AGNs) are potentially unstable to star formation at large radii. We note that when the compact objects formed from some of these stars spiral into the central supermassive black hole (SMBH), there is no radiative feedback and therefore the accretion rate is not limited by radiation forces. Using a set of accretion disc models, we calculate the accretion rate on to the central SMBH in both gas and compact objects. We find that the time-scale for an SMBH to double in mass can decrease by factors ranging from ~0.7 to as low as ~0.1 in extreme cases, compared to gas accretion alone. Our results suggest that the formation of extremely massive black holes at high redshift may occur without prolonged super-Eddington gas accretion or very massive seed black holes. We comment on potential observational signatures as well as implications for other observations of AGNs.

Key words: accretion, accretion discs – stars: black holes – quasars: supermassive black holes – cosmology: miscellaneous.

1 INTRODUCTION

The existence of supermassive black holes (SMBHs) with masses \( \gtrsim 10^9 \, M_\odot \) at redshifts \( z \gtrsim 7 \) (e.g. Mortlock et al. 2011) challenges our understanding of the formation of SMBH seeds and their subsequent growth. The fundamental issue is that even if stellar-origin black holes with masses \( M \sim 10–100 \, M_\odot \) form from the first stars at \( z \sim 20–30 \), there is not enough time to reach \( M \sim 10^9 \, M_\odot \) at \( z \sim 7 \) if accretion is Eddington limited at a standard black hole accretion efficiency \( \eta \equiv L/(M c^2) \approx 0.1 \) for a luminosity \( L \) and an accretion rate \( M \). Quantitatively, the exponential growth time for Eddington-luminosity accretion at efficiency \( \eta \) is \( \tau \approx 4.5 \times 10^7 \, \text{yr}(\eta/0.1). \) Thus, in the \( \sim 500 \, \text{Myr} \) between \( z = 20 \) and \( z = 7 \) the growth factor is only \( \sim 6 \times 10^4 \) for \( \eta = 0.1 \). This problem is usually solved by invoking massive seed black holes from direct collapses (Begelman, Volonteri & Rees 2006), Population III stars (Bromm 2013), or mergers of stars or black holes in dense clusters (Devecchi & Volonteri 2009; Davies, Miller & Beloborodov 2011); or by supposing that gas accretion can proceed at a few times the standard Eddington rate (Toyouchi et al. 2019).

Here, we consider a different scenario, in which stellar-origin black holes in gas surrounding a massive black hole migrate inward and merge with the central black hole. These mergers produce very little radiation, and therefore the effective \( \eta \) is decreased. A decrease in \( \eta \) of only an average factor of \( \sim 0.5 \) would mean that there is enough time for a hole to grow from stellar masses to \( \sim 10^9 \, M_\odot \) by \( z \sim 7 \).

In more detail, analytical models of accretion discs often predict discs unstable to gravitational perturbations at large radii. Discs become unstable when the Toomre criterion (Toomre 1964) is satisfied, \( Q \equiv c_s \kappa \sigma_1/(\pi G \Sigma) \lesssim 1 \) for a gas disc, where \( c_s \) is the sound speed, \( \kappa \) is the radial epicyclic frequency, \( \Sigma \) is the surface density, and \( G \) is Newton’s constant. In a standard accretion disc supported by gas pressure with opacity dominated by free–free absorption (Shakura & Sunyaev 1973), \( Q \propto r^{-4/3} \), so gravitational instability is expected at large radii. If the cooling time-scale in the disc is shorter than the dynamical time-scale for the disc, gravitational instability can lead to disc fragmentation into dense objects such as stars or planets (Gammie 2001). Gravitational fragmentation has also been observed in both smoothed particle and Eulerian hydrodynamics simulations (e.g. Nayakshin, Cuadra & Springel 2007; Jiang & Goodman 2011).

Considering fragment masses and balancing stellar accretion and mass-loss, typical stellar masses in the disc range from 50 to 500 \( M_\odot \), which is unlikely to be massive enough to open a gap in the accretion disc. If individual stars and black holes are not massive enough to open gaps in the disc, they migrate due to Lindblad and corotation torques with the disc, resulting in inwards migration that can be much faster than the viscous time-scale of the gas (Tanaka, Takeuchi & Ward 2002; Paardekooper & Papaloizou 2008). Once the objects are close enough to the central SMBH, torques from gravitational radiation (Peters 1964) dominate, leading to rapid mergers regardless of gap opening. These mergers can grow an SMBH without limitation by radiation forces, facilitating faster growth than gas accretion alone.

Star formation in active galactic nucleus (AGN) discs has been investigated before (e.g. Kolykhalov & Syunyaev 1980; Levin & Beloborodov 2003; Goodman & Tan 2004; Levin 2007), although our work differs in our choice of disc model and use of updated opacities in the star-forming region of the disc. Work by Inayoshi & Haiman (2016) has explored the implications of star formation in
AGN discs on limiting maximum SMBH masses. The migration of small numbers of SMBHs has been examined by Secunda et al. (2019), and the implications of SMBH mergers in AGN discs have been investigated from a variety of angles (e.g. McKernan et al. 2012, 2014; Bartos et al. 2017; Stone, Metzger & Haiman 2017; McKernan et al. 2018; Ford & McKernan 2019; McKernan et al. 2019). We in turn study the implications of star formation on the growth of high-z SMBHs. In particular, the understanding of migration has changed significantly over the last decade (see Section 2.2), whereas many previous studies have relied on more simplistic migration prescriptions that can produce migration rates differing by orders of magnitude.

In this work, we investigate a different regime of star formation in AGN discs. We use a series of star-forming steady-state disc models closely following those described in Thompson, Quataert & Murray (2005, hereafter TQM). These models connect a star-forming outer disc to a gravitationally stable inner disc. The remainder of the paper is divided as follows: in Section 2, we summarize results that we use in our analysis, comment on various assumptions, and provide a detailed description of our disc models. In Sections 3 and 4, we present the results of our analyses and discuss implications for observations, respectively. Our conclusions are in Section 5.

2 METHODOLOGY

In this section, we present a number of results that we use throughout our analysis, as well as the details of our disc models. Because of the uncertainties in many relevant aspects of astrophysics, we must make many approximations throughout this analysis. We therefore review our choices, and emphasize that although they influence the fine details of our results, our qualitative conclusions do not depend critically on our assumptions.

2.1 Disc instability

When \( Q \lesssim 1 \), discs are thought to evolve towards stability by processes such as gravitational collapse or by transport of angular momentum via global torques such as bars or spiral waves. The immediate evolution of a gravitationally unstable disc is determined by the cooling and dynamical time-scales. If the cooling time-scale \( (t_c) \) is sufficiently short compared to the dynamical time-scale \( (t_{dynam}) \), clumps of gas may continue to collapse into dense objects (Gammie 2001) such as stars or planets. Cooling time-scales predicted by AGN disc models are usually sufficiently short to facilitate the collapse of gas into dense objects. AGN disc models contrast with those describing circumstellar discs (e.g. Cai et al. 2006), where core accretion is thought to be the dominant form of initial compact object growth (Bodenheimer & Pollack 1986). We note that \( t_c \gtrsim 50 t_{dynam} \) may be necessary to suppress disc fragmentation in the presence of turbulence, such as that caused by the magnetorotational instability (Hopkins & Christiansen 2013), so our adopted conditions for fragmentation may be overly stringent.

Although the conditions \( Q \lesssim 1 \) and \( t_c \lesssim 3t_{dynam} \) may be sufficient to determine whether discs fragment, it is not clear whether the resulting fragments are massive enough to become stars. The initial mass of fragments in the disc may be approximated roughly by the Jeans mass, \( M_J \sim c_s^2 \rho^{-1/2} \), where \( \rho \) is the disc gas density. If cooling is efficient then as the initial fragments collapse, \( c_s^2 \rho^{-1/2} \) gradually decreases. As long as the thermal adjustment time-scale is shorter than the free-fall time-scale for the fragment, the collapse is approximately isothermal. Because density increases during collapse and \( c_s^2 = b_0 T/\mu \), where \( T \) is gas temperature, \( b_0 \) is the Boltzmann constant, and \( \mu \) is the mean particle mass, one expects each clump to fragment repeatedly until the assumption of isothermal collapse breaks down. The minimum fragment mass can be changed by other physics, such as the presence of magnetic fields.

To make a rough estimate of the final fragment masses in our disc models, we apply the result of Low & Lynden-Bell (1976). The minimum Jeans mass is given by

\[
M_{\text{frag}} \approx 1.54 \times 10^{-3} T_b^{2/3} \frac{\kappa_f}{\kappa_0} M_\odot, \tag{1}
\]

where \( \kappa_0 \) is the electron scattering opacity, \( \kappa_f \) is the final opacity of the fragment, and \( T_b \) is the effective temperature of the disc in Kelvin. This is appropriate for our disc models because fragments will be bathed in thermal radiation from the disc, and the discs in our models are optically thick to their own radiation. In the regions of our discs where we expect star formation, temperatures are usually between 100 and 1000 K and opacities only vary by factors of order unity. Accordingly, we assume that \( \kappa_f \) is the same as the initial opacity before collapse. Our predicted stellar masses are significantly larger than those for molecular clouds, largely because the accretion discs we consider are significantly hotter than present-day molecular clouds.

Both our predicted \( M_{\text{frag}} \) (e.g. \( \sim 150 M_\odot \) for \( T_b \sim 200 \text{ K, } \kappa \sim 1 \text{ cm}^2 \text{ g}^{-1} \) and our estimates of stellar mass based on accretion and mass-loss arguments (see Section 2.3) produce mass estimates of order \( \sim 200 M_\odot \). An upper limit on the time-scale for the collapse of protostars of mass \( m \sim 100 M_\odot \) into stars can be obtained from the Kelvin–Helmholtz time-scale (Bond, Arnett & Carr 1984; Goodman & Tan 2004)

\[
t_{KH} \approx 3300 \left( \frac{\kappa}{0.4 \text{ cm}^2 \text{ g}^{-1}} \right) \text{ yr}, \tag{2}
\]

where \( \kappa \) is the opacity of the gas in the disc, we have dropped the weak dependence on mass for stars with \( m \gtrsim 100 M_\odot \), and we use solar metallicities. This time-scale is shorter than one might think, based on less massive stars, because massive stars have smaller specific binding energies. Since we expect star formation to occur at distances from the SMBH on the order of parsecs, where \( t_{dynam} \sim 10^6 \text{ yr} \) for a \( 4 \times 10^6 M_\odot \) SMBH, \( t_{KH} \) is a few times larger than \( t_{dynam} \). We use

\[
\Omega^2 = \frac{GM_*}{r^3} + \frac{2\sigma^2}{r^2}, \tag{3}
\]

where \( M_* \) is the SMBH mass and \( \sigma \) is the stellar velocity dispersion for stars near the SMBH but not in the disc. \( \Omega^2 \sim GM_*/r^3 \) is reasonably accurate for our approximate calculations. We show in Section 3 that \( t_{KH} \) is orders of magnitude less than the characteristic time-scale for stellar growth via accretion in the outer regions of the disc, which indicates that stars do not accumulate significant additional mass before beginning fusion on the main sequence.

2.2 Migration

SMBHs in AGN discs migrate analogously to planets in circumstellar discs. Our disc models have opening angles \( H/r \sim 0.01–0.2 \). We assume that stars and black holes form in almost circular orbits with inclinations \( i \sim H/r \) and find that stellar masses are less than \( \sim 10^6 M_\odot \). Then, \( (e^2) + (i^2) \gtrsim (2m/M_\odot x)^{2/3} \), where \( m \) is the mass of the star, \( e \) is the orbital eccentricity, and \( M_\odot \) is the mass of the central SMBH, \( 4 \times 10^6 M_\odot \), so the collisional evolution of compact objects in the disc is dispersion dominated (Rafikov &
Slepian (2010). In this regime, eccentricities and inclinations follow a Rayleigh distribution (Ida & Makino 1992)
\[
f(e^2, i^2) \propto \frac{1}{(e^2)(i^2)} \exp \left(-\frac{e^2}{(e^2)} - \frac{i^2}{(i^2)}\right),
\]
which exponentially suppresses highly inclined or eccentric orbits as the system evolves. The assumption that \( i \sim Hr \) may be an overestimate if star formation occur preferentially towards the mid-plane. In either case, it is reasonable even in the absence of damping to assume that orbits are initially nearly circular and coplanar.

Collisions between compact objects excite eccentricities and inclinations over time (Stewart & Ida 2000). However, orbits passing through the gas disc excite waves which damp orbital eccentricities and inclinations (Tanaka & Ward 2004). Cresswell & Nelson (2008) investigated the competition between these two effects using hydrodynamic and N-body simulations of up to eight protoplanets, and concluded that damping forces dominate collisions, leading to nearly coplanar and circular orbits. However, this simulation included a relatively small number of objects, and the conclusion may not hold for systems including more objects, in which the frequency of interactions increases dramatically. For this work, we assume that this result also holds for much larger numbers of disc-embedded objects.

It is not clear if increases in inclination and eccentricity over time due to close gravitational encounters would significantly change migration timescales, especially if the overall growth time-scale for \( i \) or \( e \) is large compared to migration timescales. However, Papaloizou & Larwood (2000) and Cresswell & Nelson (2008) find that non-zero eccentricity and inclination can decrease inward migration rates, even reversing the sign of the torque for highly eccentric orbits. Larger eccentricities also lead to greater gravitational radiation torques, so it is not clear how overall migration time-scales would be affected. Eccentricity and inclination damping become weaker near the SMBH because the damping time-scale depends more strongly on radius than the dynamical time-scale, \( \tau_{\text{damp}}/\tau_{\text{dyn}} \propto r^{-2} \) for constant \( Hr \) (Tanaka & Ward 2004), which may lead to significant eccentricities in the regime where gravitational radiation torques dominate.

We consider multiple torques that lead to migration of compact objects through the disc. Regardless of disc structure, compact objects orbiting the SMBH will lose angular momentum to gravitational radiation at an average rate (Peters 1964)
\[
\Gamma_{\text{GW}} = -\frac{32}{5} \frac{G^7 m^2 M_\odot^2 (m + M_\odot)^{1/2}}{e^5 q^{7/2}(1 - e^2)^2} \left(1 + \frac{7}{8} e^2\right).
\]

Torques on the migrating object from the disc depend on whether the object is massive enough to open a gap. The criterion for gap opening is approximately given by \( g \lesssim 1 \), where
\[
g = \frac{3}{4} \frac{H}{r} \left(\frac{q}{3}\right)^{-1/3} + 50 \frac{a}{q} \left(\frac{H}{r}\right)^2,
\]
where \( a \) is the usual viscosity parameter (Shakura & Sunyaev 1973) and \( q = \frac{M}{M_\odot} \). Here, the first term represents the balance between the size of an object’s Hill sphere and the disc scale height, and the second term represents how the gap may be prevented from opening by viscosity.

We consider different migration torques depending on whether \( g \) is greater than or less than 1, although in our disc models gap opening is relatively rare, often happening only for objects with \( q \gtrsim 10^{-4} - 10^{-3} \) depending on the disc model (Fig. 5). Throughout most of the disc the first term in equation (6) is less than unity, so in practice the viscosity often plays the deciding role. We denote the minimum mass at a given radius which is able to open a gap as the isolation mass \( M_{\text{iso}} \), which is calculated by solving equation (6) and setting \( g = 1 \).

When no gap is present, we calculate a migration torque using the expressions for Lindblad and non-linear corotation torques found by Paardekooper et al. (2010). For azimuthally isothermal discs, the normalized torque is
\[
\frac{\Gamma_{\text{iso}}}{\Gamma_0} = 1.1 \psi + 0.9 \delta - 2.5,
\]
where
\[
\psi = -\frac{d}{dr} \ln \Omega, \quad \delta = -\frac{d}{dr} \ln \Sigma,
\]
\( \Sigma \) is the disc surface density and \( \Gamma_0 = \left(qH^2\Sigma r^2\Omega^2\right) \). Note that this notation is slightly different than used by Paardekooper et al. (2010), because we use \( \psi \) to refer to disc viscosity and we assume a rotation profile with deviations from Keplerian motion. For azimuthally adiabatic discs with adiabatic index \( \gamma \), the normalized torque is given by
\[
\frac{\Gamma_{\text{ad}}}{\Gamma_0} = 1.1 \psi - 1.7 \delta - 7.9 \left(\frac{\gamma - 1}{\gamma}\right) - 2.5.
\]

When applying these formulae, we use \( \gamma = 5/3 \), and interpolate between the two expressions following Lyra, Paardekooper & Mac Low (2010) to find the total gas torque
\[
\Gamma_g = \frac{\Gamma_{\text{iso}} + \Gamma_{\text{ad}}}{\Theta + 1} \psi,
\]
where
\[
\Theta = \frac{c_{\text{s}} \Sigma \Omega / \kappa}{12 \pi \sigma_\alpha T^3}
\]
relates the dynamical and cooling time-scales, \( c_{\text{s}} = 1.5k_B \sigma_\alpha \) is the Stefan–Boltzmann constant, and \( \tau_\varepsilon \) is the effective optical depth at the disc mid-plane. From Hubeny (1990)
\[
\tau_\varepsilon = \frac{3}{8} + \frac{\sqrt{3}}{4} + \frac{1}{4 \tau_r}.
\]

The optical depth is \( \tau = \kappa \rho H \). In principle, the total gas torque could be positive in some regions of the disc, which would push objects outwards and thus prevent inwards migration (Lyra et al. 2010). Bellovary et al. (2016) identify these ‘migration traps’ in some models of AGN discs. For each of our models, we have verified that the total torque, \( \Gamma_\text{tot} = \Gamma_g + \Gamma_{\text{GW}} \) on migrating objects is never positive, and thus migration only proceeds inwards. We plot the total torque on a migrating object with mass \( m = 10M_\odot \) as a function of radius in Fig. 1.

If a gap opens in the disc, migration changes significantly. Historically, gas was believed to be unable to cross gaps in the disc, and migration was thought to proceed at the rate of viscous gas inflow. However, numerous recent numerical studies have demonstrated that gas is able to flow through gaps on horseshoe orbits (Duffell et al. 2014; Dürmann & Kley 2015), and that migration is tied to the density of gas in the gap (Kanagawa, Tanaka & Szuszkiewicz 2018). In situations where gap opening occurs, we use the approximate radial migration formula of Kanagawa et al. (2018) in terms of the viscous inflow velocity \( v_{\text{vis}} = -\psi a H^2 \Omega / R \)
\[
v = 100 v_{\text{vis}} \Sigma r^2 / m.
\]
We find that when gaps open at radii where gravitational radiation torques are insignificant, migration timescales increase by orders of magnitude, as shown in Fig. 5.

MNRAS 493, 3732–3743 (2020)
too small to affect disc structure in major ways, we approximate accretion on to stars by the Bondi accretion rate
\[ m_B \approx \pi \rho G^2 m^2/c_s^3. \]  
\[ (14) \]

We approximate the time required for an object to double in mass via Bondi accretion as \( t_{\text{bind}} \approx c_s^2/(2\pi G^2 m \rho) \). Because in a quasi-stable disc with \( Q \sim 1 \), gas density decreases sharply with increasing distance from the SMBH, causing \( m_B \) to become quite small in the outer regions of the accretion disc. However, at radii within a parsec from the SMBH, the doubling time-scale becomes comparable to the dynamical time-scale.

Since the Bondi accretion rate is proportional to \( m^2 \), this accretion rate can become unphysically large as mass increases. In such cases, we assume that the accretion rate on to stars is limited such that the accretion luminosity, \( L_{\text{acc}} = \rho m_{\text{acc}}/r \) (for stellar radius \( r_s \)) is less than the Eddington luminosity, \( L_{\text{edd}} \approx 1.3 \times 10^{38}(m/M_\odot) \) erg s\(^{-1}\). Thus, when the opacity is dominated by electron scattering, the accretion rate is regulated to be less than
\[ \dot{m}_{\text{edd}} \approx 10^{-3} r_s r_\odot M_\odot \text{ yr}^{-1} \]  
\[ (15) \]

(see Artyomowicz, Lin & Wampler 1993 for a similar treatment). By following the procedure in Bond et al. (1984) to determine the radius of a massive star, using solar metallicity, we find that \( r_s r_\odot \) can be approximated well by \( r_s r_\odot = 0.56(m/M_\odot)^{0.51} \) in the range 100–1000 M\(_\odot\). We can thus approximate the Eddington-limited accretion rate on to stars as
\[ \dot{m}_{\text{edd}} \approx 5.6 \times 10^{-4} \left( \frac{0.4 \text{ cm}^2 \text{ g}^{-1}}{\kappa} \right) \left( \frac{m}{M_\odot} \right)^{0.51} M_\odot \text{ yr}^{-1}. \]  
\[ (16) \]

Bondi accretion itself can be modified by the stellar radiation. For example, stellar evolution and radiation hydrodynamics simulations have shown that ultraviolet feedback from accreting stars can halt growth at a few 10\(^3\) M\(_\odot\), or result in intermittent periods of strong mass-loss and accretion for stars with seed masses near \( \sim 10^3 M_\odot \) (Hosokawa et al. 2016). If stars are able to accrete sufficiently quickly compared to their thermal adjustment time-scale, their radii can increase dramatically as their effective temperature drops, facilitating stellar masses above 10\(^3\) M\(_\odot\) (Schleicher et al. 2013); we expect other factors such as the angular momentum of accreting gas to prevent runaway growth by this mechanism, as discussed later. With these considerations in mind, our first-order treatment considers spherically symmetric mass-loss and accretion.

Considering feedback on Bondi accretion from a central radiating source, the force on infalling matter is reduced by a factor of \((1 - f)\), where \( f \) is the ratio of local flux to the Eddington flux. If the accretion flow is spherically symmetric and optically thin out to the Bondi radius, then the reduction factor is \( (1 - \Gamma) \), where \( \Gamma \) is the ratio of the stellar luminosity to the Eddington luminosity. We approximate the results of Bond et al. (1984) according to a power law
\[ (1 - \Gamma) \approx 4.2(m/M_\odot)^{-0.43}. \]  
\[ (17) \]

Because mass is conserved, \( m \propto r_B^2 \rho u \), where \( u \) is the speed of the inflow at large distances and \( r_B \propto m \) is the radius inside of which gravity controls the flow. For this purpose, the factor \((1 - \Gamma)\) acts to reduce the effective mass of the object, and therefore the accretion rate becomes
\[ \dot{m}_B \approx \pi \rho G^2 m^2(1 - \Gamma)^2/c_s^3. \]  
\[ (18) \]

2.3 Stellar accretion and winds

We assume that stars form in discs as described in Section 2.1. The outer regions of our disc models have fairly large opening angles, \( H/r = 0.2 \), so the mass ratio required to open a gap is quite large. We verify this in Section 3. Since stellar masses are, at least initially,
scale as $r^{-2}$, we expect $\Gamma$ to be nearly constant throughout the entire Bondi region in the outer disc. However, as $\rho$ and $\tau_B$ increase at radii closer to the SMBH, $\tau_B$ can be greater than unity depending on the opacity of the disc. At smaller radii, differential rotation of the disc at either side of a star’s Bondi radius can also reduce the Bondi accretion rate. Moreover, if accretion on to stars deviates significantly from spherical symmetry, the impact of this mechanism may be reduced. However, we suggest that in these scenarios, stellar masses will be limited by enhanced mass-loss.

We can estimate mass-loss rates of O stars using the empirically calibrated result by Lamers & Leitherer (1993)

$$\log(M) = 1.738 \log(L) - 1.352 \log(T_{\text{eff}}) - 9.547,$$

where $M$ is measured in $M_\odot$ yr$^{-1}$, $L$ is measured in $L_\odot$, and $T_{\text{eff}}$ is measured in Kelvin. We use the result from Bond et al. (1984) that very massive Population I and II stars have effective temperatures of $\sim 10^4$ K, almost independent of mass. Then, we approximate $L/L_\odot \approx 5.5 \times 10^3 (m/M_\odot)^{1.2}$ in the range 100–1000 $M_\odot$ based on the expression for $L_{\text{edd}}$ in Bond et al. (1984). This enables us to estimate the mass-loss rate as

$$M_{\text{loss}} \approx 3.1 \times 10^{-10} (m/M_\odot)^{2.1} M_\odot \text{ yr}^{-1}. \quad (20)$$

Additionally, stellar rotation can strongly enhance mass-loss via winds. We expect that all stars born inside the disc will have some rotational angular momentum, potentially enhancing these mass-loss rates. The outer portions of our disc models have a large opening angle $H/\tau \sim 0.2$, so we do not expect mass-loss to change by factors of more than order unity. However, the thinner inner star-forming regions of the disc have $H/r \sim 0.01$. Considering the Hill radius of migrating objects, $r_H = (m/3M_\odot)^{1/3}$, and a fiducial black hole mass of $4 \times 10^6 M_\odot$, objects in the disc with mass greater than $\sim 10 M_\odot$ would experience primarily apsidal accretion. Additionally, as distance from the SMBH decreases, the difference in velocity between the outward and inward (with respect to the SMBH) Bondi radius of the migrating black hole increases due to increases in both the disc angular velocity and the Bondi radius.

It is necessary to understand the balance of accretion and mass-loss to make good estimates of the mass of stars at the end of their lives. We can approximate the final masses of stars in the disc by balancing these processes. We calculate the equilibrium mass as the mass for which equations (20) and either (16) or (18) balance. In practice, attempting to solve directly for mass using equations (18) and (20) can result in a mass lower than was assumed in writing these equations, so in these cases we expect equilibrium masses to be between 50 and 100 $M_\odot$.

However, our estimates for final masses are likely overestimates, since we have neglected enhanced mass-loss rates towards the end of stellar lifetimes, and also neglected mass-loss due to pulsational instability in stars with $m \gtrsim 100 M_\odot$. Mass-loss driven by pulsations can range from $\sim 10^{-6}$ to $10^{-4} M_\odot$ yr$^{-1}$ for stars with masses 100 $M_\odot \lesssim m \lesssim 200 M_\odot$ (Appenzeller 1970; Papaloizou 1973).

The aforementioned mass-loss rates do not depend explicitly on the angular momentum of accreting gas, but accretion on to stars in the disc is modified by shear motion of the gas disc relative to the star. The relative velocity is approximately $v_{\text{rel}} \sim c_s (rH)/(m M_\odot)^{1/3}$. This can lead to significant angular momentum accumulation for higher stellar masses and thinner discs. For less massive stars and those in the outer regions of the disc where $(H/r) \sim 0.1$, this leads to a correction factor of order unity in the denominator of equation (14). For more massive stars ($m \sim 10^3 M_\odot$) and thinner disc regions ($H/r \sim 10^{-2}$), the Bondi accretion rate can be reduced by factors near $\sim 10^2$.

Compounding this, Maeder & Meynet (2000) found analytically that mass-loss rates can become asymptotically large for stars with moderate rotation and $\Gamma \gtrsim 0.64$. Thus, using equation (17), we expect maximum masses ranging from 300 to 500 $M_\odot$, if our other considerations would predict larger masses. Balancing equations (20) and (16) informs us that the equilibrium Eddington-limited stellar mass is

$$m_{\text{edd}}/M_\odot \approx \left( 1.8 \times 10^6 \left( \frac{0.4 \text{ cm}^2 \text{s}^{-1}}{\kappa} \right) \right)^{1/1.59} \sim 10^4,$$

so we expect stars to become limited by violent mass-loss before they approach the Eddington limit.

We note that some regions of the disc, the final stellar masses are in the expected range for pair-instability supernovae. For example, Chatzopoulos & Wheeler (2012) find that pair-instability supernovae occur for stars as small as 65 $M_\odot$, although core collapse also occurred in some of their models at each mass until 80 $M_\odot$, after which either pair-instability supernovae or pulsational pair-instability supernovae are possible. However, Fryer, Woosley & Heger (2001) found that the pair-instability explosion was unable to unbind their 300 $M_\odot$ model, resulting in black hole formation. Thus, we expect some violent transients from the stars formed in our disc models, but expect that many black holes will be left behind.

### 2.4 Disc model

In our investigation, we consider the TQM accretion disc models along with the modifications to the opacity prescription described below. These models connect an outer accretion disc, where radiation pressure from massive stars or accretion on to black holes provides the support necessary to maintain $Q \sim 1$, to an inner gravitationally stable $\alpha$-disc. We find this model advantageous because it provides a straightforward way to track the gas lost from the disc during star formation. Additionally, since the model includes the effects of irradiation on disc structure, this may lead to more realistic inferences of initial fragment masses, accretion rates, and migration rates. As long as irradiation does not increase by orders of magnitude once objects migrate into the gravitationally stable region, the pressure support of the inner disc is not significantly modified by the migration of objects through it.

Note that TQM consider an $\alpha$-viscosity and increased opacities when applying the model to disc around an SMBH similar to Sagittarius A*, but use a constant radial Mach number prescription to model luminous infrared galaxies and more massive AGNs; the latter prescription reduces the star formation necessary to maintain disc stability (TQM). It is probable that a combination of local and global torques act in realistic AGNs. We attempt to address this a posteriori in Sections 3 and 4.

The TQM disc models have a number of free parameters. For our purposes, the most important are the viscosity parameter $\alpha$, the SMBH mass, the accretion rate at the outer boundary of the disc ($M_{\text{out}}$), and the efficiency with which rest mass energy from star formation is converted into radiation ($\epsilon$). We consider accretion on to a $4 \times 10^6 M_\odot$ SMBH. In principle, discs around lower mass black holes can also be gravitationally unstable, but we choose not to investigate these since the relation between central black hole mass and stellar bulge velocity dispersion is poorly measured for black holes less massive than $\sim 10^6 M_\odot$ (Gültekin et al. 2009). On the other side of the spectrum, accretion discs around more massive black holes are more susceptible to gravitational instability (see e.g. Inayoshi & Haiman 2016), although because we are concerned with
SMBH growth in the early Universe, it is pertinent to study lower mass SMBHs.

We assume a supernova feedback parameter $\xi = 1$, following TQM. Here, $\xi$ is a dimensionless parameter representing non-radiative pressure support due to feedback that is independent of optical depth. In principle, extreme mass-loss from stars or SMBH-driven outflows could increase this value significantly, causing $\xi = 1$ to overestimate the star formation rate and underestimate the gas accretion rate on to the SMBH. Since the input physics is highly uncertain, we parametrize this by considering how our results change if we have incorrectly estimated accretion rates in this manner. We also assume a stellar velocity dispersion $\sigma = 180(M_*/2 \times 10^5 M_\odot)^{0.23} \text{ km s}^{-1}$ (Kormendy & Ho 2013). It is unlikely that this empirical relation between velocity dispersion and SMBH mass holds precisely at high redshift. However, this should be sufficiently accurate, since our results do not depend qualitatively on this parameter.

We chose the outer radius, at which we set the $M_{\text{min}}$ boundary condition, to be 5 pc. This radius is consistent with observations of nearby AGNs (Burtscher et al. 2013). Because the gas accretion rate is the only disc property that depends on the disc structure exterior to a given radius, we have the freedom to excise, a posteriori, regions of the disc exterior to any given radius. Note that for some of our models, the disc can be more massive than the SMBH. However, in the most extreme cases the difference between disc mass and SMBH mass is only a factor of order unity, and our results do not depend strongly on the precise edge radius. Similarly, we neglect the mass of migrating compact objects in the disc, since this depends on disc properties both within and outside of a given radius. Additionally, if migration speeds are large, high accretion rates in compact objects can be realized with little compact object mass in the disc.

The TQM models originally used the dust and gas opacity tables from Semenov et al. (2003). Since this opacity table only extends to $10^4$ K, we smoothly connect it to the OPAL opacity tables (Iglesias & Rogers 1996) for our chosen metallicity (solar) in case temperatures exceed $10^4$ K. We use these tabulated opacities in the star-forming regions, but adopt an approximate combination of Kramers’ free–free and bound–free opacities with electron scattering opacity in the inner stable region of the disc

$$\kappa = \kappa_{\text{es}} + 4 \times 10^{25}(1 + X)(Z + 0.001)\rho T^{-7/2},$$

(22) given in cm$^2$ g$^{-1}$, and where $X = 0.7381$ and $Z = 0.0134$ are the hydrogen and metal mass fractions, respectively (Asplund et al. 2009), and $\kappa_{\text{es}} = 0.2(1 + X)$ is the electron scattering opacity. Our reasons for this choice are illustrated by a comparison between our Fig. 1, which displays disc parameters from a pair of our models, to the analogous fig. 6 in TQM. We find that using the Semenov et al. (2003) opacity tables for the inner regions of the disc leads to huge discontinuities in model parameters. These discontinuities can be multiple orders of magnitude, such as $\rho$ in TQM fig. 6, and are accompanied by sudden changes in gradients. We do not consider these changes physically meaningful; instead they are the result of applying realistic opacities to a simplified 1D disc model. It is possible that a disc model including vertical structure and radiative transport would ameliorate these issues. We find that by using a simplified opacity model, we limit discontinuities to the radius where we switch opacity prescriptions, and greatly reduce their impact compared to the discontinuities in TQM.

We note that Bellovary et al. (2016) analysed TQM disc models using the Semenov et al. (2003) opacities and found migration traps associated with the discontinuities in the disc models. Our prescription reduces the surface density gradients and finds no migration traps. However, Bellovary et al. (2016) also identified migration traps in the smooth Sirko & Goodman (2003) disc models, so it is plausible that migration traps may re-emerge in disc models with fully consistent opacity prescriptions.

We focus on the outer portion of the star-forming region of the disc, outside of the opacity gap region. We expect that there is insufficient time for star formation in the opacity gap region. To see this, note from equation (1) that when the opacity decreases with little change in temperature, the typical fragment mass drops by orders of magnitude to $\sim 1 M_\odot$. The Kelvin–Helmholtz time for such stars is on the order of $\sim 30$ Myr. Considering that the opacity gap region occurs at $\sim 1$ pc, the Kelvin–Helmholtz time for protostars to begin fusion can be tens of thousands of dynamical time-scales. Thus, it is likely that such protostars could be disrupted or accreted by more massive objects, which migrate through the disc much more quickly. Additionally, the doubling time-scale for these stars via Bondi accretion is again many dynamical time-scales. However, the gas density is sufficiently high in the opacity gap region that the doubling time-scale for stars via Bondi accretion is less than the migration time-scale ($t_{\text{mig}}$) for these objects. Thus, we expect that, on rare occasions, stars could grow quickly to $300–500 M_\odot$, limited by their luminosity and rotation, subsequently becoming black holes and contributing to disc structure in the same way as the stars that became black holes before entering the opacity gap region.

The other parameter of interest is the feedback parameter $\epsilon$. The TQM models assume local feedback, so it is useful to check if the model is compatible with migrating objects. The photon diffusion time can be approximated as

$$t_{\text{diff}} \approx h\tau/c \approx 3.26\tau(H/pc) \text{ yr.}$$

(23)

For migration time-scales on the order of $\sim$Myr, we find that the photon diffusion time-scale is much shorter than migration time-scales throughout our disc models. For example, in Fig. 1 we see that the largest $\tau$ is $\sim 10^5$ and the largest scale height is $\sim 1$ pc, which occur at very different radii. Thus, the TQM feedback model is applicable to discs supported by feedback from migrating objects, at least to first order.

We assume that irradiation in the disc is primarily from accretion on to black holes, and that the gas supply to the migrating black holes is determined by local disc parameters. This leaves our disc models independent of uncertainties in stellar accretion processes. We investigate both $\epsilon = 0.1$ and $\epsilon = 0.4$ for moderately spinning and extremely spinning black holes. We assume these values of $\epsilon$ for three reasons, two physical and one practical. The first physical reason is that accretion on to black holes is far more efficient than fusion in stars. When migration time-scales through the disc are much longer than stellar evolution time-scales, which holds for most stars under our previous assumptions, an object will spend a majority of its time in the disc as a black hole rather than as a star. The second physical reason is that this interpretation of $\epsilon$ provides a straightforward way to understand why the opacity gap controls the accretion rate on to the SMBH: as opacity decreases, radiation feedback from accretion on to black holes in the disc would have comparatively little influence on gas far from the black hole, providing black holes with ample gas supply. Such a scenario may be able to feed embedded black holes at super-Eddington rates (Jiang, Stone & Davis 2014). However, a detailed investigation into this scenario would require radiation hydrodynamics simulations.

From a practical standpoint, assuming that mass lost from the gas disc goes directly into black holes simplifies mass accounting. This is because stellar mass that does not contribute directly to black holes can be thought of as simply returning to the ambient...
gas density after supernovae. Additionally, small values of $\epsilon$ lead to enormous star formation rates that consume most of the gas flowing inwards via the disc. However, it is very difficult in such a scenario to supply the SMBH with gas at significant fraction of the Eddington rate, especially for low values of $\alpha$. Note that for a constant $M$, lower viscosity implies higher surface density, which implies more star formation and leaves almost no gas to fuel the AGNs. This could be partially mitigated, if we use $\epsilon \sim 10^{-3}$, by moving away from the $\alpha$-viscosity parametrization. Indeed, one expects additional torques in a mixed gas–stellar disc (Hopkins & Quataert 2011). However, we consider both $\alpha$-viscosity and $\epsilon \sim 10^{-1}$ to be reasonable stand-ins for the relevant physics.

We also assume negligible change in gas accretion rate due to accretion on to black holes in the gravitationally stable region of the disc. For the time being, this is a necessary assumption, as time-dependent disc models are beyond the scope of this work. This assumption is reasonable as far as pressure support is concerned, because higher gas densities and temperatures provide greater pressure in the stable disc regions than radiation pressure did in the unstable regions. This assumption is also justified after the fact by comparing the $\epsilon$-folding time of accretion on to black holes in the stable region of the disc to the migration time for our disc models, although this may not be the case for discs around more massive black holes. We perform this comparison in Section 3.

In order to gauge how sensitive our results are to our choices of model parameters, we consider different values of $\epsilon$, $\alpha$, and $M_{\text{out}}$. We choose $\epsilon = 0.1$ with $\alpha = \{0.2, 0.25, 0.3\}$, and $\epsilon = 0.4$ with $\alpha = \{0.02, 0.05, 0.1\}$. For each combination of $\epsilon$ and $\alpha$, we choose one or two values of $M_{\text{out}}$ to investigate how accretion rate, and thus disc thickness, affect migration time-scales and other results while holding other parameters constant. The range of $\alpha$ values was chosen to be consistent with a range of thin disc models, observations, and simulations (King, Pringle & Livio 2007). When quoting accretion rates without explicit units, we have normalized by the Eddington accretion rate, $M_{\text{Edd}} = 4\pi GM_r m_h/\eta \sigma_T c$, with $\eta = 0.1$, where $m_h$ is the mass of atomic hydrogen.

2.5 Summary

Our process begins with a modified version of the TQM disc models, where we have used opacities valid over a wider range of temperatures to reduce discontinuities in the model; supposed that feedback occurs primarily by accretion of gas on to black holes in the disc; and then checked using many previous results that our overall picture is self-consistent. Our approach is limited in a number of ways. For example, we explicitly require that the discs have $Q = 1$. Additionally, our model is agnostic with respect to the physical processes regulating accretion on to black holes embedded in the disc, applying the ansatz that precisely enough gas accretes on to embedded black holes to maintain $Q = 1$. Similarly, there are many uncertainties and approximations associated with the migration, feedback, and accretion prescriptions that we employ. Because of these uncertainties and others, we do not attempt to create a disc model with completely self-consistent physics.

We chose to investigate these disc models in order to assess the possibility of SMBH growth in the early Universe through mergers with compact objects formed within their accretion discs. Because the models were constructed with this in mind, they have Eddington ratios larger than are considered typical for modern AGNs. Accordingly, our disc models have larger scale heights at a given radius than most models of lower redshift AGNs. Thinner discs would lead to both enhanced gap opening and faster migration in the absence of gap opening. Thus, we expect that the results presented here are not applicable to more slowly accreting AGNs.

3 RESULTS

Fig. 1 depicts a number of disc parameters, both intrinsic and derived, over the full range of radii in our models. Note the ‘opacity gap’ near temperatures of $10^3$ K corresponding to the sublimation of dust grains. This reduction in disc opacity means that embedded black holes can be fed at super-Eddington rates more easily, primarily limited by the higher opacity of gas as it accretes rather than the opacity of the disc. We present accretion rates in Fig. 2. Considering the bottom left panel of Fig. 1, it is evident that the migration of objects is always inward, regardless of the mass of the object, so migration traps will not occur in any of the discs we consider. Note the small dips in opacity in regions where $T \sim 100–1000$ K, which occur as ice, volatile organics, or minerals such as troilite evaporate (Semenov et al. 2003).

In Fig. 2, we present both the gas accretion rate $M_{\text{gas}}$ and the black hole growth rate $M_\ast = \pi \Sigma r^2$ for one of our models, where $\Sigma$ is the accretion rate on to black holes in the disc. The other models are qualitatively similar, usually with slight differences in the location of the opacity gap or the magnitude of each curve. We chose this particular model because the change in $M_{\text{gas}}$ is easily visible, and we calculate the accretion rate in black holes on to the central SMBH as the difference between $M_{\text{gas}}$ and the value of $M_{\text{gas}}$ at the SMBH. It is clear that the star formation in the outer gap has little effect on the overall mass flowing into the AGNs. However, these black holes are able to grow significantly in the mass gap. We note that the migration time is proportional to $m^{-1}$, so as mass is accreted on to black holes the inflowing mass constituted by black holes increases quadratically in the absence of gap opening.

We present sample characteristic masses in the star-forming region for the same disc model in Fig. 3. These include the isolation mass, the initial fragment mass, and the equilibrium mass balancing Bondi-limited accretion and losses due to stellar winds. In the outer accretion disc, stellar masses are limited to the mass at which significant mass-loss starts, likely between 50 and 100 $M_\odot$. We find that instead of being limited by the Eddington accretion rate, stars will likely be limited to masses below $\sim 300–500 M_\odot$ by mass-loss enhancements from rotation at high Eddington ratios.

Figure 2. Accretion rates in the star-forming region of our $\alpha = 0.05$, $M_{\text{out}} = 3.0$ disc model. The blue solid line plots the gas accretion rate through the disc, illustrating the change in gas accretion inwards towards the SMBH due to star formation and accretion on to SMBHs. The orange dashed line plots the growth rate of black holes in the disc $M_\ast$. 

MNRAS 493, 3732–3743 (2020)
Figure 3. Characteristic masses in the star-forming region of our $\alpha = 0.05$, $M_{\text{out}} = 3.0$ disc model. The blue solid line plots the isolation mass, the mass such that $g \lesssim 1$ in equation (6). The orange dashed line plots the fragment mass given by equation (1). The dash–dotted line plots the equilibrium mass such that $g \approx \frac{\dot{M}}{\dot{M}}$. We expect other mass-loss mechanisms to limit stellar masses to $\lesssim 500 M_\odot$. Other disc models are qualitatively similar to the one presented here.

Figure 4. Characteristic time-scales in the star-forming region of our $\alpha = 0.05$, $M_{\text{out}} = 3.0$ disc model. The blue solid line indicates the Kelvin–Helmholtz time-scale for fragments with mass given by equation (1). The orange dashed line indicates the cooling time-scale in the disc. The dotted purple line indicates the Bondi doubling time-scale of a fragment with mass given by equation (1). The turquoise dash–dotted line marks the dynamical time-scale. The same hierarchy of time-scales is shared between the disc for two disc models. We choose disc models with radically different viscosities, to illustrate the effects of gap opening, as well as to demonstrate some of the effects of disc thickness. Recall from equations (7), (9), and the expression for $\Gamma_0$ that thicker discs lead to slower migration. We separate migration time into two parts: the time to migrate through the inner stable region of the disc to the SMBH, and the time to migrate from the outer edge of the disc to the inner region. There are two trends for each disc model: as expected, migration time is inversely proportional to mass, up to the point where objects become massive enough to open gaps. The migration time-scales in the outer disc are long enough for stars to become black holes and subsequently accrete. As we expect the lowest mass objects in the disc to be around $100 M_\odot$, almost all of the objects formed in the disc are able to migrate into the SMBH quickly enough to contribute to SMBH growth at high redshift. Additionally, almost all stars born in these discs should have time to evolve into black holes before reaching the SMBH.

We must verify that various time-scales relating to stellar accretion and evolution are self-consistent, as well as reasonable in the context of the stellar masses and accretion rates that we have predicted. Notably, all of the time-scales we consider in Fig. 4 are shorter than the main-sequence lifetime of massive stars, $\sim (2–3) \times 10^6$ yr (Bond et al. 1984), except for the doubling time-scale at very large radii. Additionally, it is clear that protostars require a few orbits to reach the main sequence, but do not gain appreciable mass in this time except towards the inner edge of the star-forming region. In most cases, fusion and therefore feedback processes can begin before accretion becomes significant. Additionally, the cooling time-scale in the disc is much shorter than the dynamical time-scale, as expected for a disc unstable to gravitational fragmentation.

In Fig. 5, we present the migration times of objects through the disc for two disc models. We choose disc models with radically different viscosity and material content, to illustrate the effects of gap opening, as well as to demonstrate some of the effects of disc thickness. Recall from equations (7), (9), and the expression for $\Gamma_0$ that thicker discs lead to slower migration. We separate migration time into two parts: the time to migrate from the marginally stable disc region to the central SMBH. Orange triangles correspond to the model with $\alpha = 0.02$, $M_{\text{out}} = 7.0$ and blue squares correspond to the model with $\alpha = 0.3$, $M_{\text{out}} = 3.0$. The large jumps in migration time correspond to masses that are able to open gaps in the disc.

Figure 5. Migration times for different migrating masses and different disc models. Hollow symbols represent the time required to migrate from the out to the inner edge of the unstable region of the disc, inside and outside the opacity gap region.

It is not practical to present plots of each parameter of interest for every disc model. However, we collect parameters of interest in Table 1. $M_{\text{frag, gap}}$ and $M_{\text{frag}}$ are the median fragment masses given by equation (1) in the unstable region of the disc, inside and outside the opacity gap, respectively. $M_{\text{frag, gap}}$ provides a decent sense for fragment masses in the opacity gap region. Because the Kelvin–Helmholtz times for these lower mass protostars are many times longer than those given by equation (2), collapse can take hundreds of dynamical times, making star formation in the opacity gap region unlikely. On the other hand, the fragment masses $M_{\text{frag}}$ are fairly consistent between disc models. Although we do not expect star formation in the opacity gap region to alter disc structure significantly, stars that migrate into this region might
grow to 300–500 $M_\odot$, limited by rotation-enhanced winds. Such stars may collapse directly into black holes, unbinding little of their mass (Fryer et al. 2001). In lower $\alpha$ models, such black holes may open gaps at large distances from the SMBH, $\gtrsim 10^4 R_G$, precluding migration via gas or gravitational radiation torques.

In order to understand the relevant migration time-scales, we track the total migration times for objects with constant masses of 100 $M_\odot$, $\tau_{100}$ and 500 $M_\odot$, $\tau_{500}$. To calculate this quantity, we integrate $r/\dot{r}$ across our disc models. Since we have verified that the total torque never changes sign, this is an appropriate treatment. It is important to note how the migration time-scale changes depending on object mass. For example, if the characteristic mass of migrating objects tends towards $\sim 500 M_\odot$ or higher after accretion, these objects would open gaps, slowing down significantly. Because star formation can occur at any radius of the star-forming region, the hollow symbols in Fig. 5 give an upper limit on the inward migration time for stars formed in the disc. Note that in general, $\sim 10 M_\odot$ stars should have more than enough time between $z \sim 20$ and $z \sim 7$ to reach the SMBH from the outer edge of the disc in any of our models.

Our method for calculating migration time-scales when gap opening occurs likely results in an overestimate. Looking at the $H/\dot{r}$ panel of Fig. 1 and recalling the dependence of the gap-opening criterion on $H/\dot{r}$, it is clear that once a gap-opening object migrates inward, it will eventually be unable to open a gap and resume migrating more quickly. Gravitational interactions with other migrating objects may accelerate the rate at which gap-opening objects resume type-I migration. Consider a black hole massive enough to open a gap orbiting the SMBH, migrating inwards so slowly it can be considered to have a constant semimajor axis: another black hole can then migrate towards it. In analogy to the hardening of binaries by three-body encounters, the gap-opening object can move to a lower energy orbit, transferring energy to the lighter object. Following this reasoning, a straightforward application of equation (13) may significantly overestimate the amount of time required for objects to migrate towards the SMBH is cases where gap opening occurs. Another possibility is that black holes in the disc can merge to alleviate the problem of large numbers of black holes, as each merger increases the accretion rate on to the SMBH, since $\tau \propto 1/m$. Mergers may be fairly common in the inner regions of the disc, where the number density of black holes is higher. Since we consider black holes that originate in the accretion disc, these black holes should have low relative speeds as they pass each other, which could facilitate merging. Excitation of orbital eccentricities also accelerates migration once gravitational radiation becomes a significant torque on migrating objects.

We calculate $\Delta$, the change in accretion efficiency due to the presence of black holes in the accretion disc. Here,

$$\Delta = \frac{\min(M_{\text{gas},1.0})}{\min(M_{\text{gas},1.0}) + M_{\text{bh}}}.$$  

since only this fraction of the total accretion rate results in the emission of photons. We define $\Delta_n$ by multiplying $M_{\text{bh}}$ by $1/n$ to evaluate how our estimate changes if we have overestimated the accretion rate of black holes on to the SMBH by a factor of $n$. This could occur, for example, if our assumptions of zero eccentricity and

<table>
<thead>
<tr>
<th>$\epsilon$</th>
<th>$\alpha$</th>
<th>$M_{\text{gas},100}$</th>
<th>$M_{\text{frag},100}$</th>
<th>$M_{\text{frag}}$</th>
<th>$\tau_{100}$</th>
<th>$\tau_{500}$</th>
<th>$M_{\text{gas}}$</th>
<th>$M_{\text{bh}}$</th>
<th>$\Delta$</th>
<th>$\Delta_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>0.02</td>
<td>7.0</td>
<td>169.72</td>
<td>435.85</td>
<td>31.51</td>
<td>949.88</td>
<td>1.29</td>
<td>5.71</td>
<td>0.15</td>
<td>0.47</td>
</tr>
<tr>
<td>0.4</td>
<td>0.05</td>
<td>3.5</td>
<td>15.93</td>
<td>405.59</td>
<td>18.73</td>
<td>1117.32</td>
<td>0.79</td>
<td>2.71</td>
<td>0.23</td>
<td>0.59</td>
</tr>
<tr>
<td>0.4</td>
<td>0.05</td>
<td>3.0</td>
<td>9.11</td>
<td>402.16</td>
<td>17.06</td>
<td>4517.14</td>
<td>0.50</td>
<td>2.50</td>
<td>0.17</td>
<td>0.50</td>
</tr>
<tr>
<td>0.4</td>
<td>0.10</td>
<td>2.0</td>
<td>7.22</td>
<td>379.02</td>
<td>12.52</td>
<td>2.50</td>
<td>0.55</td>
<td>1.45</td>
<td>0.28</td>
<td>0.65</td>
</tr>
<tr>
<td>0.1</td>
<td>0.20</td>
<td>4.0</td>
<td>0.64</td>
<td>379.26</td>
<td>11.81</td>
<td>2.36</td>
<td>0.24</td>
<td>3.76</td>
<td>0.06</td>
<td>0.24</td>
</tr>
<tr>
<td>0.1</td>
<td>0.25</td>
<td>2.0</td>
<td>0.28</td>
<td>360.88</td>
<td>8.91</td>
<td>1.78</td>
<td>0.26</td>
<td>1.74</td>
<td>0.13</td>
<td>0.43</td>
</tr>
<tr>
<td>0.1</td>
<td>0.30</td>
<td>1.5</td>
<td>0.58</td>
<td>351.90</td>
<td>7.94</td>
<td>1.59</td>
<td>0.31</td>
<td>1.19</td>
<td>0.21</td>
<td>0.57</td>
</tr>
<tr>
<td>0.1</td>
<td>0.30</td>
<td>3.0</td>
<td>1.24</td>
<td>364.97</td>
<td>10.51</td>
<td>2.10</td>
<td>0.45</td>
<td>2.55</td>
<td>0.15</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Table 1. Results for each model: $\epsilon$ is the efficiency parameter, $\alpha$ is the model viscosity parameter. $M_{\text{gas}}$ is the gas accretion rate used as a boundary condition for the model at 5 pc. $M_{\text{frag,gap}}$ is the fragment mass, using equation (1), in the unstable region of the disc outside of the opacity gap. $T_{\text{frag}}$ are the total migration time-scales in Myr for objects with masses 100 and 500 $M_\odot$, $M_{\text{gas}}$ is the gas accretion rate on to the central SMBH, $M_{\text{bh}}$ is the accretion rate, in black holes, on to the central SMBH. $\Delta$ is the fraction of the total Eddington-limited accretion rate comprised by gas. $\Delta_n$ is the same quantity if the accretion rate of black holes on to the SMBH has been overestimated by a factor of 5. Tabulated masses are in $M_\odot$, time-scales in Myr, and accretion rates as a ratio to that of Eddington-limited accretion with $\eta = 0.1$. 

We calculate $\Delta$, the change in accretion efficiency due to the presence of black holes in the accretion disc. Here, 

$$\Delta = \frac{\min(M_{\text{gas},1.0})}{\min(M_{\text{gas},1.0}) + M_{\text{bh}}}.$$  

since only this fraction of the total accretion rate results in the emission of photons. We define $\Delta_n$ by multiplying $M_{\text{bh}}$ by $1/n$ to evaluate how our estimate changes if we have overestimated the accretion rate of black holes on to the SMBH by a factor of $n$. This could occur, for example, if our assumptions of zero eccentricity and...
zero inclination orbits break down before torque due to gravitational waves becomes significant, or if black holes are ejected from the system.

We have verified many of the assumptions made during the construction of our disc models. For the purposes of growing high-$z$ SMBHs, we have verified the necessary conditions that $t_{kh} < t_{mag}$ and that fragment masses in the disc are large enough to form stars as opposed to Jupiter-like objects. We find a range of potential factors that could decrease the radiative efficiency of SMBH growth. Because of the numerous unknowns in our treatment, such as the effects of feedback on migration torques and the dynamics of large number of migrating black holes in accretion discs, we consider it plausible that our values of $\Delta$ are overestimates. Even under these conservative considerations, the potential impact on high-$z$ SMBH growth is significant.

4 DISCUSSION

Let us turn to a more quantitative interpretation of our results. Consider our $\epsilon = 0.4, \alpha = 0.1$ case, which has both a sub-Eddington gas accretion rate and a comparatively mild accretion rate in black holes. In this case, a total accretion rate of about twice-Eddington is achieved even though the luminosity of the AGNs would only be about half of the Eddington limit. This could effectively reduce the SMBH e-folding time by a factor of 2.

Very optimistically, one can interpret the results of other models as reducing the e-folding time by factors as low as 0.06. Such scenarios are extreme, and could not represent SMBH growth over long periods of time. However, such conditions might represent SMBH growth during a brief period of time, a few tens of Myr, over which an SMBH is able to consume a large fraction of its disc emitting little radiation, possibly leaving behind rings of black holes and stars that formed but did not have time to migrate before the disc dissipated. If many black holes are ejected from the disc, if other stresses such as those from magnetic fields support the disc, or if global torques such as those associated with bars or spiral arms act to keep $Q \sim 1$ through the marginally stable disc region, our $\Delta_{\epsilon}$ figures could represent a modification to accretion efficiency valid over long periods of time. Even in these situations, the accretion efficiency could change by factors $\sim 0.5$, which would significantly ease the growth of very massive AGNs at high redshift.

We expect that our analysis is also applicable to accretion discs around more massive SMBHs than the $4 \times 10^6 M_{\odot}$ case that we consider. The key consideration is how the Type-I migration timescale varies with central object mass, $\tau \propto M_*$. For discs around less massive SMBHs, stars in the disc may not have time to exhaust their fuel before reaching the SMBH, being fully or partially disrupted, or undergoing stable Roche transfer (e.g. Dai & Blandford 2013; Metzger & Stone 2017), which would leave the accretion efficiency unchanged. Stellar-origin black holes, on the other hand, cannot undergo stable Roche transfer to the SMBH, and would spiral into the SMBH largely unaffected once gravitational radiation torques dominate (Yunes et al. 2011). However, growth via mergers with compact objects formed in accretion discs could potentially occur around more massive black holes without significant issue. Thus, although we do not think it is possible to grow a seed black hole around more massive black holes without significant issue. Thus, we consider only this change, the growth factor using $\eta = 0.1$ for the first 345 Myr after $z \sim 20$ is approximately $2.1 \times 10^6$. If instead $\epsilon = 0.1$ and $\alpha = 0.2$, then an SMBH can grow from $10^6$ to $10^9 M_{\odot}$ in only 78 Myr by accreting gas at $\sim 0.25$ times the Eddington rate, facilitating its growth by a factor of $1.2 \times 10^7$ over the $\sim 500$ Myr between $z \sim 20$ and $z \sim 7$. Thus, it is possible that SMBHs with masses $\gtrsim 10^9 M_{\odot}$ grew from $\sim 100 M_{\odot}$ black holes via approximately steady accretion. Similarly, starting with a more massive seed black hole, mergers with black holes formed in accretion discs allow SMBHs to form at high redshift without invoking constant accretion. Note that if black holes are left orbiting the SMBH at the end of an accretion episode, they will experience drag as their orbits cross gas during the next accretion episode and resume migration along with the next generation of black holes formed within the disc, analogous to the capture of stars into AGN discs discussed in Artymowicz et al. (1993).

4.1 Observables

Plentiful star formation in AGN discs has a number of other implications beyond growing SMBHs by $z \sim 7$. These include, among others: metallicity enrichment of the AGNs and surrounding galaxy (Artymowicz et al. 1993); extreme mass ratio inspiral (EMRI) production; and transients such as superluminous supernovae.

It is very difficult to measure the metallicities of the few AGNs at $z \gtrsim 6$ because most of the lines typically used to infer metallicity are redshifted to ranges difficult to observe from the ground. However, studies of quasar metallicity at $2.25 \lesssim z \lesssim 5.25$ have indicated that quasar broad-line region (BLR) metallicities do not vary with redshift, although correlations between SMBH mass and quasar metallicity have been observed (Xu et al. 2018). We note that if SMBH growth occurs significantly by accretion of compact objects formed in the disc, this AGN mass–metallicity relation follows naturally with little redshift dependence. Additionally, Xu et al. (2018) found that BLR metallicities were systematically larger than host galaxy metallicities by factors larger than can be explained by uncertainties in their measurements. The $z \sim 7.5$ quasar host galaxy J1342+0928 has a metallicity near solar (Novak et al. 2019), which would imply supersolar AGN metallicity. We suggest that AGNs should naturally follow a mass–metallicity relationship regardless of redshift, provided the AGN disc can support star formation and those stars have time to reach the end of their lives ($M_* \gtrsim 10^8 M_{\odot}$). The James Webb Space Telescope will facilitate direct metallicity measurement in $z \gtrsim 7$ AGNs.

Based on the accretion rates of black holes discussed in the previous section, we can make illustrative estimates again using the $\epsilon = 0.4, \alpha = 0.1$ model. Assuming that the black holes in the disc have masses $\sim 500 M_{\odot}$, we expect an EMRI rate of $\sim 10^{-4}$ yr$^{-1}$ per AGNs. Thus, space-based missions such as the Laser Interferometer Space Antenna (Amaro-Seoane et al. 2017) or the Chinese Taiji program (e.g. Ruan et al. 2018) could detect such events with regularity. If we assume that the eccentricities and inclinations of stellar-origin black holes continue to damp as they migrate through the inner regions of the disc, many of these EMRI events would occur with nearly zero eccentricity. This would distinguish them from other EMRI mechanisms, which are expected to have large eccentricities when in the $\sim 10^{-4}$–$10^{-2}$ Hz frequency range (Amaro-Seoane et al. 2007). Capture of stars into the disc may also occur (e.g. McKernan et al. 2011a, b), although we expect captured stars to have retrograde as well as prograde orbits, in contrast to the exclusively prograde orbits we expect from stars formed in situ (assuming minimal perturbations). For retrograde orbits, there exists the intriguing possibility of negative dynamical friction.
As stars born in AGN accretion discs exhaust their fuel, we expect violent transients. As supernovae occur, we expect significant changes in AGN light curves. Typical Type-II supernovae rise in luminosity over the course of months, a signal that could be evident when superimposed over a stochastic AGN light curve. It is difficult to spectrally identify supernovae in AGNs since Type-II supernovae generally have spectral lines that are similar to those in quasars. However, as the ejecta from these supernovae interact with the surrounding accretion disc, the flares could be significantly brighter and longer lived, along the lines of the transients observed at lower redshift in Graham et al. (2017). The rate of such flares in the light curves of high-$z$ AGN could provide insight into star formation processes in the disc. AGN discs also provide a natural birthplace for pair-instability supernovae, as accretion on to stars could make up for the significant mass-loss from stars above $\sim 50 M_\odot$. Such events will likely be observable at $z \gtrsim 7$ by the James Webb Space Telescope (Smidt et al. 2015). Additionally, pair-instability supernovae could account for why the flares noted in Graham et al. (2017) are more luminous than normal superluminous supernovae.

For the purposes of growing SMBHs at high redshift, it is pertinent to study near-Eddington accretion. However, there are multiple reasons that it is unlikely that the method for growing SMBHs described here will be similarly effective in local AGNs. Putting aside dependences on SMBH mass, the time-scale for type-I migration scales $\propto (H/r)^3$ (Tanaka, Takeuchi & Ward 2002), as does the time-scale for gas inflow due to viscous torques. Considering the simplified radiation-pressure supported disc models of Shakura & Sunyaev (1973), $H \propto M/M_{\text{Edd}}$. Thus, for accretion discs with lower dimensionless accretion rates, stars may be more likely to reach the SMBH before reaching the end of their lives. Additionally, from equation (6), it is clear that gap opening is more prevalent in thinner discs. If gap opening becomes the norm rather than an exception for migrating objects, we expect significant departures from the results presented here.

Finally, we can view our proposed mechanism in light of the Sołtan argument (Sołtan 1982), which connects the evolution of AGN luminosity and SMBH masses over cosmic time. Recent estimates (e.g. Marconi et al. 2004) suggest that SMBH mass is accreted with an average efficiency of $\eta \sim 0.04$--0.16, and that a typical AGN accretes at a fraction of $\sim 0.1$--1.7 times the nominal Eddington rate. Typical radiative efficiencies for pure gas accretion in this range of Eddington ratios are $\sim 0.06$ for slowly rotating black holes to $\sim 0.3$ for rapidly rotating holes. Our finding that accretion of compact objects is potentially significant suggests that the derived average efficiencies imply a gas accretion efficiency that is a few times larger than 0.04--0.16. Thus, we suggest that the spin parameters of SMBH could be systematically larger than normally inferred. Note that observational constraints on $\eta$ inform us that the smaller values we find for $\Delta$ cannot represent a significant fraction of SMBH growth in the Universe, although such discs may still exist infrequently. We note that lower radiative efficiencies can also be expected during super-Eddington accretion for other reasons (see e.g. Jiang, Stone & Davis 2019).

5 CONCLUSIONS

We find that mergers between SMBHs and objects formed in or captured into the accretion disc can facilitate SMBH growth significantly beyond the standard Eddington-limited rate, even if the luminosity is capped at Eddington. This is therefore a potential channel for the production of $M \gtrsim 10^9 M_\odot$ black holes by $z \sim 7$ from $\sim 100 M_\odot$ seeds.

As previously suggested, stars can help explain the observed early enhancement of metallicity in discs (Artymowicz et al. 1993), and we also suggest that in situ star formation can explain the observed luminous long time-scale transients in AGNs (Graham et al. 2017). This is also an interesting channel for EMRs, which would have high enough secondary masses to produce events that can be detected easily and characterized precisely.

There are numerous unknowns in the physics relevant to our treatment. Examples include the balance between accretion and mass-loss from stars, the torques in a mixed gas–stellar discs, and the thermal balance in discs that have as important elements both stars and accreting compact objects. None the less, compact objects in discs around AGNs are likely to play important roles in many aspects of disc structure, SMBH growth, and the production of gravitational wave sources.

ACKNOWLEDGEMENTS

The authors thank Nick Stone for many stimulating discussions on accretion disc models. The authors are also thankful to Brian Metzger, Yuri Levin, and Alberto Sesana for their comments on feedback, EMRs, migration, and many other remarks that enhanced the clarity of this paper. The authors are grateful to Saavik Ford and Barry McKernan for useful discussions on the migration and orbital dynamics of black holes embedded in AGN discs. We also thank the referee for thoughtful comments which improved the paper. MCM is grateful for the hospitality of Perimeter Institute, where part of this work was carried out. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development Canada and by the Province of Ontario through the Ontario Ministry of Economic Development, Job Creation and Trade.

REFERENCES

Amaro-Seoane P., Gair J. R., Freitag M., Miller M. C., Mandel I., Cutler C. J., Babak S., 2007, Class. Quantum Gravity, 24, R113
This paper has been typeset from a TeX/LaTeX file prepared by the author.